

Affective and Sensorimotor Components of Emotional Prosody Generation

Swann Pichon^{1,2} and Christian A. Kell³

¹Laboratory for Behavioral Neurology and Imaging of Cognition, Department of Neuroscience, Medical School, University of Geneva, 1211 Geneva 4, Switzerland, ²Swiss Center for Affective Sciences, University of Geneva, 1211 Geneva 4, Switzerland, and ³Brain Imaging Center and Department of Neurology, Goethe University, 60528 Frankfurt, Germany

Although advances have been made regarding how the brain perceives emotional prosody, the neural bases involved in the generation of affective prosody remain unclear and debated. Two models have been forged on the basis of clinical observations: a first model proposes that the right hemisphere sustains production and comprehension of emotional prosody, while a second model proposes that emotional prosody relies heavily on basal ganglia. Here, we tested their predictions in two functional magnetic resonance imaging experiments that used a cue-target paradigm, which allows distinguishing affective from sensorimotor aspects of emotional prosody generation. Both experiments show that when participants prepare for emotional prosody, bilateral ventral striatum is specifically activated and connected to temporal poles and anterior insula, regions in which lesions frequently cause dysprosody. The bilateral dorsal striatum is more sensitive to cognitive and motor aspects of emotional prosody preparation and production and is more strongly connected to the sensorimotor speech network compared with the ventral striatum. Right lateralization during increased prosodic processing is confined to the posterior superior temporal sulcus, a region previously associated with perception of emotional prosody. Our data thus provide physiological evidence supporting both models and suggest that bilateral basal ganglia are involved in modulating motor behavior as a function of affective state. Right lateralization of cortical regions mobilized for prosody control could point to efficient processing of slowly changing acoustic speech parameters in the ventral stream and thus identify sensorimotor processing as an important factor contributing to right lateralization of prosody.

Introduction

Prosody, the speech melody, corresponds to slow modulations of pitch, rhythm, stress, or loudness, and conveys linguistic information but also most of the information regarding speakers' intentions or emotional state. Although recent advances have been made regarding how the brain perceives emotional prosody, the neural bases involved in its generation remain unclear and debated. Two models have been proposed and forged on the basis of clinical observations: a first model proposes that the right hemisphere sustains production and comprehension of emotional prosody, while a second model proposes that emotional prosody relies heavily on basal ganglia (BG) function. Here, we aimed at testing predictions arising from these models using functional magnetic resonance imaging (fMRI).

The lateralization hypothesis considers that affective prosody is a lateralized and dominant function of the right hemisphere (Ross, 1981; Shapiro and Danly, 1985; Borod et al., 2002). Re-

garding comprehension, several fMRI studies have emphasized the role of a right-lateralized network involving the superior temporal sulcus (STS; BA22/42) and inferior frontal cortex (BA45/47), with no clear influence of emotional valence or category of emotion (for review, see Wildgruber et al., 2006; Kotz et al., 2006). On the production side, prosodic control during speech production has either been related to a right-lateralized (Riecker et al., 2002) or to a bilateral perisylvian network (Aziz-Zadeh et al., 2010), and right hemisphere damage often induces a monotonous speech pattern evoking "flat affect." Despite these observations, dysprosody is also found after left hemisphere damage and the lateralization theory remains debated (Baum and Pell, 1999; Kotz et al., 2003; Van Lancker et al., 2006; Ross and Monnot, 2008).

An alternative model proposes that BG play an essential role in the generation of prosody and also mediate emotional expression and comprehension (Cancelliere and Kertesz, 1990; Pell and Leonard, 2003; Van Lancker et al., 2006; Paulmann et al., 2011). Cancelliere and Kertesz (1990) found that stroke patients who presented clinically with dysprosody suffered most frequently from left or right BG lesions, followed by anterior temporal lobe and insula lesions. Moreover, dysprosody is frequently associated with the disruption of striatal dopaminergic innervation in Parkinson's disease (PD) patients (Caekebeke et al., 1991; Benke et al., 1998). Although PD patients speak monotonously, they are able to imitate emotional prosody reasonably well (Möbes et al., 2008). This suggests that BG play a prominent role in modulating the speech motor plan as a function of the affective state.

Received July 24, 2012; revised Nov. 16, 2012; accepted Nov. 23, 2012.

Author contributions: S.P. and C.A.K. designed research; S.P. and C.A.K. performed research; S.P. and C.A.K. analyzed data; S.P. and C.A.K. wrote the paper.

S.P. is supported by the Academic Society of Geneva (Fund Foremane) and by the Swiss Centre for Affective Sciences. S.P. would also like to thank Patrik Vuilleumier for his support. C.A.K. is supported by the Medical Faculty of Goethe University, Frankfurt. We thank Christiane Arnold, Marion Behrens, Sebastian Hoffer, Christian Keller, Maritza Darquea, and Frederic von Wegner for help.

Correspondence should be addressed to Dr. Swann Pichon, Laboratory for Behavioral Neurology and Imaging of Cognition, Department of Neuroscience, Medical School, University of Geneva, 1 rue Michel-Servet, 1211 Geneva 4, Switzerland. E-mail: swann.pichon@gmail.com.

DOI:10.1523/JNEUROSCI.3530-12.2013

Copyright © 2013 the authors 0270-6474/13/331640-11\$15.00/0

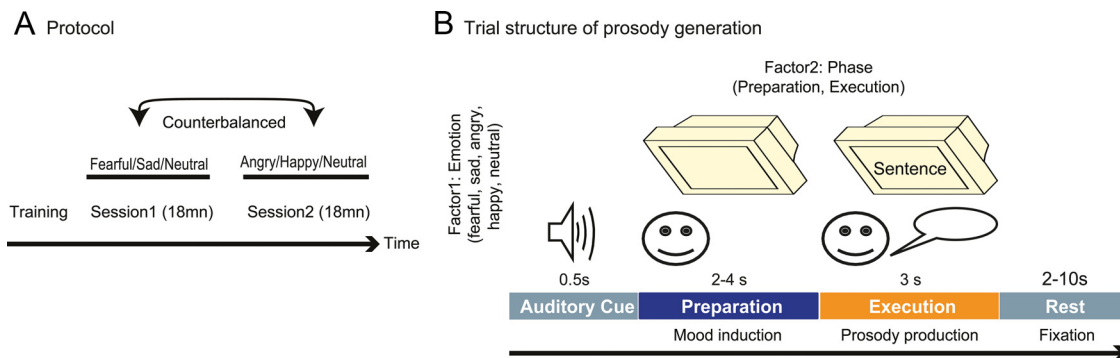


Figure 1. Experimental protocol. **A**, The protocol was composed of two successive sessions during which participants generated either fearful, sad, and neutral prosody or angry, happy, and neutral prosody. **B**, In each trial of prosody generation, participants were cued with the emotion to be produced and were given a varying preparatory time (preparation phase) for emotion induction, after which a semantically neutral sentence was displayed and was articulated according to the relevant emotional state (execution phase).

Anatomy indicates that the affective and motor components of prosody should be handled by two adjacent BG territories: dorsal striatum exchanges information with sensorimotor and associative cortices while ventral striatum is linked to the limbic system (Yelnik, 2008). We thus used fMRI in healthy participants to test which processing steps during emotional prosody generation recruit BG and the right hemisphere. We explored a cue-target paradigm that allows distinguishing the affective from the sensorimotor aspects of emotional prosody generation (see Materials and Methods). In a first experiment, participants were scanned while reading aloud sentences adopting either an emotional or a neutral prosody. In a second experiment, we tested the specificity of our findings and dissociated generation of emotional from generation of linguistic prosody by comparing happy with interrogative intonation.

Materials and Methods

Procedure of Experiment 1. During the first experiment (Fig. 1), subjects performed two sessions of a reading task adapted from Kell et al. (2011). In each trial, they read sentences aloud and modulated prosody according to a brief auditory word cue (fearful, sad, angry, happy or neutral, lasting ~500 ms), which preceded sentences by a variable delay (discrete uniform distribution) ranging from 2 to 4 s. The use of a jittered instruction delay allowed us to temporally dissociate speech preparation from speech execution (factor trial phase) and assess the influence of emotional versus neutral prosody (factor emotion) within each phase. Because during the preparation phase-specific speech planning was impossible, it reflected only intention to speak and induction of emotional state. All sensorimotor aspects of speech processing thus fall into the execution phase. The affect associated with the cue was randomized between trials. To avoid frequent trial-to-trial switching between the five emotional states, fearful and sad trials were pooled in the same session, separately from angry and happy trials that appeared together in another session. A neutral condition was included in each counterbalanced session. Manipulation of emotion yielded a total of six experimental conditions, each containing 30 trials (total duration per session 17 min 40 s). For each condition, we modeled the preparation and the execution phase separately, yielding a total of 12 experimental conditions. Thirty written semantically neutral German declarative sentences with identical syntactical structure were presented for 3 s each (e.g., “Grüne Marken kleben vorne auf dem Umschlag,” translated “Green stamps are attached to the front of the envelope”). Intertrial intervals were pseudorandomly generated using a discrete uniform distribution with a mean of 5.6 s, lower and upper bounds of 2 and 10 s, respectively. Before scanning, participants were familiarized with the experimental setting and practiced by reading aloud a training set of sentences that were not used in the main experiment. Training was repeated until participants’ performances were deemed satisfactory by the experimenter, and were recorded for later acoustic and perceptual analysis.

Participants. Twenty right-handed volunteers (nine females, mean age \pm SD: 26.5 years \pm 3.4; and nine males: 27.3 years \pm 2.5) without neurological or psychiatric history participated in a 3 T fMRI study. All provided written informed consent according to guidelines of the local research ethics committee and were paid for their participation.

fMRI data acquisition. Gradient-echo T_2^* -weighted transverse echo-planar images (EPI) with blood oxygenation level-dependent (BOLD) contrast were acquired with a 3 T Magnetom TIM Trio scanner (Siemens). Each volume contained 33 axial slices acquired in a sequential manner (TR/TE/flip angle = 2000 ms/30 ms/90°, FOV = 192 mm, resolution = 64°64, isotropic voxels size of (3 mm)³, distance factor 25%). We collected a total of 1060 functional volumes for each subject, as well as a high-resolution T_1 -weighted anatomical image (TR/TI/TE/flip angle = 2250 ms/900 ms/2.6 ms/9°, FOV = 256 mm, resolution = 256°256, slice thickness = 1.1 mm, 144 sagittal slices). We administered the behavioral protocol using Presentation software (Neurobehavioral Systems).

Acoustic and perceptual analysis of speech data from the training phase. The initial training phase was audio-recorded with a standard microphone plugged to a computer equipped with Adobe Audition. Pauses between utterances were cut out and sentences grouped according to the category (fearful, sad, angry, happy, neutral). Using PRAAT (Boersma, 2001; <http://www.fon.hum.uva.nl/praat>), we extracted the intensity contours of each sentence and obtained mean values by converting the sound pressure level measured in decibels (dB) into linear sound pressure values. After creating the mean for all sentences of one condition, values were reconverted to dB. To analyze how much a participant modulated his voice in terms of frequency we calculated the SD of each sentence’s fundamental frequency. The SDs from all sentences belonging to one condition were averaged. Statistics were performed using paired *t* tests in SPSS (SPSS Inc.). Averaged acoustic parameters were then used for contrast weighting with functional imaging group data (see below), solely for the analysis rendered in Figure 2. Finally, the audio speech samples were presented in randomized order to seven lay raters who judged the prosody of the sentences as neutral, fearful, sad, angry, or happy. Percentage correct answers for each prosody type were compared against each other using paired *t* tests in SPSS.

fMRI image processing. Image processing and statistical analyses were performed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; www.fil.ion.ucl.ac.uk/spm). Functional images were spatially realigned to the first volume by rigid body transformation, corrected for time differences in slice acquisition using the middle slice in time as reference, spatially normalized to the standard Montreal Neurological Institute EPI template, resampled to an isotropic voxel size of 2 mm, and spatially smoothed with an isotropic 8 mm full-width at half-maximum (FWHM) Gaussian kernel (Friston et al., 1995).

fMRI whole-brain analysis. We performed standard analyses using the general linear model (GLM) as implemented in SPM8, where event-related signal changes were modeled separately for each subject. For each session, we specified a linear model including seven event types. A first condition modeled transient effects related to the auditory cue (all cues

were pooled in a single regressor). Three other regressors (two emotions, one neutral condition) modeled the preparatory period as a function of the primed affect, while three additional regressors modeled the presentation of the sentence and its articulation by the subject. We decided to keep both neutral conditions apart because of possible influence by session-specific emotions. For each condition, a covariate was calculated by convolving delta functions with a canonical hemodynamic response function (HRF), over the duration of the phase (preparation or execution). As explained earlier, the use of a jittered cue allowed us to decorrelate preparation from execution and thus temporally dissociate them. Twelve additional covariates were included in the GLM to model movement-related artifacts (realignment parameters and their temporal derivatives). As each session encompassed a neutral condition, this yielded a total of 12 images of interest per subject.

For the whole-brain group level analysis, images of parameter estimates were entered in a repeated-measures ANOVA crossing the factors trial phase (preparation, execution) and emotion (fearful, sad, neutral1, angry, happy, neutral2). In this way, the variance estimates incorporated appropriately weighted within-subject and between-subject variance effects. Given the absence of emotion-specific effects, we averaged emotions together at the second level by weighting the contrasts preparation (emotion > neutral) and execution (emotion > neutral) accordingly (Table 1, Table 2, respectively). We preferred to keep the neutral baseline apart to estimate session-specific contrasts although no significant difference was detected when comparing both conditions against each other. For informative purposes, we also indicate in Table 1 and Table 2 which regions show a phase*emotion interaction pattern specific to preparation or execution of emotional prosody. For Table 1, this corresponds to a greater differential response between emotional and neutral trials during the preparation phase (execution phase for Table 2) as compared with the same difference in the execution phase (preparation phase for Table 2). A nonsphericity correction was applied to prevent inhomogeneity in variance of differences among pairs of conditions. Anatomical labeling was performed over the average T_1 -weighted brain of the group and all parametric maps were rendered on the single-subject T_1 -weighted brain available in SPM8 using MRIcron (Rorden et al., 2007). To identify brain regions that generally contribute to acoustic modulations underlying prosody, the emotion-specific regressors (execution phase) were weighted (in the group analysis) with the averaged acoustic parameters from prescanning recordings (Fig. 2). This contrast corresponds to a simple correlation between condition-specific first level β images and the pattern of acoustic parameters extracted from the training set of sentences. It is important to note that the correlation is not with the actual prosody of the sentences spoken during scanning, but with the prescanning recordings, which were different sentences, spoken in different conditions. All whole-brain statistical maps were corrected for multiple comparisons using standard Bonferroni correction ($p < 0.01$ familywise error rate; FWE).

This analysis was completed by region of interest (ROI) analyses targeting the amygdala and the striatum. The amygdala is well known for its involvement in processing emotional visual and vocal stimuli (Phillips et al., 1998; Fröhlich et al., 2012), notably fearful, angry, and happy expressions (Vuilleumier, 2005; Sergerie et al., 2008; Pichon et al., 2012a,b). Together with the hippocampus, they appear as good candidates for mood induction as both structures are known to be engaged during fear and happiness induction (Damasio et al., 2000). We first extracted β values using maximum probability maps of the left and right amygdala

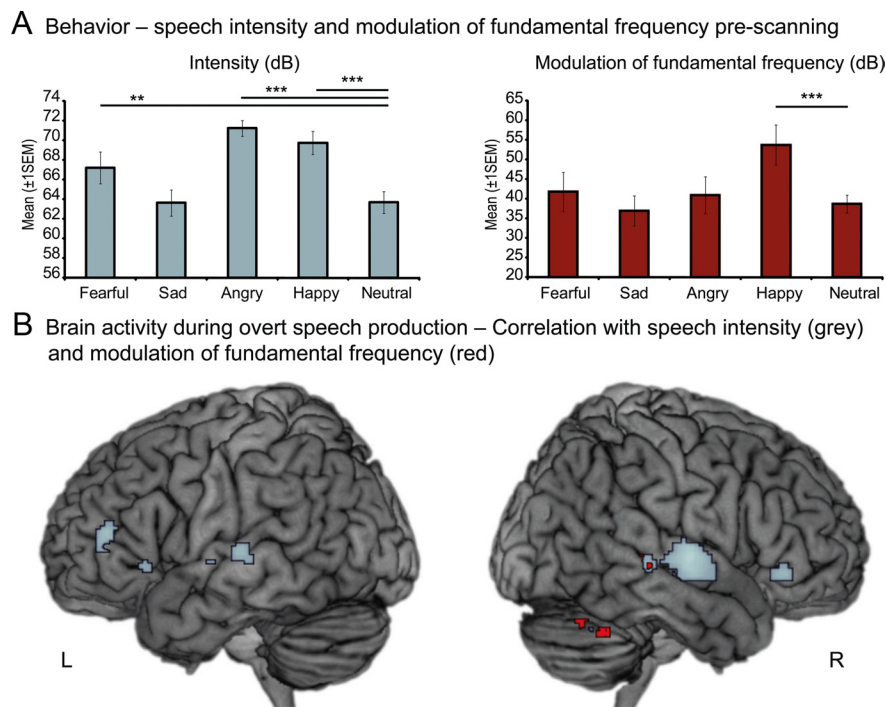


Figure 2. Acoustic speech parameters and their neural signatures. **A**, Results of the acoustic analysis of the speech data obtained in the training phase before the MRI experiment. Except for sad prosody, all other emotions were articulated more intensely than neutral prosody (gray). Variation of the fundamental frequency (F0, red) was especially high for happy prosody. **B**, A parametric contrast (weighted according to these parameters) revealed a similar signature in right posterior STS, overlapping for intensity (gray contrast displayed at $p < 0.01$ FWE) and F0 variation (red contrast, whose activations were significant at $p < 0.01$ FWE, rendered at $p < 0.05$ FWE for display purpose).

provided in the anatomy toolbox (Eickhoff et al., 2005). We then examined the degree of activation (extracted β values averaged across all voxels of each mask) evoked by each trial type using a repeated-measures ANOVA crossing the factors laterality, trial phase and emotion (α threshold at $p < 0.01$). The ROI analysis in the striatum aimed at studying whether the previously described dissociable loops through the BG (Alexander et al., 1990; Yelnik, 2008) manifested in functional striatal subdivisions or rather in a continuous dorsoventral gradient in limbic processing (see Fig. 6). We extracted, in each hemisphere and for each trial phase, contrast values for six voxels located between the dorsal (± 22 8 14) and the ventral striatum (± 22 8 –6), and separated each by 4 mm on the z-axis. These values were submitted to an ANOVA crossing the factors coordinate (1, 2, 3, 4, 5, 6), trial phase (preparation, execution), and laterality (left, right). We performed a specific contrast on the factor coordinate to test for the presence of a significant linear term (increase) along the z-axis (from dorsal to ventral).

Finally, to visualize the time courses of the ventral striatum (that we found markedly recruited during preparation for emotional prosody) as compared with time courses of the STS (that was recruited during both phases) we computed the group average time courses in both regions over a period of 14 s after the trial onset (cue presentation). For each subject, adjusted time courses were extracted using a 2 mm radius sphere centered at the group activation coordinate peaks in the right hemisphere. Given the absence of emotion-specific effects in our results, we averaged the time courses of all emotions together and also plotted the average time course during generation of neutral prosody.

fMRI lateralization analysis. We tested for lateralization of brain activity during emotional prosody preparation and production by comparing individual contrast images with their flipped counterparts. To account for minor spatial differences in exact localization of corresponding brain regions, we additionally smoothed images with an isotropic 6 mm FWHM Gaussian kernel (Kell et al., 2011). Results of the paired t tests are reported at $p < 0.01$, corrected for multiple comparisons at the voxel level (FWE) over the entire brain.

Table 1. Significant changes in brain activation when preparing emotional (vs neutral) prosody and results of lateralization analysis

	Anatomical region	Preparation (emotional > neutral) prosody				Phase (preparation) * emotion interaction				Lateralization			
		MNI coordinates				MNI coordinates				MNI coordinates			
		x	y	z	T value	x	y	z	T value	x	y	z	T value
L & R	Inferior frontal gyrus (dorsal BA44)	±54	14	30	5.77/5.25								
L & R	Inferior frontal sulcus	±44	18	32	5.31/5.32								
L & R	Pre-SMA	±6	4	64	7.13/6.53								
L & R	Pre-SMA	±8	18	50	7.36/6.56								
L & R	SMA	±8	−18	52	7.44/6.35								
L	Superior frontal sulcus	−24	24	52	5.74	±24	24	50	5.97/5.93				
L & R	Middle frontal gyrus	±30	0	54	7.39/6								
L	Middle frontal gyrus	−26	38	40	5.28								
L & R	Dorsal striatum	±16	10	2	6.98/6.65								
L & R	Ventral striatum	±22	8	−6	6.86/6.42								
L & R	Motor cortex	±28	−24	50	7.60/7.57								
L & R	Retrosplenial cortex/posterior cingulate	0	−58	16	8.08	0	−58	16	8.31				
L & R	Precuneus	±8	−42	52	7.11/7.46	2	−66	46	7.53				
R	Precuneus					6	−42	50	7.32				
L	Postcentral gyrus					−42	−28	60	5.59				
L & R	Superior parietal gyrus	±22	−64	54	6.22/7.63								
L & R	Superior parietal gyrus	±22	−64	54	7.11/7.63								
L & R	Inferior parietal cortex	±24	−72	36	6.87/7.31	±34	−76	40	7.36/7.7				
L & R	Supramarginal gyrus	±52	−24	28	6.19/6.73								
L & R	STS	±52	−52	4	6.50/5.57								
R	STS-horizontal segment	50	−62	12	6.94								
R	STS	58	−30	8	6.17								
L	Planum temporale	−38	−30	6	5.79								
R	Heschl's gyrus	38	−32	10	5.85								
L & R	Amygdala	Anatomical ROI											
L & R	Posterior hippocampus	±24	−38	−2	5.56/6.66								
R	Posterior Insula	38	−30	20	5.97								
L & R	Pulvinar	±10	−24	2	6.36/6.70								
L & R	Calcarine sulcus	±26	−64	4	5.72/5.85								
R	Calcarine sulcus	18	−86	4	5.63					14	−88	2	8.44
L & R	Parieto-occipital fissure	±16	−84	32	6.70/7.17								
R	Parieto-occipital fissure	18	−62	16	5.77								
L	Middle occipital gyrus	−24	−90	12	5.51								

$p < 0.01$ FWE whole-brain corrected.

Table 2. Significant changes in brain activation when executing emotional (vs neutral) prosody and results of lateralization analysis

R/L	Anatomical region	Execution (emotional > neutral) prosody				Phase (execution) * emotion interaction				Lateralization			
		MNI coordinates				MNI coordinates				MNI coordinates			
		x	y	z	T value	x	y	z	T value	x	y	z	T value
R	Inferior frontal gyrus (BA44)	58	14	6	6.67								
L & R	Inferior frontal gyrus (BA45/BA47)	±48	26	−6	8.9/8.92	48	34	2	5.09				
R	Anterior insula	28	30	−2	7.44								
L	Dorsal anterior cingulate cortex	−12	18	34	5.9								
R	Pre-SMA	6	6	66	5.25								
R	Hippocampus (anterior)	38	−2	−26	5.83								
L & R	Thalamus (ventrolateral) and globus pallidus	±10	−4	2	7.44/7.25								
L	SPT/TPJ	−50	−40	26	5.88								
L & R	STS	±50	−30	0	6.05/8.50					42	−26	−6	8.1
L & R	STS	±62	−22	4	6.64/7.59								
R	Middle temporal gyrus	50	−56	0	5.3								
L & R	Primary visual cortex	±18	−70	−10	5.42/5.08								
L & R	Cerebellum	±34	−56	−34	7.73/9.85								
	Cerebellum (vermis)	0	−64	−34	7.18								
L & R	Substantia nigra	±10	−20	−14	5.19/5.04								

$p < 0.01$ FWE whole-brain corrected.

Psychophysiological interactions analysis. We conducted psychophysiological interaction (PPI) analyses (Friston et al., 1997) to compare the connectivity of ventral and dorsal striatum with limbic and motor regions, during preparation and execution of emotional relative to neutral prosody. PPIs estimate effective connectivity via changes in inter-

regional covariance as a function of different experimental manipulations or tasks. First eigenvariate values were extracted for each participant from the filtered BOLD signal of the right and left ventral and dorsal striatum, separately. Regions were identified at the level of the group by contrasting the preparation for emotional (vs neutral) prosody

(see Table 1). Signal was extracted for each region from a 2 mm radius sphere centered at the group activation coordinate peak. Each time series was adjusted to exclude effects of no interest (movement parameters) and mean corrected. The regressors were deconvolved to obtain an estimate of the neural response, multiplied by the psychological context of interest (i.e., emotion vs neutral, either in the preparation or in the execution phase), and reconvolved using the canonical HRF to obtain a PPI regressor. According to the standards, we set up for each subject 8 GLMs ($2 \text{ laterality} \times 2 \text{ regions} \times 2 \text{ phases}$). Each contained the PPI regressor as well as 15 other regressors of no interest, including the time series of the seed region, the psychological factor convolved with the canonical HRF, a regressor modeling transient responses induced by the auditory cues, and 12 covariates modeling the movement parameters and their temporal derivatives. The model was then fitted to each brain voxel. At the group level, one-sample t tests over the PPI regressor of each model were used to reveal brain regions that showed significant increases in functional coupling with the seed region during preparation or production of emotional relative to neutral prosody. Consistent with the previous analysis, results were reported at $p < 0.01$ (FWE).

Using paired t tests, we also performed laterality analyses to compare the intrahemispheric and interhemispheric connectivity differences between (PPI maps of) the striatal seed regions in each hemisphere. We found none, even at more lenient thresholds ($p < 0.05$ FWE). These analyses were performed separately for the ventral or dorsal striatum, and separately for each trial phase. This required comparing the PPI map of a given seed region with the flipped version of the homolog region's PPI map in the other hemisphere for the same phase. As in the fMRI lateralization analysis, PPI contrast images were smoothed before performing statistical comparisons, using an isotropic 6 mm FWHM Gaussian kernel to account for minor spatial differences in exact localization of corresponding brain regions.

Procedure of Experiment 2. We ascertained whether our findings in ventral and dorsal striatum from Experiment 1 were specific for emotional prosody by comparing emotional (happy) and nonemotional linguistic (interrogative) prosody. We explored an existing dataset that had been acquired before Experiment 1 and that has partly been used for publication previously (Kell et al., 2011). While emotional and linguistic tasks are both known to engage superior temporal lobe, perceptual tasks requiring judgments on linguistic prosody engage other brain regions than affective judgments (Wildgruber et al., 2006). This experiment was initially designed to test if such differences could be detected between generation of emotional and nonemotional linguistic prosody. Twenty-six participants were scanned (13 females, mean age 29, range 19–44). Similar to Experiment 1, they were asked to prepare according to four different cues (factor 1): “happy,” “question,” “neutral,” and “covert” (i.e., neutral silent reading) before reading the sentence that was displayed after the jittered interval. fMRI data were analyzed separately for the preparation and execution phases (factor 2), producing eight conditions of interest. We applied the same methodological procedure as the one described in Experiment 1. At the group level, we performed a whole-brain analysis (thresholded at $p < 0.01$, FWE corrected) and ROI analyses using the coordinates of the ventral and dorsal striatum subpeaks isolated in Experiment 1. We extracted β values for each of the eight conditions of interest and each of the four (left and right, ventral and dorsal) striatal ROIs using a 2 mm radius sphere centered on each coordinate as estimated above (β values were averaged across all voxels of the sphere). We then performed *post hoc* tests between the emotional and linguistic conditions within each trial phase. Similarly to Experiment 1, we applied a Bonferroni correction for the number of ROIs ($n = 4$) used within each contrast and report adjusted p values thresholded at $p < 0.01$ corrected (corresponding to an α threshold of $p < 0.0025$). We also performed laterality analyses on these ROI data for the comparison of emotional and linguistic versus neutral prosody.

Results

Behavioral results

Participants in Experiment 1 produced perceptually distinguishable prosody: seven independent raters attributed speech samples from the training phase correctly to the neutral condition or one

of the emotions in 67% of trials, which is more than three times above chance level and congruent with previous literature (Scherer et al., 2003). Angry (76%) and happy (74%) prosody were most easily identified, while recognition scores were lower for sad (66%), neutral (64%), and fearful (54%) prosody (the latter significantly different from the others at $p < 0.01$), again replicating previous studies (Banse and Scherer, 1996; Wildgruber et al., 2005). Acoustic analysis of participants' speech production during the training phase confirmed that they adequately modulated speech acoustics to express emotions. Respective of the neutral condition, mean speech amplitude (intensity, Fig. 2A) was stronger when producing fearful, happy, or angry prosody ($t_{(19)} = 2.98/8.22/12.7$, respectively, adjusted $p < 0.01$). No intensity modulation was observed for sad prosody (adjusted $p = 0.48$). This is consistent with previous reports indicating that the speech amplitude is less pronounced for sad, fearful, and neutral prosody than for happy and angry prosody (Pfitzinger and Kaernbach, 2008). We observed no difference in the variability of fundamental frequency (SD of F0) for fearful, sad, or angry prosody ($t_{(19)} = 0.95/0.72/1.17$, adjusted $p = 1$). Only happy prosody differed from neutral ($t_{(19)} = 7.07$, adjusted $p < 0.001$; Fig. 2A), again reproducing previous results (Pell, 1999).

fMRI results of Experiment 1

Effects of emotion induction and prosody preparation (preparation phase)

During the preparation for emotional (vs neutral) prosody, we observed a strong activation of the bilateral dorsal and ventral striatum, consistent with the prediction that BG are important in the production of prosody (Fig. 3A; Cancelliere and Kertesz, 1990; Van Lancker Sidtis et al., 2006). As part of the limbic system, the bilateral amygdala (as confirmed by the ROI analysis below) and posterior insula increased their activity during emotion induction. Notably, the strongest activation was located medially in the retrosplenial and the posterior cingulate cortices. The retrosplenial cortex is densely connected to the hippocampal formation (Vann et al., 2009), which was also activated bilaterally. The involvement of this region during emotion induction is consistent with the role of the retrosplenial cortex in retrieving autobiographical memory and its involvement in processing self-related information (Maddock, 1999; Vann et al., 2009; Qin and Northoff, 2011).

In addition to emotion induction, preparation for emotional prosody also involved speech-related preparatory activity. Stronger activity for preparing for emotional than for neutral prosody was found in frontal and prefrontal motor-related cortices including left inferior frontal gyrus (dorsal BA44) and bilaterally in supplementary motor area (SMA), pre-SMA, and primary motor cortex. Also auditory regions pre-activated more strongly in the anticipation of auditory feedback, including the bilateral STS, left planum temporale, and right primary auditory cortex (Heschl's gyrus). Except for a cluster in the right primary visual cortex, no significant lateralization was found, even at lower threshold ($p < 0.05$ FWE). We also checked whether there were some emotion-specific patterns of responses by contrasting each emotion to all other emotions during the preparation phase but found none. Details of all activations are presented in Table 1.

In bilateral amygdala, the ROI analysis revealed no clear sign of lateralization. The ANOVA confirmed a main effect of emotion ($F_{(2,74,52)} = 10.32$, $p < 0.001$) with no other significant main effects or interactions [main effect of laterality ($p = 0.94$), main effect of phase ($p = 0.16$), laterality-by-emotion interaction ($p = 0.35$)], indicating that emotional prosody did not lateralize amygdala activity.

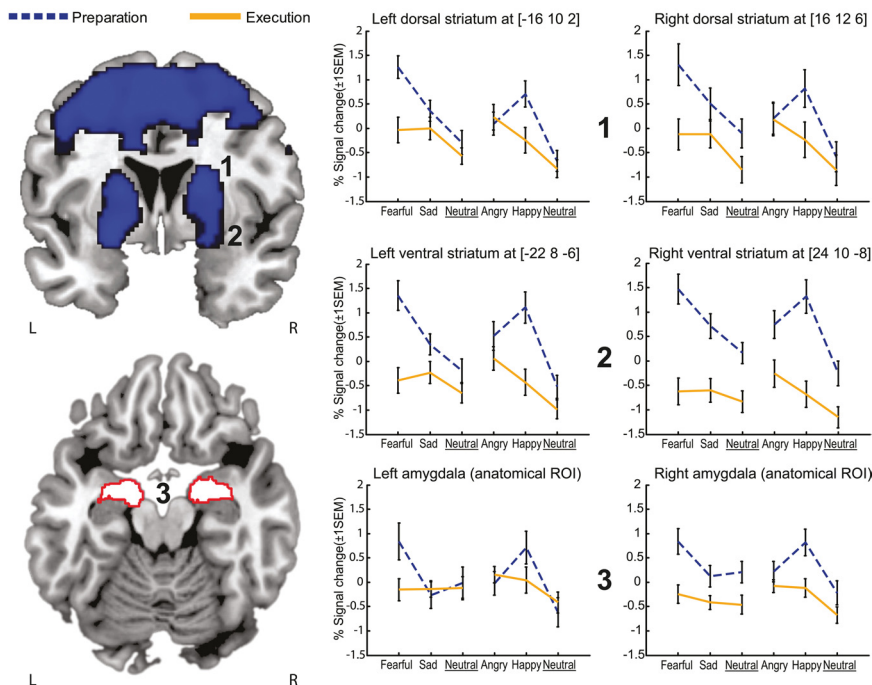
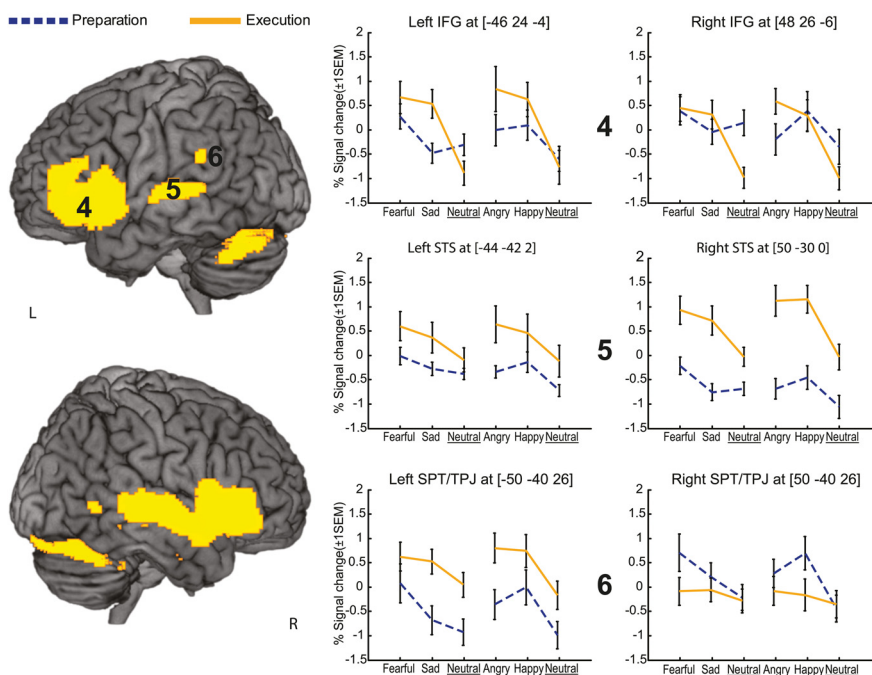
A Preparation of emotional (vs. neutral) speech ($P < 0.01$ FWE)**B** Execution of emotional (vs. neutral) speech ($P < 0.01$ FWE)

Figure 3. Affective and motor components of emotional prosody generation. **A**, (1–2) Preparation for emotional (vs neutral) prosody increased activity in bilateral dorsal (1) and ventral striatum (2), as well as limbic (retrosplenial, hippocampus, and amygdala) and sensorimotor cortices. Graphs on the right show mean group activation estimates (grand-mean centered) in all experimental conditions at the maximum peaks of each striatal cluster (see Materials and Methods). (3) Mean group activation estimates extracted from the amygdala using anatomical masks (displayed on the left), showing increased activity for fearful and happy prosody only during the preparation phase. **B**, Execution of emotional (vs neutral) prosody engaged a bilateral network involving the bilateral inferior frontal gyrus (IFG, 4), the STS (5), and the SPT/TPJ (6). Only the posterior STS showed right lateralization (see text). All activations were significant at $p < 0.01$ FWE.

Striatal connectivity during emotion induction (preparation phase)

We compared changes in effective connectivity between the ventral and the dorsal striatum and other brain regions while subjects

prepared for emotional (vs neutral) prosody (see Fig. 5; see Materials and Methods). The right ventral and dorsal striatum showed stronger ipsilateral connectivity between each other and with the anterior superior temporal gyrus, temporal pole, and right anterior insula. The right dorsal striatum additionally connected to the right Broca homolog (BA44) and to the contralateral orbitofrontal cortex and inferior temporal gyrus. Given that patients who suffer from lesions in these brain regions are dysprosodic, insular and anterior temporal integrity is necessary to generate or at least mediate information necessary for generation of emotional prosody (Cancelliere and Kertesz, 1990). Connectivity results are detailed in Table 3.

Effects of emotional prosody production (execution phase)

Production of emotional prosody compared with neutral speech production increased BOLD responses bilaterally in inferior frontal gyrus (Fig. 3B; BA44, BA45, and BA47, extending into anterior insula), superior cerebellum, thalamus and globus pallidus, substantia nigra, and STS with a much larger extent from anterior to posterior in the right hemisphere. Lateralization analyses revealed stronger activation of the posterior right (vs left) STS ($p < 0.01$ FWE). We observed no other lateralization, even at lower threshold ($p < 0.05$ FWE). Note that these results are consistent with the nonlateralized bilateral inferior prefrontal network observed in perception studies (Wildgruber et al., 2006; Ethofer et al., 2009). Interestingly, the left sylvian parietotemporal area (SPT/TPJ), as an important component of the dorsal stream, was more strongly activated during production of emotional than of neutral prosody, but no other regions belonging to this fast sensorimotor translating system (i.e., dorsal premotor cortex, dorsal Broca's region; Hickok and Poeppel, 2007) showed this activity pattern. Details of activations are presented in Table 2.

During speech production, activity in right STS ($xyz = [48 - 32 0]$, $t = 5.01$) and cerebellar vermis ($xyz = [2 - 64 - 34]$, $t = 5.44$) reflected modulation of fundamental frequency (Fig. 2B). This suggests that right lateralization of prosody-related activity is directly related to the acoustic features of the speech signal. Also condition-specific changes in the other studied acoustic parameter, speech intensity, affected right STS ($xyz = [48 - 32 0]$, $t = 5.95$), and cerebellar vermis activity ($xyz = [2 - 62 - 34]$, $t = 6.88$), but also activity in the bilateral inferior frontal gyrus (BA47; $xyz = [\pm 48 32 - 6]$, $t = 5.15/5.45$) and superior temporal

Table 3. Increased connectivity with right dorsal and ventral striatum during preparation or execution of emotional (vs neutral) prosody

R/L	Anatomical region	MNI coordinates			T value
		x	y	z	
Right dorsal striatum—preparation for emotional prosody					
L&R	Dorsal striatum	± 18	14	4	8.44/9.99
R	BA 44	60	18	4	8.79
L&R	Ventral striatum	± 20	− 2	− 6	9.31/10.18
R	Superior temporal gyrus (anterior)/temporal pole	56	4	− 14	10.7
L	Inferior temporal gyrus	− 42	− 6	− 36	8.27
L	Orbitofrontal cortex	− 24	30	− 16	8.42
R	Anterior insula	38	14	0	9.12
Right ventral striatum—preparation for emotional prosody					
L&R	Ventral striatum	± 24	8	− 8	5.03/6.16
R	Superior temporal gyrus (anterior)/temporal pole	50	12	− 22	6.12
R	Anterior insula	38	10	− 12	5.06
Right dorsal striatum— execution of emotional prosody					
L&R	Amygdala	± 26	2	− 14	8.79/12.98
R	Hippocampus	28	− 10	− 12	8.96
R	Primary auditory cortex	56	− 12	2	10.06
R	Caudate nucleus	18	12	10	9.76
L&R	Primary motor cortex	± 52	− 6	34	8.84/9.47
L	Occipitotemporal cortex	− 24	− 84	− 12	9.49
L&R	Primary visual cortex	± 18	− 100	− 4	8.17/8.44
Right ventral striatum— execution of emotional prosody					
R	Dorsal striatum	24	10	− 10	5.58
R	Ventral striatum	20	12	4	5.1

$p < 0.01$ FWE whole-brain corrected.

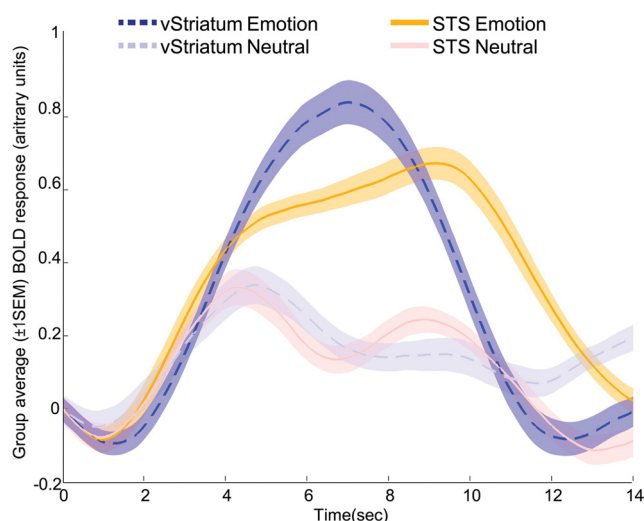


Figure 4. Group average time courses illustrating the response of the right ventral striatum (vStriatum, in blue) and right STS (in orange) during generation of emotional (dark color) and neutral (light color) prosody. Note that the BOLD response in the ventral striatum decreases ~3 s earlier than in the STS, confirming its primary involvement in emotion induction during the preparation phase.

gyrus ($xyz = [\pm 64 -12 2]$, $t = 5.04/6.28$). The time courses of BOLD responses in ventral striatum and STS (Fig. 4) confirmed an involvement of the limbic region during preparation only while STS activated during preparation and actual production.

Striatal connectivity during production of emotional prosody (execution phase)

The right ventral striatum mainly connected to the dorsal striatum (Fig. 5). The dorsal striatum in turn was strongly connected to the bilateral articulatory motor and right auditory cortices but was also connected with limbic structures such as the right anterior hippocampus and bilateral amygdala. Connectivity results are fully detailed in Table 3.

Striatal gradient in affective processing

We tested whether there were distinguishable motor-related dorsal and limbic ventral striatal subdivisions. We hypothesized that if BG loops were spatially segregated (Alexander et al., 1990), we should observe separable activations in the dorsal versus ventral striatum during emotion induction only. Nevertheless, the ANOVA revealed a significant main effect of coordinate ($p < 0.001$) characterized by a significant linear increase ($p < 0.001$) of bilateral striatal preparatory activity from dorsal to ventral striatum (Fig. 6). This linear increase disappeared during prosody production ($p > 0.2$). There was also a laterality-by-coordinate interaction ($p = 0.008$), driven by a marginally stronger response of the left (vs right) ventral striatum ($p = 0.13$) while no difference was present in dorsal striatum ($p = 0.4$).

fMRI results of Experiment 2

Specificity of effects related to emotional prosody

We controlled in 26 additional subjects whether the activity pattern observed in Experiment 1 was-specific to emotional prosody or could also be detected for linguistic prosody. We used an existing dataset that had been acquired before Experiment 1 that has partly been used for publication previously (Kell et al., 2011). The whole-brain analysis showed that the ventral striatum was again more active during preparation for happy than neutral prosody (Kell et al., 2011), and dissociated preparation for emotional from preparation for linguistic prosody ($p < 0.01$ FWE), while no difference was detected during the execution of emotional versus linguistic prosody. We also reproduced our results for the contrast comparing the execution of happy versus neutral prosody except for the activation in dorsal ventral premotor cortex (which activated at $p < 0.001$, uncorrected).

The ROI analyses (Fig. 7) revealed that although the bilateral ventral striatum was also slightly pre-activated for linguistic and neutral prosody (with no significant difference between neutral and linguistic prosody, adjusted $p > 0.01$) it was mostly involved in the preparation for happy prosody (preparation for happy >

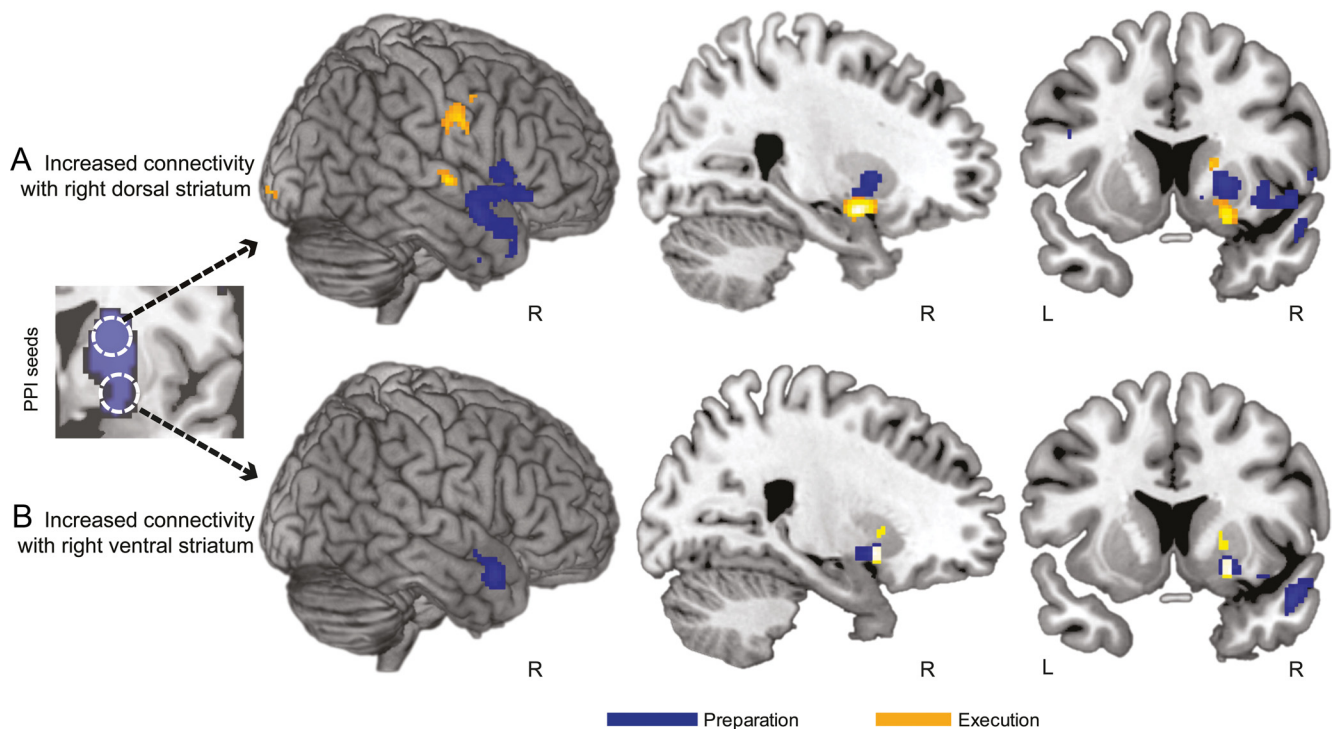


Figure 5. Connectivity of the right striatum specific to emotional prosody ($p < 0.01$ FWE). **A**, During preparation, both the dorsal (**A**) and ventral striatum (**B**) increased their connectivity with the temporal pole and the anterior insula. During the execution phase, the dorsal striatum was more connected to the auditory and motor cortices, as well as limbic regions (amygdala and hippocampus, same conventions as in Fig. 3).

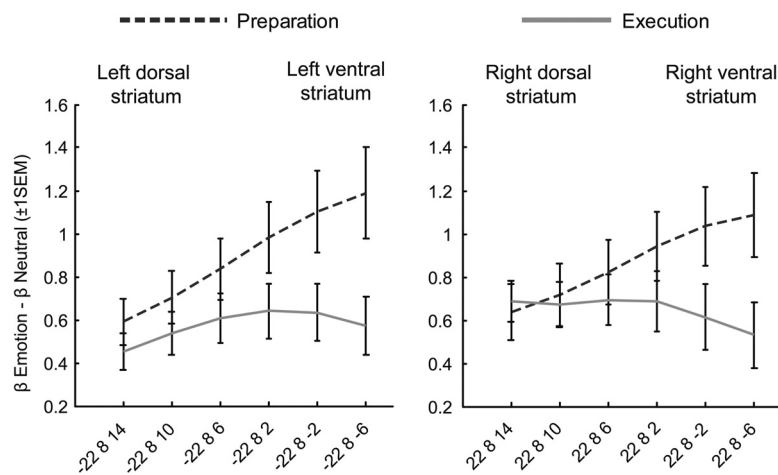


Figure 6. Preparatory striatal response to emotion induction increases linearly from dorsal to ventral striatum. These values were extracted from 6 voxels located between the dorsal ($\pm 22 8 14$) and the ventral striatum ($\pm 22 8 -6$), and separated each by 4 mm on the z-axis.

linguistic prosody: left: $t_{(25)} = 3.99$, adjusted $p = 0.002$; right: $t_{(25)} = 3.63$, adjusted $p = 0.005$). The bilateral dorsal striatum was equally pre-activated for emotional and linguistic prosody (adjusted $p > 0.16$), suggesting a more cognitive role of the dorsal compared with the ventral striatum. During production of happy prosody, the bilateral ventral and dorsal striatum did not increase activity compared with linguistic baseline ($p > 0.01$). Again, we observed no striatal lateralization ($p > 0.15$).

Discussion

The present study investigated the neural substrates involved in the generation of emotional prosody. We dissociated affective

from sensorimotor aspects of this speech behavior to test the assumptions of two existing mostly lesion-based models in emotional prosody generation: a first model which emphasizes the role of BG, together with anterior temporal regions and insula (Cancelliere and Kertesz, 1990), and a second model proposing that emotional prosody is a dominant function of the right hemisphere (Ross, 1981; Shapiro and Danly, 1985; Borod et al., 2002). We found support for both models as we show a strong bilateral involvement of BG, limbic regions, temporal pole, and anterior insula during preparation for emotional prosody and right-lateralized auditory feedback-related processing during actual production of prosodic speech.

Our results support the BG model, as this structure appears particularly involved in the early phase of emotional prosody generation preceding linguistic or sensorimotor processing. Although we did not find support in favor of completely segregated BG loops, we attribute different roles to the ventral and dorsal parts of the striatum: while the ventral striatum is specifically related to emotion induction and not to linguistic or motor aspects of emotional prosody generation, the dorsal striatum is particularly sensitive to the latter. The limbic activity observed during the preparation phase in ventral striatum and amygdala is accompanied by processing in bilateral retrosplenial cortex, hippocampi, and anterior temporal poles, which in turn could be interpreted as self-referential semantic processing subserving emotion induction. This network has been

consistently found in tasks requiring retrieval of autobiographical experience, processing of emotional stimuli and episodic memory (Maddock, 1999; Frith and Frith, 2003; Vann et al., 2009). Lesions of this network (temporal pole and anterior insula) have been identified as the second most frequent cause of dysprosody after BG lesions (Cancelliere and Kertesz, 1990). Note also that we cannot exclude the alternative explanation that the ventral striatum activation may reflect the mismatch of reading a neutral sentence aloud with a strong emotion rather than mood induction per se. Participants may have experienced the emotion condition as motivating and funny, which may explain why we failed to observe significant emotion-specific effects. We also show that the ventral striatum does not lateralize when preparing for emotional prosody and is functionally connected via the dorsal striatum to a bilateral fronto-insular network that likely codes task rules (Dosenbach et al., 2006). Importantly, we did not find increased connectivity with attentional networks that would have been expected if the emotional tasks were more difficult than the neutral condition. This does not exclude entirely the possibility that difficulty modulated responses in the observed networks. We assume nevertheless that training the participants considerably attenuated difficulty effects.

The ventral striatum receives information from a considerable array of limbic structures (Cardinal et al., 2002). It is tightly linked to motivation, reward processes, and pleasure (Kringelbach and Berridge, 2009). It has also been related to recognition and expression of affects. Patients with lesion in the ventral striatum (but not dorsal parts of BG) show deficits in emotion recognition (whether expressed through the face or voice) as well as increased or decreased emotional experiences (Calder et al., 2004). Note that although we lack the sufficient spatial resolution to identify with precision which BG nuclei are involved during emotion induction, the cluster's maximum appears centered on the ventral pallidum rather than the nucleus accumbens. Ventral pallidum, which is part of the limbic BG territory (Parent and Hazrati, 1995; Yelnik, 2008; Smith et al., 2009), receives inputs from other limbic regions including nucleus accumbens, orbito-frontal cortex, amygdala, and lateral hypothalamus. It connects to the brainstem and to the prefrontal cortex indirectly via the thalamus (Smith et al., 2009). The ventral pallidum thus likely plays a considerable role in translating limbic signals into motor output (Mogenson and Yang, 1991).

Furthermore, alteration of BG dopaminergic transmission in PD patients has often been associated with changes in emotional experience (both in terms of subjective experience and physiological arousal) and impaired recognition of emotions conveyed by faces or voices (for review, see Péron et al., 2012). Note that deficits in emotion recognition are not specific to PD and have also been reported in other pathologies involving disturbed dopaminergic functioning such as autism, Huntington's disease, or schizophrenia (Edwards et al., 2002; Johnson et al., 2007; Harms

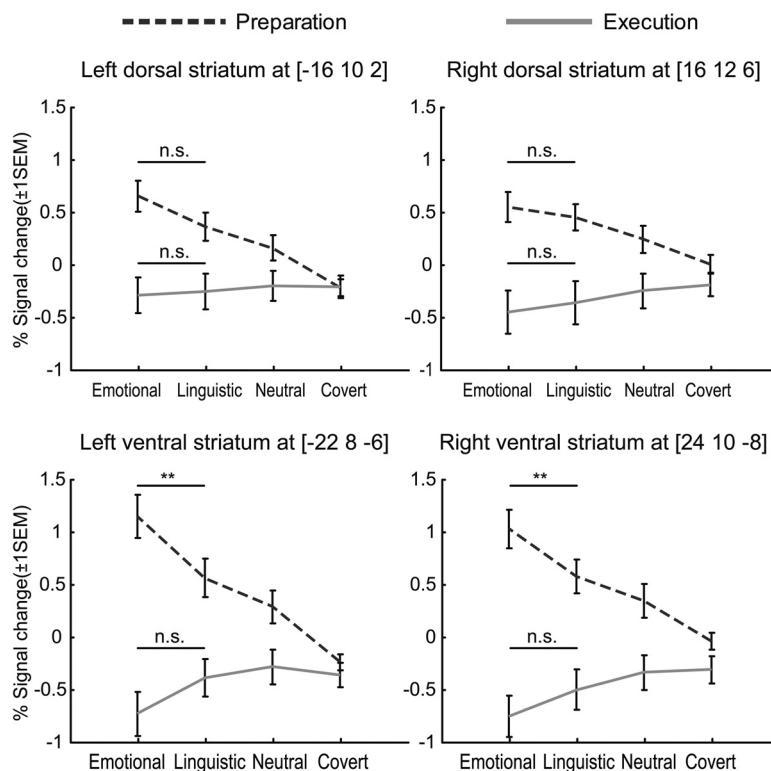


Figure 7. Control experiment assessing the specificity of striatal involvement for emotional prosody. Task parameters were identical to those of the main experiment except that subjects ($n = 26$) were asked to read sentences with emotional (happy) prosody, nonemotional linguistic (interrogative) prosody, with neutral prosody or covertly (neutral intonation). During the preparation phase, involvement of the ventral striatum was stronger for emotional prosody than for linguistic or neutral prosody. $**p < 0.01$.

et al., 2010). In PD, deficits in spontaneous expression of emotional prosody arise independently from affective state and without alteration of semantic understanding or ability to imitate emotional prosody when an external model is available (Blonder et al., 1989; Smith et al., 1996; Simons et al., 2003; Möbes et al., 2008). It indicates that the disease cannot be regarded solely as purely motor, but involves additional deficits at the interface between the limbic and the motor system. Our data indicate that a dysfunctional interplay between ventral and dorsal striatum could underlie PD patients' dysprosody.

Once emotion is induced and speech material available for articulation (execution phase), subcortical integration of affective with sensorimotor speech processing seems to center around the dorsal striatum. The cortical network engaged in sensorimotor processing underlying the slow prosodic speech modulations involves bilateral inferior frontal gyri and superior temporal gyri and sulci, as well as superior cerebellum. This network, together with BG, has previously been associated with control of speech rhythm and speed during syllable repetitions (Ackermann and Riecker, 2010). Interestingly, it is only in this sensorimotor network that we found support for the right-lateralization model of prosody. Right lateralization in the entire system was confined to right posterior STS when participants produced emotionally charged speech. The right STS receives direct input from auditory cortex, is integrated in the ventral speech processing stream (Hickok and Poeppel, 2007), and is sensitive to prosodic speech features during speech perception (Wildgruber et al., 2006). Strikingly, it is this brain region (together with the cerebellum) that is consistently sensitive to acoustic speech features in our study and a previous perceptual study (Wiethoff et al., 2008). It is

thus conceivable that this brain region monitors slow (prosodic) modulations of auditory feedback during ongoing speech production. This suggests that previously reported right lateralization of prosody processing (be it frontal or temporal) may relate more to sensory features of the acoustic signal rather than to emotional valence. Please note that in addition to intensity and modulation of fundamental frequency, other correlated acoustic speech parameters could contribute to this right lateralization as hidden variables.

One attractive model that could explain the right-hemisphere advantage of prosody control is the double filtering by frequency hypothesis (Ivry and Robertson, 1999) and its expansion in the speech domain, the asymmetric sampling in time hypothesis (Poeppel, 2003): both models propose an efficient processing of relative lower frequencies in the right hemisphere (suprasegmental processing) while the left hemisphere would have an advantage for the processing of relative higher frequencies (segmental processing). The slow modulations in fundamental frequency and intensity underlying prosody could thus be efficiently decoded in the right hemisphere while the left hemisphere succeeds better in phonematic processing that in turn relies more heavily on higher relative frequencies of the speech signal. This would suggest that during speech production, the right hemisphere could provide auditory feedback regarding the relatively slower modulations of the auditory stream (i.e., prosody) while relatively faster modulations of the speech signal are fed back from the left auditory cortex. Consistent with this model, we observed activation of the right-lateralized posterior STS during prosody production. Yet, in addition to this activity in the right ventral speech processing stream, prosody also recruited left hemispheric ventral speech regions including the superior temporal gyrus and sulcus extending into the insula and ventral portions of the inferior frontal gyrus. The bilateral ventral speech stream is primarily thought to map auditory speech representations onto lexical conceptual representations for speech comprehension (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Our data indicate that during higher prosodic demands, the bilateral ventral streams become engaged in feedback processing (Hickok, 2012). We therefore hypothesize that both ventral streams are able to integrate relatively slower modulations of the speech signal that characterize prosody. Nevertheless, right lateralization of the ventral streams is observed during increased processing due to prosodic control. This could result from that fact that during speech production the left hemisphere is primarily engaged in fast sensorimotor processing in the left-lateralized dorsal stream, which could map rapidly changing (high-frequency) properties of the acoustic speech signals to frontal lobe articulatory networks (Hickok and Poeppel, 2007). Supposing parallel processing in the two hemispheres, the right ventral stream could have consequently become specialized in analyzing the slow modulations of the acoustic speech spectrum that underlie prosody and in mapping these sensory representations to motor representations. It is interesting to note that increased prosodic processing does not further mobilize the left-lateralized dorsal stream except for its proposed input region, area (SPT)/(TPJ). This region could potentially contribute to sensorimotor processing in the ventral stream in addition to its known involvement in sensorimotor mapping within the dorsal stream. Yet, given that right SPT/TPJ is not engaged during prosody generation, its activity does not seem to constitute a prerequisite for ventral stream processing.

Conclusion

Our findings support both the BG and the lateralization hypothesis, yet for different aspects of prosody processing. Emotional prosody generation involves affective components that are linked to bilateral ventral striatum and other limbic and paralimbic structures. On the other hand, dorsal striatum is more sensitive to cognitive and motor aspects of emotional prosody preparation and production and is more strongly connected to the sensorimotor speech network compared with the ventral striatum. Right lateralization is limited to the right STS in which activity reflects acoustic speech parameters, suggesting that it is sensorimotor and not affective processing that right lateralizes prosodic computations. Our results also suggest that dysprosodic syndromes could arise from two different lesions. First, a first lesion that perturbs a limbic-semantic network, which comprises anterior insular and temporal cortices connected to limbic (ventral) portions of the BG, and that would yield an emotional prosodic production deficit on the basis of disturbed integration of affect into motor routines. A second type of lesion involving a right-lateralized sensorimotor network more strongly connected to dorsal striatum would lead to dysprosody on the basis of perturbed sensorimotor processing of relatively slower modulations of the speech signal.

References

- Ackermann H, Riecker A (2010) The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct Funct* 214:419–433. [CrossRef Medline](#)
- Alexander GE, Crutcher MD, DeLong MR (1990) Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, “pre-frontal” and “limbic” functions. *Prog Brain Res* 85:119–146. [Medline](#)
- Aziz-Zadeh L, Sheng T, Gheytanchi A (2010) Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PloS one* 5:e8759. [CrossRef Medline](#)
- Banase R, Scherer KR (1996) Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol* 70:614–636. [CrossRef Medline](#)
- Baum S, Pell M (1999) The neural bases of prosody: insights from lesion studies and neuroimaging. *Aphasiology* 13:581–608. [CrossRef](#)
- Benke T, Bösch S, Andree B (1998) A Study of Emotional Processing in Parkinson's Disease. *Brain Cogn* 38:36–52. [CrossRef Medline](#)
- Blonder LX, Gur RE, Gur RC (1989) The effects of right and left hemiparkinsonism on prosody. *Brain Lang* 36:193–207. [CrossRef Medline](#)
- Boersma P (2001) Praat, a system for doing phonetics by computer. *Glott International* 5:341–345.
- Borod JC, Bloom RL, Brickman AM, Nakhutina L, Curko EA (2002) Emotional processing deficits in individuals with unilateral brain damage. *Appl Neuropsychol* 9:23–36. [CrossRef Medline](#)
- Caekebeke JF, Jennekens-Schinkel A, van der Linden ME, Buruma OJ, Roos RA (1991) The interpretation of dysprosody in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 54:145–148. [CrossRef Medline](#)
- Calder AJ, Keane J, Lawrence AD, Manes F (2004) Impaired recognition of anger following damage to the ventral striatum. *Brain* 127:1958–1969. [CrossRef Medline](#)
- Cancelliere AE, Kertesz A (1990) Lesion localization in acquired deficits of emotional expression and comprehension. *Brain Cogn* 13:133–147. [CrossRef Medline](#)
- Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002) Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 26:321–352. [CrossRef Medline](#)
- Damasio AR, Grabowski TJ, Bechara A, Damasio H, Ponto LL, Parvizi J, Hichwa RD (2000) Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat Neurosci* 3:1049–1056. [CrossRef Medline](#)
- Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE (2006) A core system for the implementation of task sets. *Neuron* 50:799–812. [CrossRef Medline](#)
- Edwards J, Jackson HJ, Pattison PE (2002) Emotion recognition via facial

- expression and affective prosody in schizophrenia: a methodological review. *Clin Psychol Rev* 22:789–832. [CrossRef Medline](#)
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25:1325–1335. [CrossRef Medline](#)
- Ethofer T, Kreifelts B, Wiethoff S, Wolf J, Grodd W, Vuilleumier P, Wildgruber D (2009) Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *J Cogn Neurosci* 21:1255–1268. [CrossRef Medline](#)
- Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210.
- Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6:218–229. [CrossRef Medline](#)
- Frith U, Frith CD (2003) Development and neurophysiology of mentalizing. *Philos Trans R Soc London B Biol Sci* 358:459–473. [CrossRef Medline](#)
- Frühholz S, Ceravolo L, Grandjean D (2012) Specific Brain Networks during Explicit and Implicit Decoding of Emotional Prosody. *Cereb Cortex* 22:1107–1117. [CrossRef Medline](#)
- Harms MB, Martin A, Wallace GL (2010) Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychol Rev* 20:290–322. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402. [CrossRef Medline](#)
- Ivry RB, Robertson LC (1999) The two sides of perception. Cambridge: MIT.
- Johnson SA, Stout JC, Solomon AC, Langbehn DR, Aylward EH, Cruce CB, Ross CA, Nance M, Kayson E, Julian-Baros E, Hayden MR, Kieburz K, Guttman M, Oakes D, Shoulson I, Beglinger L, Duff K, Penziner E, Paulsen JS (2007) Beyond disgust: impaired recognition of negative emotions prior to diagnosis in Huntington's disease. *Brain* 130:1732–1744. [CrossRef Medline](#)
- Kell CA, Morillon B, Kouneiher F, Giraud AL (2011) Lateralization of speech production starts in sensory cortices—a possible sensory origin of cerebral left dominance for speech. *Cereb Cortex* 21:932–937. [CrossRef Medline](#)
- Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, Friederici AD (2003) On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang* 86:366–376. [CrossRef Medline](#)
- Kotz SA, Meyer M, Paulmann S (2006) Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Prog Brain Res* 156:285–294. [CrossRef Medline](#)
- Kringelbach ML, Berridge KC (2009) Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn Sci* 13:479–487. [CrossRef Medline](#)
- Maddock RJ (1999) The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci* 22:310–316. [CrossRef Medline](#)
- Möbes J, Joppich G, Stiebritz F, Dengler R, Schröder C (2008) Emotional speech in Parkinson's disease. *Mov Disord* 23:824–829. [CrossRef Medline](#)
- Mogenson GJ, Yang CR (1991) The contribution of basal forebrain to limbic-motor integration and the mediation of motivation to action. *Adv Exp Med Biol* 295:267–290. [CrossRef Medline](#)
- Parent A, Hazrati LN (1995) Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Res Brain Res Rev* 20:128–154. [Medline](#)
- Paulmann S, Ott DV, Kotz SA (2011) Emotional speech perception unfolding in time: the role of the basal ganglia. *PLoS ONE* 6:e17694. [CrossRef Medline](#)
- Pell MD (1999) Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere-damaged speakers. *Brain Lang* 69:161–192. [CrossRef Medline](#)
- Pell MD, Leonard CL (2003) Processing emotional tone from speech in Parkinson's disease: a role for the basal ganglia. *Cogn Affect Behav Neurosci* 3:275–288. [CrossRef Medline](#)
- Péron J, Dondaine T, Le Jeune F, Grandjean D, Vèrin M (2012) Emotional processing in Parkinson's disease: a systematic review. *Mov Disord* 27:186–199. [CrossRef Medline](#)
- Pfützinger HR, Kaernbach C (2008) Amplitude and amplitude variation of emotional speech. In *Proc. International Conference of the Interspeech*, 2008, Brisbane, Australia, p 1036–1039.
- Phillips ML, Young AW, Scott SK, Calder AJ, Andrew C, Giampietro V, Williams SC, Bullmore ET, Brammer M, Gray JA (1998) Neural responses to facial and vocal expressions of fear and disgust. *Proc Biol Sci* 265:1809–1817. [CrossRef Medline](#)
- Pichon S, de Gelder B, Grèzes J (2012a) Threat prompts defensive brain responses independently of attentional control. *Cereb Cortex* 22:274–285. [CrossRef Medline](#)
- Pichon S, Rieger SW, Vuilleumier P (2012b) Persistent affective biases in human amygdala response following implicit priming with negative emotion concepts. *Neuroimage* 62:1610–1621. [CrossRef Medline](#)
- Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time' 41:245–255.
- Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? *Neuroimage* 57:1221–1233. [CrossRef Medline](#)
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724. [CrossRef Medline](#)
- Riecker A, Wildgruber D, Dogil G, Grodd W, Ackermann H (2002) Hemispheric lateralization effects of rhythm implementation during syllable repetitions: an fMRI study. *Neuroimage* 16:169–176. [CrossRef Medline](#)
- Rorden C, Karnath HO, Bonilha L (2007) Improving lesion-symptom mapping. *J Cogn Neurosci* 19:1081–1088. [CrossRef Medline](#)
- Ross ED (1981) The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol* 38:561–569. [CrossRef Medline](#)
- Ross ED, Monnot M (2008) Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain Lang* 104:51–74. [CrossRef Medline](#)
- Scherer KR, Johnstone T, Klasmeyer G (2003) Vocal expression of emotion. In: *Handbook of affective sciences* (Davidson RJ, Scherer KR, Goldsmith HH eds), pp 433–456. New York: Oxford UP.
- Sergerie K, Chochol C, Armony JL (2008) The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 32:811–830. [Medline](#)
- Shapiro BE, Danly M (1985) The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain Lang* 25:19–36. [CrossRef Medline](#)
- Simons G, Ellgring H, Smith Pasqualini M (2003) Disturbance of spontaneous and posed facial expressions in Parkinson's disease. *Cogn Emot* 17:759–778. [CrossRef](#)
- Smith KS, Tindell AJ, Aldridge JW, Berridge KC (2009) Ventral pallidum roles in reward and motivation. *Behav Brain Res* 196:155–167. [CrossRef Medline](#)
- Smith MC, Smith MK, Ellgring H (1996) Spontaneous and posed facial expression in Parkinson's disease. *J Int Neuropsychol Soc* 2:383–391. [CrossRef Medline](#)
- Van Lancker Sidsis D, Pachana N, Cummings JL, Sidsis JJ (2006) Dysprosodic speech following basal ganglia insult: toward a conceptual framework for the study of the cerebral representation of prosody. *Brain Lang* 97:135–153. [CrossRef Medline](#)
- Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? *Nat Rev Neurosci* 10:792–802. [CrossRef Medline](#)
- Vuilleumier P (2005) How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci* 9:585–594. [CrossRef Medline](#)
- Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T (2008) Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. *Neuroimage* 39:885–893. [CrossRef Medline](#)
- Wildgruber D, Riecker A, Hertrich I, Erb M, Grodd W, Ethofer T, Ackermann H (2005) Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24:1233–1241. [CrossRef Medline](#)
- Wildgruber D, Ackermann H, Kreifelts B, Ethofer T (2006) Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog Brain Res* 156:249–268. [CrossRef Medline](#)
- Yelnik J (2008) Modeling the organization of the basal ganglia. *Rev Neurol* 164:969–976. [CrossRef Medline](#)