



# Humans in Love Are Singing Birds: Socially-Mediated Brain Activity in Language Production

Clara Martin<sup>1,2</sup> , Ileana Quiñones<sup>1</sup> , and Manuel Carreiras<sup>1,2,3</sup> 

<sup>1</sup>Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain

<sup>2</sup>Ikerbasque, Basque Foundation for Science, Bilbao, Spain

<sup>3</sup>University of the Basque Country, Bilbao, Spain

**Keywords:** anterior neural pathway, cross-species similarity, oral production, social context

## ABSTRACT

This functional magnetic resonance imaging (fMRI) study investigated whether and how the human speech production circuit is mediated by social factors. Participants recited a poem in the MRI scanner while viewing pictures of their lover, unknown persons, or houses to simulate different social contexts. The results showed, as expected, the recruitment of the speech production circuit during recitation. However, for the first time, we demonstrated that this circuit is tightly linked to the network underlying social cognition. The socially relevant contexts (familiar and unfamiliar persons) elicited the recruitment of a widespread bilateral circuit including regions such as the amygdala, anterior cingulate, and orbitofrontal cortex, in contrast to the non-socially relevant context (houses). We also showed a neural gradient generated by the differences in the social relevance of affective and nonaffective contexts. This study opens up a novel line of research into socially mediated speech production, revealing drastic differences in brain activation when performing the same speech production task in different social contexts. Interestingly, the analogous avian anterior neural pathway in the zebra finch is also differentially activated when the bird sings facing a (potential) mate or alone. Thus, this study suggests that despite important phylogenetic differences, speech production in humans is based, as in songbirds, on a complex neural circuitry that is modulated by evolutionarily primordial aspects such as the social relevance of the addressee.

## INTRODUCTION

Vocal communication in vertebrates is a major cognitive function shared by phylogenetically distant social animal species, such as birds and primates (Hauser et al., 2014; Seyfarth & Cheney, 2010). Evolutionarily, this function allows mating, warning calls, localizing food resources, and social learning among others (Jarvis, 2019). Considering human speech as the physical act of producing communicative sounds and sound patterns, and given that songbirds have been considered an ideal animal model to understand the neural basis of speech communication, here we will draw a parallel between brain mechanisms involved in bird vocalization and human speech production in the context of socially mediated communication (see Figures 1 and 2 in Simonyan et al., 2012). In this regard, the avian anterior neural pathway of songbirds—involved in sound pattern learning and production—is the

**Citation:** Martin, C., Quiñones, I., & Carreiras, M. (2023). Humans in love are singing birds: Socially-mediated brain activity in language production. *Neurobiology of Language*, 4(3), 501–515. [https://doi.org/10.1162/nol\\_a\\_00112](https://doi.org/10.1162/nol_a_00112)

**DOI:**  
[https://doi.org/10.1162/nol\\_a\\_00112](https://doi.org/10.1162/nol_a_00112)

**Supporting Information:**  
[https://doi.org/10.1162/nol\\_a\\_00112](https://doi.org/10.1162/nol_a_00112)

**Received:** 28 March 2022  
**Accepted:** 22 May 2023

**Competing Interests:** The authors have declared that no competing interests exist.

**Corresponding Authors:**  
Clara Martin  
[c.martin@bcbl.eu](mailto:c.martin@bcbl.eu)  
Ileana Quiñones  
[i.quinones@bcbl.eu](mailto:i.quinones@bcbl.eu)

**Handling Editor:**  
Steven Small

Copyright: © 2023  
Massachusetts Institute of Technology  
Published under a Creative Commons  
Attribution 4.0 International  
(CC BY 4.0) license

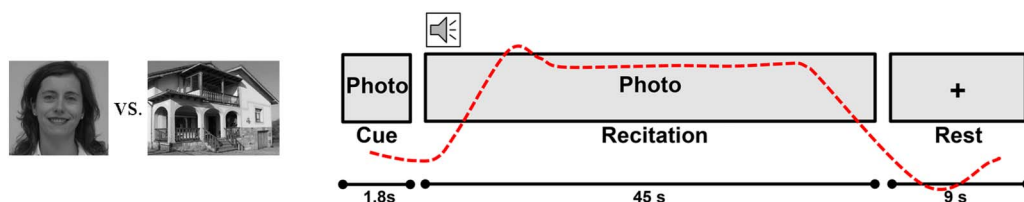
analog of the mammalian cortical-striatal-thalamocortical loop (Alexander et al., 1986; Doupe et al., 2005; Jarvis, 2004, 2019; Jarvis et al., 2005). The present functional magnetic resonance imaging (fMRI) study will investigate whether and how the human speech production circuit is mediated, as in songbirds, by social factors (Jarvis, 2004, 2019; Jarvis et al., 2005). Importantly, previous studies have shown that several neural circuits are differentially engaged in socially mediated perception in humans, specifically, by tasks that demand inferences about or empathy for another person's mental or emotional states (Lieberman, 2010; Martin et al., 2016; Molnar, 2018; Molnar et al., 2015; Woumans et al., 2015). By drawing a parallel between perception and production, we hypothesized that socially mediated speech production (i.e., taking the interlocutor into account during language processing) might produce an interaction between regions that have been related to speech production and some of the core areas of the right-lateralized social brain network (Lieberman, 2010).

#### **Avian Anterior Neural Pathway Underlying Bird Vocalization**

Male zebra finches produce both directed and undirected songs, using the same motif (i.e., syllabic ordering or ordering of elements) in two different social contexts (Dunn & Zann, 1996; Morris, 1954). A directed song is performed for a mate whereas an undirected song is produced during learning and solitary practice (see Zann, 1996, for a review). Directed songs have a stereotypical acoustic structure and ordering of syllables and syllabic patterns (Williams, 2004). Undirected songs have the same structure but are produced with larger variability and are more likely to include truncated patterns and atypical transitions (Jarvis et al., 1998; Williams, 2004). Directed and undirected vocalizations are underpinned by different patterns of neural activity along the avian anterior vocal pathway (Hessler & Doupe, 1999). This pathway runs from frontal cortical regions (the magnocellular nucleus of the anterior neostriatum; MAN) to the striatum (Area X), the dorsolateral thalamus, and back (Jarvis, 2004). Recent studies have specified the role of the different areas within the anterior pathway and their respective functions in socially mediated singing: Inactivation of the MAN region reduces variability in undirected song patterns to a stereotyped level similar to that in directed song patterns (Kojima et al., 2018; Stepanek & Doupe, 2010). Similarly, the inactivation of Area X reduces within-syllable variability (in fundamental frequency) typically observed in undirected singing (Kojima et al., 2018). Consistently, Singh Alvarado et al. (2021) recently showed that neural activity in Area X is strongly suppressed during directed song production.

#### **Mammalian Cortical-Striatal-Thalamocortical Loop Underlying Speech Production**

Speech production is a complex cognitive function that requires the processing of linguistic information as well as precise sensorimotor integration. This function is supported by a circuitry that includes the primary motor cortex, regions of the lateral inferior and medial frontal cortex, premotor cortex, supplementary motor cortex, cerebellum, and subcortical structures such as parts of the basal ganglia and the thalamus (Alexander et al., 1986; Doupe et al., 2005; Jarvis, 2004, 2019; Jarvis et al., 2005). There is evidence demonstrating that this circuit has direct projections from the facial area of the primary motor cortex in Brodmann's area 4 (the so-called laryngeal motor cortex) to the nucleus ambiguus in the brainstem (Jarvis, 2019). Interestingly, Westermann et al. (2022) demonstrated that humans have an additional affective or innate limbic vocal production pathway that involves connections from the amygdala, orbitofrontal cortex, and anterior cingulate cortex to the periaqueductal gray in



**Figure 1.** Schematic representation of the experimental design.

the brainstem (Hage & Nieder, 2016; Jarvis, 2019). This circuit supports, for example, threatening (Hage & Nieder, 2016) or laughing (Westermann et al., 2022) vocalizations used in emotional situations. However, so far there is no experimental evidence for the activation of this circuit during human language production or comprehension.

### Present Study

In the present study, we draw a parallel between human and avian vocal production, exploring whether neural circuit activation in the human brain also differs depending on the social context of speech production. We created a socially mediated language production task in which participants had to recite a previously learned poem, in the MRI scanner, in various social conditions (see Figure 1A). In order to mimic as closely as possible directed and undirected singing in songbirds we carried out an fMRI experiment in which participants recited a poem in one of two types of directed conditions (i.e., affective directed to their lover or non-affective directed to unfamiliar persons) or an undirected condition (i.e., without an addressee). (See Table S1 of the Supporting Information available at [https://doi.org/10.1162/nol\\_a\\_00112](https://doi.org/10.1162/nol_a_00112).) A social gradient was simulated by manipulating the amount of social information contained in each of the experimental conditions. As shown in Table 1, the image of the lover elicits an emotional charge associated with the face identity that makes its social relevance greater than that of the unfamiliar persons. The houses, on the other hand, do not contain social information, so they are used as a baseline—equivalent to the undirected singing in songbirds.

Our main goal was to investigate whether socially mediated speech production directed toward a lover triggered differential activation in humans similar to that observed in singing birds. Since seduction is a trans-species universal behavior and the anterior avian vocal pathway has a mammalian analog (see Alexander et al., 1986; Doupe et al., 2005), we expected brain activity during human speech production might also differ with social context: Specifically, based on previous research on songbirds, we hypothesized the cortical-striatal-thalamocortical loop might be differentially activated in directed than undirected poem recitation. Furthermore, we also explored whether the simulated social gradient elicits differential activation in brain areas involved in language processing and socially mediated behavior (see Table 1). To differentiate face-specific and face-familiarity neural effects from those related to socially mediated speech, the contrasts between faces versus houses and familiar versus unfamiliar faces were tested while the participants passively viewed the visual stimuli, without the linguistic task. This allowed us to interpret the potential results excluding the effects driven by the processing of the social context (i.e., the visual cue) rather than by socially mediated speech production in a directed way.

**Table 1.** Main contrast and specific predictions.

Main contrasts (directed vs. undirected recitation)	Linguistic manipulation	Manipulating the social context through face familiarity	Hypothesis	
			The human speech production circuit is NOT mediated by social factors	The human speech production circuit is NOT mediated by social factors
Reciting facing unfamiliar faces vs. houses	Same poem	+	<i>Equivalent to the contrast faces vs. houses in the pre-recitation time window.</i>	Face perception network (right-lateralized)
			↓	+
Reciting facing familiar faces (their lover) vs. houses		++	Face perception network (right-lateralized)	Speech production circuit (left-lateralized)
			<i>Equivalent to the contrast familiar faces vs. houses in the pre-recitation time window.</i>	Face perception network (right-lateralized)
			↓	+
			Face perception network (right-lateralized)	Regions underlying the processing of the affective information (right-lateralized)
			+	+
			Regions underlying the processing of the affective information (right-lateralized)	Speech production circuit (left-lateralized)

*Note.* Faces constitute a socially relevant stimulus. However, familiar faces, unlike the faces of strangers, allow us to anchor the face to a personal identity, which implies access to semantic, biographical, and autobiographical information. In the present case, the lover's face also contains emotional information due to the existing affective bond. Therefore, the symbol "+" (in the third column) quantifies the socially relevant information associated with each contrast.

## MATERIAL AND METHODS

### Participants

Forty-one native Spanish speakers (mean age = 28 years,  $SD = 5.37$ , range = 20–46, 21 females) took part in the study as paid volunteers. They were all right-hand dominant, had normal or corrected to normal vision, and reported no history of psychiatric or neurological disease, learning disabilities, or hearing impairments. The study protocol was conducted according to the Declaration of Helsinki guidelines and approved by the Ethics and Scientific Committees of the Basque Center on Cognition, Brain and Language (BCBL). After the experimental procedure and MRI cautions were explained, all participants signed an informed consent form and completed a standard MRI safety declaration before the scanning session.

### Stimuli

The poem to be learned and recited was “Necesito de ti” (*I need you*) by the Spanish poet Rafael de León (see the poem and an English translation in Table S1). Gray-scale pictures were used for the experimental contexts. Image size (217 × 285 pixels) and resolution (72 pixels per inch) were controlled across conditions. For each participant, we selected 10 face pictures of their lover (same photo session conditions and neutral background); 10 face pictures of a young adult of the same gender as the participant’s lover with a neutral expression for the unfamiliar face condition—same gender (see Figure 1 for an example); 10 face pictures of a young adult of a different gender to the participant’s lover for the unfamiliar face condition—different gender; and 10 pictures of houses.

Pictures of young adult faces were selected after controlling for attractiveness in an earlier test. Twenty participants (not involved in the scanning session) were presented with 60 faces in random order and asked to rate each of them for attractiveness (1: *unattractive*, 10: *highly attractive*). We selected pictures of unfamiliar faces for both genders (10 men and 10 women) within the average range of these ratings (mean = 5.64,  $SD = 0.92$ ), to avoid gender- and attractiveness-related effects in the unknown person condition.

### Experimental Procedure

The complete fMRI session consisted of two different production modes: reciting the previously memorized poem or reading a previously unseen short poem displayed on screen. Only the poem recitation is presented in this report. Each participant participated in four consecutive sessions, consisting of four randomized repetitions of a block-design functional task (see Figure 1A). Each block lasted 55.8 s (31 volumes), starting with a 1.8 s picture (visual cue) display, followed by an audio signal cueing the participant to start producing the poem, 45 s of poem recitation (with continuing picture display), followed by a 9 s (5 volumes) fixation-cross display during the interblock interval. During the four runs, participants recited the memorized poem while viewing pictures of their lover, unfamiliar faces (same or different gender as the lover), or houses. Each block (i.e., picture–poem combination) was repeated 10 times for a total of 40 blocks (10 blocks/condition).

### Instructions

Before attending the scanner session, participants were asked to memorize a short poem by imitating a digital recording of a same-sex speaker. They were instructed to listen to and repeat the recording as often as needed until they could recite the poem fluently. This process of learning by listening to and mimicking a same-sex speaker was designed to approximate the learning experience of birds, who learn song patterns by imitating a same-sex bird singing (Slater et al., 1988).

During the scanning session, participants were instructed to recite the poem as if they were reciting it to their lover when a picture of the lover appeared on the screen; as if they were reciting it to an unknown person when a picture of an unknown person appeared on the screen; and as if they were reciting it to themselves/rehearsing when a picture of a house appeared on the screen.

### MRI Acquisition

Structural and functional MRI sessions for each participant were acquired using a Siemens 3T MAGNETOM Prisma Fit scanner. High-resolution T1- and T2-weighted images were acquired with a 3D ultrafast gradient echo (MPRAGE) pulse sequence using a 64-channel head coil with the following acquisition parameters for T1: 176 contiguous sagittal slices; voxel resolution  $1 \times 1 \times 1 \text{ mm}^3$ ; repetition time (RT) = 2,530 ms, echo time (ET) = 2.36 ms; image columns = 256; image rows = 256; flip angle (flip) =  $7^\circ$  and for T2: 176 contiguous sagittal slices; voxel resolution  $1 \times 1 \times 1 \text{ mm}^3$ ; RT = 3,390 ms, ET = 389 ms; image columns = 204; image rows = 256; flip =  $120^\circ$ . The origin of the T1/T2 weighted images was set to the anterior commissure. The two structural images were then co-registered and used as priors for the segmentation. Functional volumes consisted of 618 T2\*-weighted echoplanar images, acquired with the following multiband sequence specifications: RT = 1.8 s with no time gap; ET = 29 ms; number of axial slices = 72, isotropic voxel size =  $2 \text{ mm}^3$ ; percent phase field of view (FOV) = 100% (192 mm); number of phase encoding steps = 84; pixel bandwidth = 2170; flip angle =  $73^\circ$ ; GRAPPA 4 in plane (see Table S2 in the Supporting Information for a full description of the sequence).

### fMRI Data Analysis

Functional data were analyzed using SPM12 (Ashburner et al., 2021) and related toolboxes (<https://www.fil.ion.ucl.ac.uk/spm>). Raw functional volumes were slice-time corrected taking the middle slice as reference, spatially realigned, unwarped, co-registered with the anatomical T1, and normalized to MNI (Montreal Neurological Institute) space using the unified normalization segmentation procedure. Global effects were then removed using a voxel-level linear model of the global signal proposed by Macey et al. (2004). Detrending fMRI time series were then smoothed using an isotropic 8 mm Gaussian kernel. The resulting time series from each voxel were high-pass filtered (128 s cut-off period). Statistical parametric maps were generated by modeling a univariate general linear model (GLM), using a regressor obtained by convolving the canonical hemodynamic response function with delta functions at stimulus onsets for each stimulus type (i.e., Lover's face, Unfamiliar face same gender, Unfamiliar face different gender, and House), and also including the six motion-correction parameters as regressors. In order to increase the reliability of the first-level analysis, GLM parameters were estimated with the FAST model, which uses a dictionary of covariance components based on exponential covariance functions in the context of the restricted maximum likelihood estimation (Olszowy et al., 2019). Contrast images for each of the two critical conditions compared to Houses were submitted into the second-level design. A grey matter inclusive mask was defined using individual grey matter segmentations. Only those voxels with grey matter probability value higher than 0.6 in at least 50% of participants were included. Only peaks or clusters with a significant  $p$  value after correction for multiple comparisons using family-wise error rate (FWER [ $p < 0.05$ ]; Nichols & Hayasaka, 2003) are reported.

## RESULTS

### Affective Directed (Facing Their Lover) Versus Undirected (Facing a House) Recitation

Differential brain activation was found during lover-directed versus undirected recitation in a bilaterally distributed cortico-subcortical network. This network included subcortical areas

**Table 2.** Significant clusters resulting from the comparisons “recitation directed to the lover versus viewing a house” and “recitation directed to an unknown person versus viewing a house.”

Regions	Cluster size	Coordinates			Lover vs. Houses			Unfamiliar vs. Houses		
		x	y	z	Cluster size	t	z	Cluster size	t	z
Left										
Fusiform gyrus	5683	-40	-42	-16	5683	7.70	6.00	2067	8.81	6.53
Inferior temporal gyrus		-48	-16	-18		4.84	4.27			
Cerebellum 6		-42	-60	-20		8.87	6.56			
Cerebellum crus1		-50	-58	-30		7.21	5.74		7.21	5.74
Posterior middle temporal gyrus	2523	-54	-62	18	2523	9.29	6.74			
Middle temporal gyrus		-68	-44	-4		5.74	4.88	846	6.58	5.39
Supramarginal gyrus		-54	-46	32		5.32	4.60		3.41	3.17
Cuneus	1074	-4	-70	34	1074	5.26	4.56			
Precuneus		-16	-58	34		3.41	3.18	2655	5.15	4.48
Posterior cingulum		-4	-40	32					6.68	5.44
Precentral	2119	-40	-6	42	2119	6.10	5.10			
Rolandic operculum		-52	-6	16		5.87	4.96			
Insula		-48	10	8		5.44	4.68			
Inferior frontal gyrus (triang)		-54	20	4		4.64	4.13			
Middle frontal gyrus		-50	24	40		3.94	3.60			
Caudate	628	-18	6	20	628	4.38	3.93			
Thalamus		-14	-6	12		4.12	3.74	1059	5.23	4.54
Putamen		-22	-4	12		4.25	3.83		4.12	3.74
Superior parietal gyrus	182	-40	-58	60	182	5.65	4.82			
Inferior parietal gyrus		-40	-62	46		4.70	4.17			
Angular gyrus		-56	-58	32		4.14	3.75	846	5.06	4.42
Right										
Fusiform gyrus	5683	42	-48	-18	5683	10.43	7.21	3493	11.64	7.65
Cerebellum crus1		30	-56	-30		6.38	5.27		4.19	3.79
Cerebellum crus2		4	-82	-26		5.23	4.54		4.14	3.75
Middle temporal gyrus	2720	62	-42	12	2720	9.42	6.80		7.55	5.92
Superior temporal gyrus		66	-26	22		4.87	4.29			
Precuneus	1074	0	-48	28	1074	7.98	6.14	2655	7.78	6.04
Caudate	628	14	-8	22	628	5.82	4.92			

Downloaded from [http://direct.mit.edu/nol/article-pdf/4/3/501/2155991/nol\\_a\\_00112.pdf](http://direct.mit.edu/nol/article-pdf/4/3/501/2155991/nol_a_00112.pdf) by guest on 10 September 2023

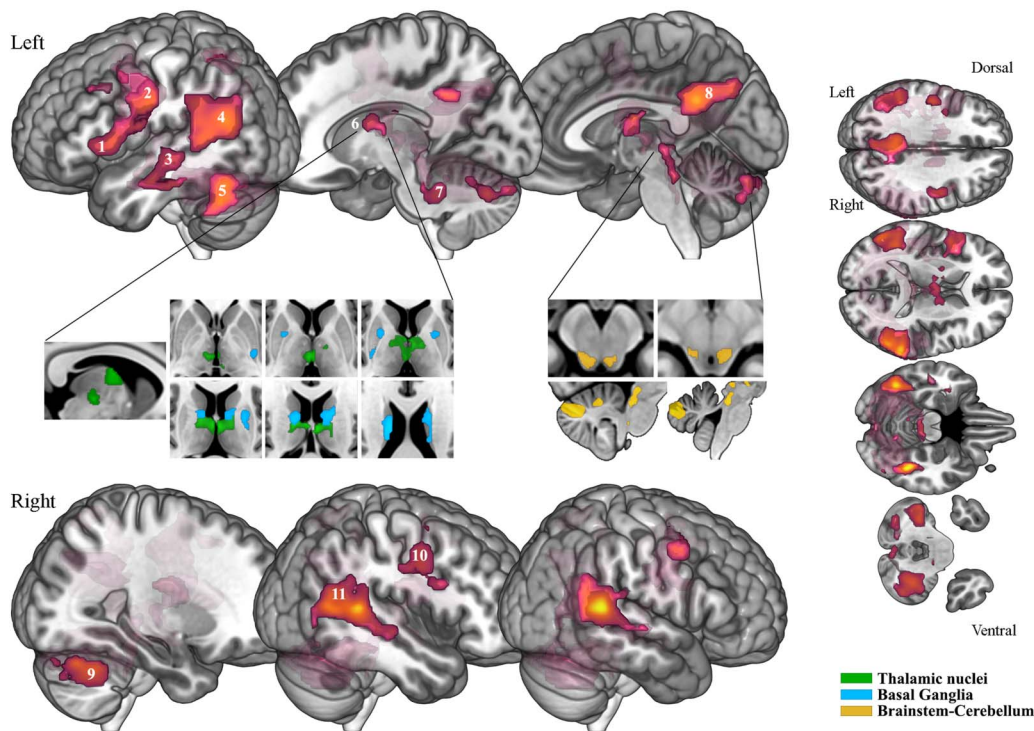
Table 2. (continued)

Regions	Cluster size	Coordinates			Lover vs. Houses			Unfamiliar vs. Houses		
		x	y	z	Cluster size	t	z	Cluster size	t	z
Thalamus		0	-10	6		3.90	3.57	1059	5.76	4.89
Putamen		18	2	14		4.60	4.10		4.90	4.31
Caudate									5.10	4.45
Globus pallidus		20	2	8		3.84	3.52			
Precentral	752	50	6	50	752	5.68	4.84			

Note. x, y, z (mm) = Coordinates in Montreal Neurological Institute space of local maxima.

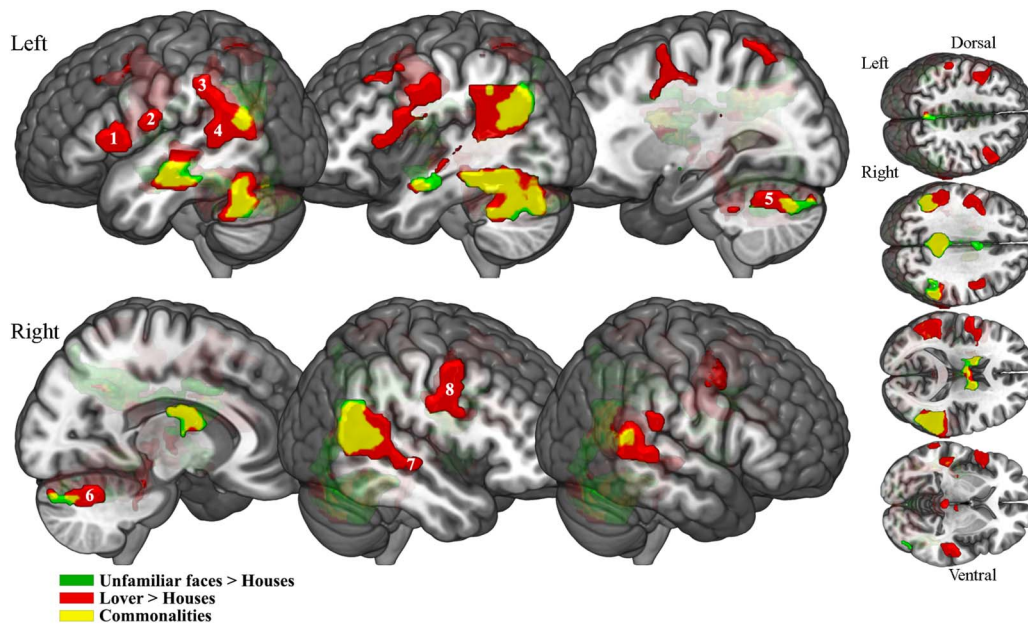
similar to those differentially activated in directed versus undirected singing in songbirds, namely the dorsal striatum (putamen and caudate nuclei), globus pallidus, and thalamus. Most strikingly, the left perisylvian language-specific system was also differentially activated by social context: Significant differences emerged in the inferior frontal gyrus (IFG), insula, anterior and posterior middle temporal gyrus (MTG), superior temporal gyrus (STG). Functional responses from additional areas such as the fusiform gyrus, angular gyrus, supramarginal gyrus, inferior and superior parietal gyri, and cerebellum also varied with social context (see Table 2 for a detailed list of regions and Figure 2 for a representation of the response pattern).

### Recitation directed to the lover versus a house



**Figure 2.** Contrast for directed versus undirected recitation (facing a house). 1: left inferior frontal gyrus; 2: left premotor/supplementary motor area; 3: left middle temporal gyrus; 4: left inferior parietal gyrus/angular gyrus/supramarginal gyrus; 5: left fusiform gyrus/cerebellum; 6: left basal ganglia/thalamus; 7: left cerebellum; 8: left posterior cingulate/precuneus; 9: right cerebellum; 10: right premotor; 11: right superior temporal sulcus/temporoparietal junction.





**Figure 3.** Comparison between recitation viewing the lover versus a house and recitation viewing an unfamiliar person versus a house. 1: left inferior frontal gyrus; 2: left middle frontal gyrus; 3: left superior parietal gyrus; 4: left superior temporal sulcus (STS); 5: left cerebellum; 6: right cerebellum; 7: right STS; 8: right premotor.

### Contrasting the Two Directed Conditions (Affective and Nonaffective)

Figure 3 displays the superimposition of the contrasts for recitation facing lovers versus houses—represented in red—and recitation facing unfamiliar persons versus houses—represented in green (also see Table 2). Activation common to the two contrasts is represented in yellow. Interestingly, while some areas were common to both contrasts, others appeared only in the affective directed condition. The left and right IFG, left and right middle frontal gyrus, left and right posterior part of the STG, left MTG, and left inferior parietal gyrus (IPG) emerged as significant only for the affective directed condition.

Possible gender-related effects concerning the unfamiliar faces were tested by contrasting unfamiliar persons of the same gender as the participant’s lover with unfamiliar persons of a different gender as the participant’s lover. This contrast yielded no significant results, so the plots were restricted to the comparison between lovers’ and same gender unfamiliar faces.

### Controlling for Face-Specific and Familiarity Effects Using the Pre-Recitation Time Window

To disentangle whether the differential activation in socially mediated speech circuitry detected by the critical comparisons was driven by the social context (i.e., exposure to different faces) we tested for face-specific and face-familiarity responses in the pre-recitation time window—while the participants passively viewed the visual stimuli, without the linguistic task. Face-specific voxels were defined by larger blood oxygen level dependent responses for faces than for houses (Table 3). The posterior fusiform and inferior occipital gyri presented regional maxima in both cerebral hemispheres, corresponding to the fusiform face area (FFA) and occipital face area, two face-specific areas previously described in other studies (Kanwisher et al., 1997). Responses to lovers’ faces were larger than responses to

**Table 3.** Control for face-specific and face-familiarity activations.

Faces > Houses						
Regions	Cluster size (voxels)	x (mm)	y (mm)	z (mm)	T scores	Z scores
Fusiform right	5739	42	-48	-18	10.43	7.21
Fusiform left		-42	-56	-18	7.51	5.90
Precuneus right	1534	0	-50	26	7.36	5.82
Cingulum_post left		-8	-46	32	6.20	5.16
Cuneus left		-4	-70	34	5.26	4.56
Middle temporal left	936	-52	-66	20	6.58	5.39
Angular left		-56	-58	32	4.10	3.72
Supramarginal left		-50	-46	34	3.41	3.17
Caudate right	657	14	-8	22	5.62	4.80
Thalamus left		-4	-10	10	4.42	3.97
Familiar > Unfamiliar faces						
Fusiform right	9847	42	-48	-18	11.64	7.65
Fusiform left		-40	-40	-20	8.81	6.53
Middle temporal right		54	-56	20	7.55	5.92
Precuneus right	3121	0	-52	24	7.78	6.04
Precuneus left		4	-62	36	7.18	5.72
Cingulum_post left		-4	-40	32	6.68	5.44
Middle temporal left	1110	-52	-66	20	6.58	5.39
Occipital_mid left		-42	-66	34	5.06	4.42
Supramarginal left		-50	-46	34	3.41	3.17
Insula left	445	-40	4	22	4.76	4.21
Precentral right	445	52	8	50	4.46	3.99
Middle frontal right		34	26	52	4.18	3.78
Middle frontal left		30	14	60	3.97	3.62

Note. Face-specific voxels were defined by larger blood oxygen level dependent responses for faces than for houses whereas brain areas related to face familiarity were defined by comparing lover and unfamiliar faces. Both contrasts were tested in the pre-recitation time window.

unfamiliar faces in several regions, including the right FFA, right inferior temporal gyrus, and right parahippocampus, regions known to be related to face identity recognition (Gobbini & Haxby, 2007; Table 3). Strikingly, when these two statistical parametric maps were used as a restrictive mask for the contrasts between directed (affective and nonaffective) and undirected recitation we did not find any significant voxels, which indicates that during poem recitation the speech production circuit interacted with the social brain network beyond the face-specific areas.

## DISCUSSION

This study explored brain activation during language production in various socially mediated contexts, by testing participants reciting a poem in the MRI scanner while facing a picture of their lover, an unknown person, or a house. The results showed, as expected, the recruitment of the speech production circuit during recitation. However, for the first time, we demonstrated that this cortical-striatal-thalamocortical loop underlying speech production is linked to the network underlying social cognition. The socially relevant contexts (familiar and unfamiliar persons) elicited the recruitment of a widespread bilateral circuit including regions such as the amygdala, anterior cingulate, and orbitofrontal cortex, whereas the noninformative context (houses) showed no effect on the speech production network. We also showed a neural gradient generated by the differences in the social relevance of affective and nonaffective contexts.

While the exact function of striatal substructures in human language has not yet been clearly defined, their role in motor control has been repeatedly observed. Specifically, the putamen and caudate nuclei of the striatum and the thalamus have been shown to be linked to memory and language monitoring (Hamm & Mattfeld, 2019; Saalman & Kastner, 2015). The putamen is also involved in the control of learned vocal patterns and the initiation and execution of speech (Price, 2010; Simonyan et al., 2012). Basal ganglia nuclei are also involved in the regulation of motor actions in response to cognitive activities and the processing of emotional and motivational stimuli (Heimer et al., 1982; Turner & Desmurget, 2010). Here, we demonstrate that the activation of this cortical-striatal-thalamocortical loop during production is also socially mediated. Interestingly, some of those regions which were differentially activated in language production facing a lover or a stranger (e.g., striatum and thalamus) have also been shown to be related to “intense love,” in studies presenting participants with pictures of a long-term partner versus highly and low-familiar persons (Acevedo et al., 2012).

Cortical brain areas that were differentially activated in the directed poem recitation included the core areas of the ventral and dorsal streams of the speech production system (Jarvis, 2019): the IFG, premotor cortex, and insula, which subserve articulatory operations (Basilakos et al., 2018; Price, 2010); the anterior MTG and angular gyrus, which work together to support lexicosemantic combinatorial processes (Hickok & Poeppel, 2007), and the superior temporal sulcus (STS), which codes phonological information as well as sensory-based representations of speech within the auditory–motor integration network (Hickok & Poeppel, 2007; Houde & Nagarajan, 2011). However, importantly, differential activations were not only observed in the cortical-striatal-thalamocortical loop but also in regions that have been previously related to the social brain (Lieberman, 2010). This shows that while social context modulates analogous ancestral brain regions and circuits in humans and birds, additional cortical circuits are recruited in humans, reflecting their unique capacity for language. We also found differential activation for directed versus undirected speech production in the IPG, amygdala, orbitofrontal cortex, and anterior cingulate cortex. Since these areas have been proposed to constitute a core social cognition network in perception (Lieberman, 2010), this result reveals a striking parallel between the processing of socially mediated stimuli and socially oriented language production. In the same direction, previous studies have shown that nonhuman primates have an additional affective or innate vocal production pathway involving connections from the amygdala, orbitofrontal cortex, and anterior cingulate cortex to the periaqueductal gray in the brainstem (Hage & Nieder, 2016; Jarvis, 2019). This circuit has been identified in humans, but so far mainly in the context of threatening (Hage & Nieder, 2016) or laughing (Westermann

et al., 2022) vocalizations used in emotional situations. Crucially, our results reveal the existence of a similar functional pathway in language production. The activation of primitive regions dedicated to the processing of social and emotional information during the recitation of a poem supports theories of language modularity by revealing that the language network is closely related to our emotional ancestral brain (Allott & Smith, 2021; Mesulam et al., 2021).

Previous evidence for the social modulation of language production has come from behavioral studies showing that some acoustic aspects of adult speech depend on factors related to the addressee (see Cooke et al., 2014, for a review). Regarding the association between face perception and language production, previous studies have shown that faces can act as cues for language selection in bilinguals prior to speech production (Woumans et al., 2015). Here, we add novel insight into this research on socially mediated speech production, showing that neural activations differ during speech production of the same text, depending on the social context (i.e., directed to a lover versus an unfamiliar face or house).

In the present study, responses of the human speech production network based on the addressee (being a lover or the picture of a house) mimic modulations previously observed in the analogous pathway in songbirds. Thus, this study opens a novel link between research on animals and humans, suggesting that social context might modulate brain activation in (at least partially) equivalent ways in human speech production as that already established for birdsong.

Despite the interesting parallel between brain circuit activation in humans and songbirds identified in the present study, we acknowledge some limitations that require further investigation: Participants in our study learned a poem by listening to a recorded version of it, and we did not explore brain activation modulation during learning, which was performed at home prior to MRI testing.

Although we tried to approximate poem learning to song pattern learning as much as possible, we acknowledge the limitation that songbirds learn by imitating a physical conspecific and not via a tape recording (Chen et al., 2016; Derégnaucourt et al., 2013). It would be interesting to explore the learning process with MRI data acquisition and with participants mimicking a “tutor” present in the room. Several studies on songbirds have shown that the presence of a tutor has a strong influence on vocal learning: Juvenile zebra finches are quicker and more accurate in learning to mimic songs when interacting with a conspecific tutor than when passively listening to the song through a loudspeaker (Chen et al., 2016; Derégnaucourt et al., 2013).

Furthermore, it would be interesting to test participants reciting the poem to their lover, an unknown person, or alone, while the intended listener was physically present in the room instead of represented by pictures. Modulation of the socially mediated brain circuits should certainly be larger in such a more naturalistic setting. However, the fact that we were able to detect significant differences in both social and language networks despite using simulated/virtual tutors and addressees points to the robustness of our findings.

To conclude, the current study (1) opens up a novel line of research on socially mediated language production, showing that cortical and subcortical regions are differentially activated depending on the addressee; (2) highlights the benefits of comparing sophisticated communication practices, such as birdsong and human speech production, to better understand human language; and (3) reveals the contribution of cortical and subcortical components and overall complexity of brain circuits underlying human language.

## ACKNOWLEDGMENTS

The authors would like to thank all the participants who agreed to take part in this study. We would also like to thank Caroline Handley for her help with editing this work, and Sandra Gisbert and David Carcedo for their help with data acquisition.

## FUNDING INFORMATION

Manuel Carreiras, Eusko Jaurlaritza (<https://dx.doi.org/10.13039/501100003086>), Award ID: BERC 2018 2021. Manuel Carreiras, Agencia Estatal de Investigación (<https://dx.doi.org/10.13039/501100011033>), Award ID: CEX2020-001010-S. Manuel Carreiras, Ministerio de Economía y Competitividad (<https://dx.doi.org/10.13039/501100003329>), Award ID: RTI2018 093547 B I00. Ileana Quiñones, Ministerio de Economía y Competitividad (<https://dx.doi.org/10.13039/501100003329>), Award ID: RTI2018 093547 B I00. Clara Martín, Ministerio de Economía y Competitividad (<https://dx.doi.org/10.13039/501100003329>), Award ID: PID2020-113926GB-I00. Clara Martín, H2020 European Research Council (<https://dx.doi.org/10.13039/100010663>), Award ID: 819093.

## AUTHOR CONTRIBUTIONS

**Clara Martín:** Conceptualization; Formal analysis; Funding acquisition; Project administration; Writing – original draft; Writing – review & editing. **Ileana Quiñones:** Conceptualization; Formal analysis; Funding acquisition; Investigation; Visualization; Writing – original draft; Writing – review & editing. **Manuel Carreiras:** Conceptualization; Funding acquisition; Writing – review & editing.

## DATA AND CODE AVAILABILITY STATEMENT

Following BCBL data sharing policy, raw data used in this study have been shared upon specific request to the authors and according to a formal data sharing agreement ([https://www.bcbl.eu/DataSharing/NeuroLang2022\\_Birds/](https://www.bcbl.eu/DataSharing/NeuroLang2022_Birds/)). However, MATLAB codes used for preprocessing and statistical analyses are available at <https://github.com/iquinones1959/singingbirds>. Individual statistical models have also been published as open data, including T-maps for each contrast cited in the text and beta maps associated with each regressor (10.5281/zenodo.7927725).

## REFERENCES

- Acevedo, B. P., Aron, A., Fisher, H. E., & Brown, L. L. (2012). Neural correlates of long-term intense romantic love. *Social Cognitive and Affective Neuroscience*, 7(2), 145–159. <https://doi.org/10.1093/scan/nsq092>, PubMed: 21208991
- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9(1), 357–381. <https://doi.org/10.1146/annurev.ne.09.030186.002041>, PubMed: 3085570
- Allott, N., & Smith, N. (2021). Chomsky and Fodor on modularity. In N. Allott, T. Lohndal, & G. Rey (Eds.), *A companion to Chomsky* (pp. 529–543). John Wiley & Sons. <https://doi.org/10.1002/9781119598732.ch34>
- Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K., Gitelman, D., Glauche, V., Henson, R., Hutton, C., Jafarian, A., Kiebel, S., Kilner, J., Litvak, V., Mattout, J., Moran, R., Penny, W., Phillips, C., Razi, A., ... Zeidman, P. (2021). *SPM12 manual*. Wellcome Centre for Human Neuroimaging.
- Basilakos, A., Smith, K. G., Fillmore, P., Fridriksson, J., & Fedorenko, E. (2018). Functional characterization of the human speech articulation network. *Cerebral Cortex*, 28(5), 1816–1830. <https://doi.org/10.1093/cercor/bhx100>, PubMed: 28453613
- Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences*, 113(24), 6641–6646. <https://doi.org/10.1073/pnas.1522306113>, PubMed: 27247385
- Cooke, M., King, S., Garnier, M., & Aubanel, V. (2014). The listening talker: A review of human and algorithmic context-induced modifications of speech. *Computer Speech & Language*, 28(2), 543–571. <https://doi.org/10.1016/j.csl.2013.08.003>
- Derégnaucourt, S., Poirier, C., Van der Kant, A., Van der Linden, A., & Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *Journal of Physiology–Paris*, 107(3), 210–218. <https://doi.org/10.1016/j.jphysparis.2012.08.003>, PubMed: 22982543

- Doupe, A. J., Perkel, D. J., Reiner, A., & Stern, E. A. (2005). Bird-brains could teach basal ganglia research a new song. *Trends in Neurosciences*, 28(7), 353–363. <https://doi.org/10.1016/j.tins.2005.05.005>, PubMed: 15935486
- Dunn, A., & Zann, R. (1996). Undirected song encourages the breeding female zebra finch to remain in the nest. *Ethology*, 102(4), 540–548. <https://doi.org/10.1111/j.1439-0310.1996.tb01146.x>
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45(1), 32–41. <https://doi.org/10.1016/j.neuropsychologia.2006.04.015>, PubMed: 16797608
- Hage, S. R., & Nieder, A. (2016). Dual neural network model for the evolution of speech and language. *Trends in Neurosciences*, 39(12), 813–829. <https://doi.org/10.1016/j.tins.2016.10.006>, PubMed: 27884462
- Hamm, A. G., & Mattfeld, A. T. (2019). Distinct neural circuits underlie prospective and concurrent memory-guided behavior. *Cell Reports*, 28(10), 2541–2553. <https://doi.org/10.1016/j.celrep.2019.08.002>, PubMed: 31484067
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., & Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, 5, Article 401. <https://doi.org/10.3389/fpsyg.2014.00401>, PubMed: 24847300
- Heimer, L., Switzer, R. D., & Van Hoesen, G. W. (1982). Ventral striatum and ventral pallidum: Components of the motor system? *Trends in Neurosciences*, 5, 83–87. [https://doi.org/10.1016/0166-2236\(82\)90037-6](https://doi.org/10.1016/0166-2236(82)90037-6)
- Hessler, N. A., & Doupe, A. J. (1999). Social context modulates singing-related neural activity in the songbird forebrain. *Nature Neuroscience*, 2(3), 209–211. <https://doi.org/10.1038/6306>, PubMed: 10195211
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/nrn2113>, PubMed: 17431404
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, Article 82. <https://doi.org/10.3389/fnhum.2011.00082>, PubMed: 22046152
- Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016(1), 749–777. <https://doi.org/10.1196/annals.1298.038>, PubMed: 15313804
- Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, 366(6461), 50–54. <https://doi.org/10.1126/science.aax0287>, PubMed: 31604300
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., Medina, L., Paxinos, G., Perkel, D. J., Shimizu, T., Striedter, G., Wild, J. M., Ball, G. F., Dugas-Ford, J., Durand, S. E., Hough, G. E., Husband, S., Kubikova, L., Lee, D. W., ... Butler, A. B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2), 151–159. <https://doi.org/10.1038/nrn1606>, PubMed: 15685220
- Jarvis, E. D., Scharff, C., Grossman, M. R., Ramos, J. A., & Nottebohm, F. (1998). For whom the bird sings: Context-dependent gene expression. *Neuron*, 21(4), 775–788. [https://doi.org/10.1016/S0896-6273\(00\)80594-2](https://doi.org/10.1016/S0896-6273(00)80594-2), PubMed: 9808464
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>, PubMed: 9151747
- Kojima, S., Kao, M. H., Doupe, A. J., & Brainard, M. S. (2018). The avian basal ganglia are a source of rapid behavioral variation that enables vocal motor exploration. *Journal of Neuroscience*, 38(45), 9635–9647. <https://doi.org/10.1523/JNEUROSCI.2915-17.2018>, PubMed: 30249800
- Lieberman, M. D. (2010). Social cognitive neuroscience. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 1, pp. 143–193). Wiley. <https://doi.org/10.1002/9780470561119.socpsy001005>
- Macey, P. M., Macey, K. E., Kumar, R., & Harper, R. M. (2004). A method for removal of global effects from fMRI time series. *NeuroImage*, 22(1), 360–366. <https://doi.org/10.1016/j.neuroimage.2003.12.042>, PubMed: 15110027
- Martin, C. D., Molnar, M., & Carreiras, M. (2016). The proactive bilingual brain: Using interlocutor identity to generate predictions for language processing. *Scientific Reports*, 6(1), Article 26171. <https://doi.org/10.1038/srep26171>, PubMed: 27173937
- Mesulam, M. M., Coventry, C. A., Rader, B. M., Kuang, A., Sridhar, J., Martersteck, A., Zhang, H., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2021). Modularity and granularity across the language network: A primary progressive aphasia perspective. *Cortex*, 141, 482–496. <https://doi.org/10.1016/j.cortex.2021.05.002>, PubMed: 34153680
- Molnar, M. (2018). The role of voice familiarity in bilingual speech processing. *Journal of the Acoustical Society of America*, 144(3), 1864–1865. <https://doi.org/10.1121/1.5068198>
- Molnar, M., Ibáñez-Molina, A., & Carreiras, M. (2015). Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language*, 81, 91–104. <https://doi.org/10.1016/j.jml.2015.01.002>
- Morris, D. (1954). The reproductive behaviour of the zebra finch (*Poephila guttata*), with special reference to pseudofemale behaviour and displacement activities. *Behaviour*, 6(1), 271–322. <https://doi.org/10.1163/156853954X00130>
- Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, 12(5), 419–446. <https://doi.org/10.1191/0962280203sm341ra>, PubMed: 14599004
- Olzoway, W., Aston, J., Rua, C., & Williams, G. B. (2019). Accurate autocorrelation modeling substantially improves fMRI reliability. *Nature Communications*, 10(1), Article 1220. <https://doi.org/10.1038/s41467-019-09230-w>, PubMed: 30899012
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191(1), 62–88. <https://doi.org/10.1111/j.1749-6632.2010.05444.x>, PubMed: 20392276
- Saalmann, Y. B., & Kastner, S. (2015). The cognitive thalamus. *Frontiers in Systems Neuroscience*, 9, Article 39. <https://doi.org/10.3389/fnsys.2015.00039>, PubMed: 25852498
- Seyfarth, R. M., & Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain and Language*, 115(1), 92–100. <https://doi.org/10.1016/j.bandl.2009.10.003>, PubMed: 19944456
- Simonyan, K., Horwitz, B., & Jarvis, E. D. (2012). Dopamine regulation of human speech and bird song: A critical review. *Brain and Language*, 122(3), 142–150. <https://doi.org/10.1016/j.bandl.2011.12.009>, PubMed: 22284300
- Singh Alvarado, J., Goffinet, J., Michael, V., Liberti, W., Hatfield, J., Gardner, T., Pearson, J., & Mooney, R. (2021). Neural dynamics underlying birdsong practice and performance. *Nature*, 599(7886), 635–639. <https://doi.org/10.1038/s41586-021-04004-1>, PubMed: 34671166
- Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches (*Taeniopygia guttata*): Progress and prospects. In J. S. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater (Eds.),

- Advances in the study of behavior* (Vol. 18, pp. 1–34). Academic Press. [https://doi.org/10.1016/S0065-3454\(08\)60308-3](https://doi.org/10.1016/S0065-3454(08)60308-3)
- Stepanek, L., & Doupe, A. J. (2010). Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *Journal of Neurophysiology*, *104*(5), 2474–2486. <https://doi.org/10.1152/jn.00977.2009>, PubMed: 20884763
- Turner, R. S., & Desmurget, M. (2010). Basal ganglia contributions to motor control: A vigorous tutor. *Current Opinion in Neurobiology*, *20*(6), 704–716. <https://doi.org/10.1016/j.conb.2010.08.022>, PubMed: 20850966
- Westermann, B., Lotze, M., Varra, L., Versteeg, N., Domin, M., Nicolet, L., Orbist, M., Klepzig, K., Marbot, L., Lämmler, L., Fiedler, K., & Wattendorf, E. (2022). When laughter arrests speech: fMRI-based evidence. *Philosophical Transactions of the Royal Society B*, *377*(1863), Article 20210182. <https://doi.org/10.1098/rstb.2021.0182>, PubMed: 36126674
- Williams, H. (2004). Birdsong and singing behavior. *Annals of the New York Academy of Sciences*, *1016*(1), 1–30. <https://doi.org/10.1196/annals.1298.029>, PubMed: 15313767
- Woumans, E., Martin, C. D., Vanden Bulcke, C., Van Assche, E., Costa, A., Hartsuiker, R. J., & Duyck, W. (2015). Can faces prime a language? *Psychological Science*, *26*(9), 1343–1352. <https://doi.org/10.1177/0956797615589330>, PubMed: 26209531
- Zann, R. A. (1996). *The zebra finch: A synthesis of field and laboratory studies*. Oxford University Press.