



# Auditory observation of infant-directed speech by mothers: experience-dependent interaction between language and emotion in the basal ganglia

Yoshi-Taka Matsuda<sup>1,2,3,4\*</sup>, Kenichi Ueno<sup>5</sup>, Kang Cheng<sup>5</sup>, Yukuo Konishi<sup>4</sup>, Reiko Mazuka<sup>3</sup> and Kazuo Okanoya<sup>1,2,6</sup>

<sup>1</sup> Okanoya Emotional Information Project, Exploratory Research for Advanced Technology (ERATO), Japan Science and Technology Agency (JST), Wako, Saitama, Japan

<sup>2</sup> Cognition and Behavior Joint Research Laboratory, RIKEN Brain Science Institute, Wako, Saitama, Japan

<sup>3</sup> Laboratory for Language Development, RIKEN Brain Science Institute, Wako, Saitama, Japan

<sup>4</sup> Center for Baby Science, Doshisha University, Kyoto, Japan

<sup>5</sup> Support Unit for Functional MRI, RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

<sup>6</sup> Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

## Edited by:

Lynne E. Bernstein, George Washington University, USA

## Reviewed by:

Yang Zhang, University of Minnesota, USA

Malathi Thothathiri, George Washington University, USA

## \*Correspondence:

Yoshi-Taka Matsuda, Center for Baby Science, Doshisha University, 4-1-1 Kizugawadai, Kizugawa, Kyoto, 619-0225, Japan  
e-mail: matsuda@brain.riken.jp

Adults address infants with a special speech register known as infant-directed speech (IDS), which conveys both linguistic and emotional information through its characteristic lexicon and exaggerated prosody (e.g., higher pitched, slower, and hyperarticulated). Although caregivers are known to regulate the usage of IDS (linguistic and emotional components) depending on their child's development, the underlying neural substrates of this flexible modification are largely unknown. Here, using an auditory observation method and functional magnetic resonance imaging (fMRI) of four different groups of females, we revealed the experience-dependent influence of the emotional component on linguistic processing in the right caudate nucleus when mothers process IDS: (1) non-mothers, who do not use IDS regularly, showed no significant difference between IDS and adult-directed speech (ADS); (2) mothers with preverbal infants, who primarily use the emotional component of IDS, showed the main effect of the emotional component of IDS; (3) mothers with toddlers at the two-word stage, who use both linguistic and emotional components of IDS, showed an interaction between the linguistic and emotional components of IDS; and (4) mothers with school-age children, who use ADS rather than IDS toward their children, showed a tendency toward the main effect of ADS. The task that was most comparable to the naturalistic categories of IDS (i.e., explicit-language and implicit-emotion processing) recruited the right caudate nucleus, but it was not recruited in the control, less naturalistic condition (explicit-emotion and implicit-language processing). Our results indicate that the right caudate nucleus processes experience- and task-dependent interactions between language and emotion in mothers' IDS.

**Keywords:** basal ganglia, emotion, experience dependency, functional magnetic resonance imaging (fMRI), infant-directed speech, language, maternal behavior

## INTRODUCTION

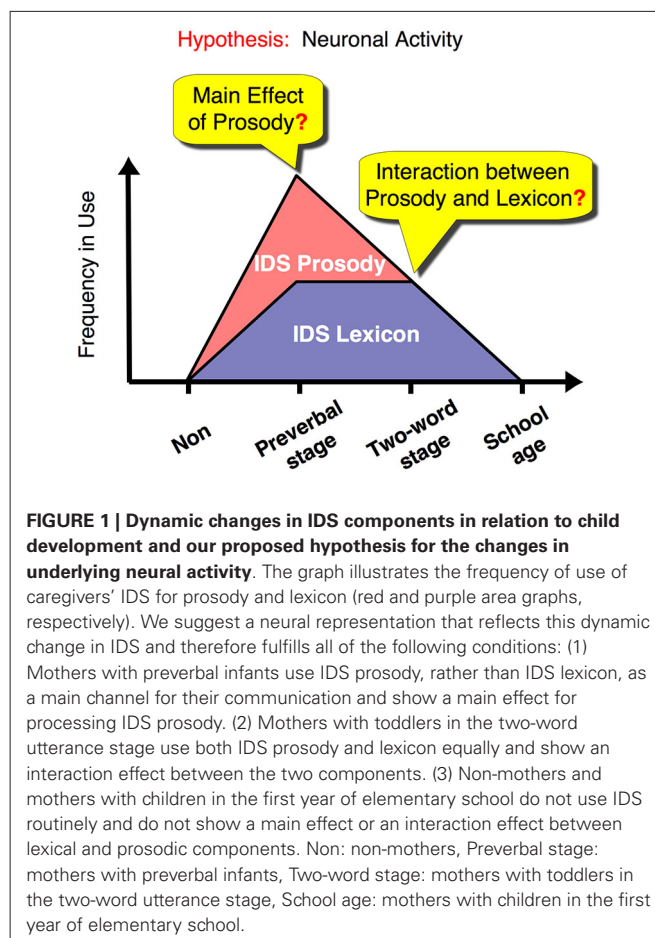
Across languages and cultures, adults often use a speech register known as infant-directed speech (IDS) when talking to infants. IDS typically involves modifications of speech prosody, such as a higher pitch, an exaggerated intonation contour, and lexical, syntactic, and other modulations (Snow and Ferguson, 1977; Soderstrom, 2007; Saint-Georges et al., 2013). The lexical modifications often include a sizeable number of "baby-talk" words and frequently include onomatopoeia (Mazuka et al., 2008). These modifications draw infants' attention to speech (Werker and McLeod, 1989; Cooper and Aslin, 1990; Fernald, 1991; Barker and Newman, 2004; Zangl and Mills, 2007) and may play a facilitating role in the emotional and linguistic development of

infants (Thiessen et al., 2005; Vallabha et al., 2007; Kuhl and Rivera-Gaxiola, 2008; Taumoepeau and Ruffman, 2008). Recent neuroimaging studies support these IDS functions. The perception of IDS activated the orbitofrontal cortex and language areas in infants (Dehaene-Lambertz et al., 2002; Pena et al., 2003; Saito et al., 2007), where the affective and linguistic properties are encoded, respectively (Minagawa-Kawai et al., 2007, 2009). Interestingly, even when affective exaggeration is removed in IDS, formant-exaggeration alone could elicit enhanced neural activities in the infant brain (Zhang et al., 2011).

Because IDS conveys both affective and linguistic information towards infants (Kuhl et al., 1997; Trainor et al., 2000; Burnham et al., 2002), caregivers are known to regulate the

proportion of the two types of information depending on the child's level of development to obtain a sufficient reaction from developing infants (Snow and Ferguson, 1977; Saint-Georges et al., 2013). Emotionally biased IDS toward preverbal infants gradually shifts to both emotionally and linguistically balanced speech toward talking infants at approximately the two-word utterance stage (Penman et al., 1983; Kitamura and Burnham, 2003; Kajikawa et al., 2004; Amano et al., 2006; Ogura, 2006). As the infant's language ability develops, IDS changes to a less emotional and linguistically more complex style. By the time caregivers are speaking to school-age children, the speech is comparable to adult-directed speech (ADS; Garnica, 1977). Little is known, however, as to what neural processes in adult caregivers encode this flexible modification of IDS. We previously reported the experience-dependent processing of IDS in language areas (Matsuda et al., 2011). Mothers with preverbal infants showed an enhanced activity of the emotional (prosodic) component of IDS. This group is known to use the emotional component of IDS more heavily compared with non-mothers or mothers with older children (Amano et al., 2006). Similarly, mothers with infants in the two-word utterance stage and mothers with preverbal infants showed an enhanced activity of the linguistic (lexical) component of IDS. These mother groups are known to use the linguistic component of IDS frequently (Ogura, 2006). Mothers with school-age children and non-mothers do not use emotional or linguistic components of IDS regularly and did not show enhanced activity. Thus, these results indicate the presence of use-dependent processing of IDS in the language areas, which reflects the current usage of IDS properties. However, this study only investigated the main effect of either the prosody (i.e., IDS prosody vs. ADS prosody) or lexicon (i.e., IDS lexicon vs. ADS lexicon) separately without thorough consideration of the mutual interaction between prosody and lexicon, i.e., (IDS prosody, ADS prosody)  $\times$  (IDS lexicon, ADS lexicon). In the naturalistic environment, mothers use both prosodic and lexical properties simultaneously to communicate with their infants and regulate the proportion used depending on the child's level of development (Snow and Ferguson, 1977; Saint-Georges et al., 2013); thus, it remains unclear as to what neural process encodes the dynamic interaction and integration between the two types of IDS depending on the parenting experience.

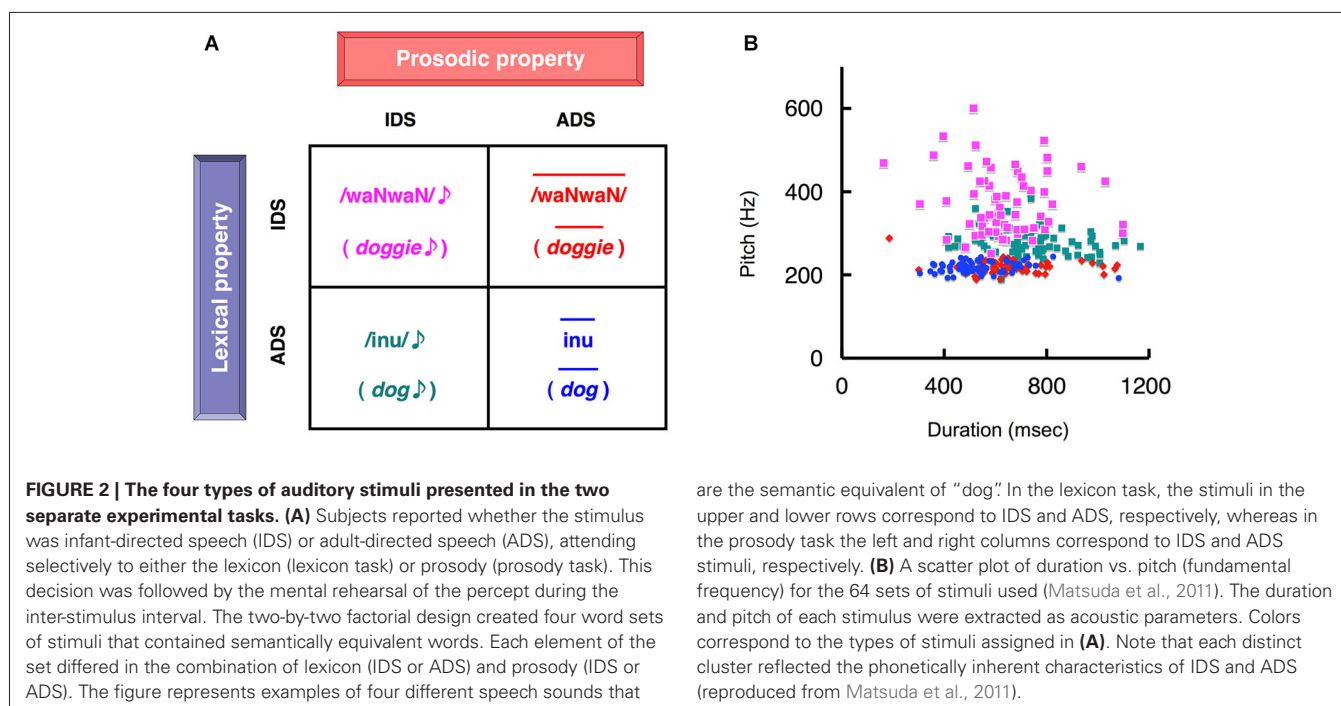
Here, we used functional magnetic resonance imaging (fMRI) to investigate the specific neural substrates that showed a dynamic change corresponding to the interaction between emotion (prosody) and language (lexicon) depending on the child's developmental stage (see time course in **Figure 1**). To address this issue, we reanalyzed our previous data (Matsuda et al., 2011), which comprised four groups of female adult participants, including women without children and women with children in different stages of language development: (1) non-mothers; (2) mothers with preverbal infants; (3) mothers with toddlers speaking in two-word utterances; and (4) mothers with children in elementary school. Each group of mothers represents a distinct stage in the time course of IDS usage. Of the four groups, the mothers with preverbal infants are the heaviest users of prosodic IDS (Amano et al., 2006). Around the time children are in



the two-word utterance stage, mothers proportionally suppress prosodic modulation (Amano et al., 2006), while they continue to use lexical IDS (Ogura, 2006). This is the period when emotionally and linguistically balanced IDS is produced. Subsequently, mothers virtually cease to use both prosodic and lexical IDS toward elementary-school children. In this study, non-mothers were used as a control group.

The experiment consisted of two separate tasks, which contrasted a more naturalistic environment with a less naturalistic control condition in regard to the actual usage of IDS. As caregivers are known to address infants with a conscious conversational and sub-conscious emotional speech style in IDS (Snow and Ferguson, 1977), the explicit-linguistic and implicit-emotional processing (the lexicon task) resembled a naturalistic condition. However, the explicit-emotional and implicit-linguistic processing (the prosody task) was different from the actual environment and served as a control condition in this study.

In the lexicon task (naturalistic condition), the subjects judged whether auditory stimuli were IDS or ADS by focusing on the word while ignoring prosody (**Figure 2A**). In the prosody task (control condition), the subjects judged whether auditory stimuli were IDS or ADS by focusing on only prosody while ignoring word content (**Figure 2A**).



We used an IDS perception experiment rather than a production experiment to avoid the effect of head movement on fMRI signal. One property of IDS production is hyperarticulation, and speakers move their mouths more dynamically as opposed to ADS production. This hyperarticulation could inevitably cause head movement and subsequently decrease the signal-to-noise ratio of the fMRI signal. Owing to a fine-tuned sensorimotor “resonance” system, which develops as a consequence of extensive motor practice (Aglioti et al., 2008), we conducted an IDS perception experiment rather than a production experiment.

## METHODS

All data sets were obtained as part of our previous study (Matsuda et al., 2011) and were reanalyzed with different methods to address the issues raised in the current study.

## PARTICIPANTS

Sixty-nine participants comprised four different groups, which included 15 non-mothers, 20 first-time mothers with preverbal infants, 16 first-time mothers with toddlers who spoke two-word utterances, and 18 mothers with children in the first grade. We confirmed the presence of two-word utterances in each toddler via parental report on the toddler form of the MacArthur Communicative Development Inventory (Watamaki and Ogura, 2004). No parent was single, and all mothers were the primary caregivers. All subjects were neurologically normal, right-handed Japanese speakers, and no one was professionally involved in infant care. All groups were matched for education and verbal IQ. All subjects provided informed consent according to the procedures approved by the RIKEN Brain Science Institute Ethics Committee and the Functional MRI Safety and Ethics Committee in Wako (for further details, see Matsuda et al., 2011).

## AUDITORY STIMULI

The stimulus sets were the same as those used in our previous study (Matsuda et al., 2011). Briefly, we prepared 64 pairs of word items with semantically equivalent meanings but different vocabularies: words utilized in the context of (i) IDS and (ii) ADS, e.g., /waNwaN/ (doggie) and /inu/ (dog), respectively. The stimuli were recordings of a professional actress speaking each word in two prosodic conditions: (i) IDS prosody in the presence of her preverbal infant son and (ii) ADS prosody directed at an adult, which resulted in 256 stimuli (64 lexical items  $\times$  2 linguistic conditions  $\times$  2 prosodic conditions). This balanced two-by-two factorial design, which used lexicon or prosody to identify an item as IDS or ADS, resulted in four different types of stimuli (**Figure 2A**). Congruous ADS speech (ADS words with ADS prosody) had a shorter duration and lower pitch, congruous IDS speech (IDS words with IDS prosody) had a varying duration and higher pitch, and the two incongruous speech-stimuli (IDS words with ADS prosody and ADS words with IDS prosody) fell between the two other forms of speech (**Figure 2B**). Additional details have been provided in our previous work (Matsuda et al., 2011).

## EXPERIMENTAL DESIGN

Each subject performed two separate tasks (prosody and lexicon). In the prosody task, the subjects judged whether binaurally presented auditory stimuli were IDS or ADS by focusing on only prosody while ignoring word content. In the lexicon task, the subjects judged whether the stimuli were IDS or ADS by focusing on the word while ignoring prosody. The participants were asked to respond as quickly and as accurately as possible after the recognition of each stimulus by pressing a button with either the left or right hand. Within each task, 128 words were

presented in quasi-random order. The order of the two tasks was counterbalanced across subjects (for additional details, see Matsuda et al., 2011).

### MRI SCANNING PROCEDURE

fMRI experiments were performed on a 4T Agilent whole-body MRI system (Agilent Inc., Santa Clara, CA, USA). Twenty-three axial slices (24-cm FOV,  $64 \times 64$  matrix, 5-mm thickness, 0-mm gap) parallel to the anterior commissure-posterior commissure (AC-PC) plane were acquired using a two-shot Echo Planar Imaging (EPI) pulse sequence (volume TR 2,600 ms, TE 25 ms, flip angle  $40^\circ$ ) for the two functional runs; each run consisted of 298 volumes. Prior to and between the functional runs, a set of high-resolution ( $1 \text{ mm}^3$ ) and low-resolution ( $1.72 \text{ mm}^3$ ) whole-brain anatomical images were acquired using a magnetization-prepared 3D FLASH (fast low-angle shot) pulse sequence (for additional details, see Matsuda et al., 2011).

### fMRI DATA ANALYSIS

EPI image reconstruction was followed by respiratory- and cardiac-fluctuation removal using a previously described retrospective estimation and correction method (Cheng et al., 2001). The data were then preprocessed and analyzed using the BrainVoyager QX software package (Brain Innovation, Maastricht, Netherlands).

After the transformation of each subject's registered functional images into the Talairach space (Talairach and Tournoux, 1988), whole-brain activation maps were obtained using a standard voxel-wise general linear model (GLM) at the single-subject level. To extract the brain regions involved in the perceptual processing of each speech sound (IDS or ADS), multiple regressors were generated by convolving a boxcar representation of each type of stimulus-presentation period prior to the subject response with a theoretical two-gamma hemodynamic response function with an event-related design. A second-stage random-effect analysis (RFX) was then performed across subjects. Activated clusters with less than  $281 \text{ mm}^3$  (4 voxels;  $1 \text{ voxel} = 3.75 \times 3.75 \times 5 \text{ mm}^3$ ) were omitted from our results (for additional details, see Matsuda et al., 2011).

To compare contrasted activations among the subject groups, an analysis of variance (ANOVA) was conducted with subject-wise contrast, i.e., a 2 (prosody type)  $\times$  2 (lexicon type) design, as the repeated-measures factor and group as the between-subject factor for the prosody or lexicon tasks. We had a strong a priori hypothesis regarding how different groups will show different activation patterns that would show a dynamic change corresponding to the interaction between emotion (prosody) and language (lexicon) depending on the child's developmental stage (Figure 1). Thus, we conducted a three-way ANOVA (2 (prosody type)  $\times$  2 (lexicon type)  $\times$  4 (group type)) for each task to determine the brain areas with a significant "interaction" ( $q < 0.05$ , False Discovery Rate or FDR correction for whole brain (Genovese et al., 2002)). Then, we confirmed whether each detected brain area (region of interest or ROI) satisfied our hypothesized patterns of activation (Figure 1) using *post hoc* comparisons between the subject groups. Because we did not

assume bias in the selection of each ROI, the *post hoc* comparisons are valid and different from a circular analysis (Kriegeskorte et al., 2009).

## RESULTS

### BEHAVIORAL RESULTS

We analyzed the reaction time and accuracy of judgment for IDS or ADS stimuli in the lexicon and prosody tasks for each participant. The reaction time was measured from the end point of each stimulus presentation to compensate for stimulus time variations. The percent correct data were arcsine transformed to stabilize variance when we conducted statistical analysis.

First, the difference of task difficulty was investigated by direct comparison between the lexicon and prosody tasks across subject groups and stimulus types. Analysis of the reaction time data revealed a significant difference between the tasks (one-way ANOVA, 2 tasks) ( $F_{(1,68)} = 20.02$ ,  $p < 0.001$ ), which indicates faster responses to the prosodic judgment (mean, 622 ms) compared with the lexical judgment (mean, 712 ms) across subject groups. The prosody information, e.g., voice pitch, was included from the initial part of the stimuli, and thus subjects could detect and discriminate the prosody component of the stimuli more quickly. However, the lexical information enabled subjects to understand meanings when they listen to each stimulus to the end, and thus the subject's reaction time might be slower than in the prosody-judgment task. There was no significant difference among the subject groups ( $F_{(3,65)} = 0.81$ ,  $p = 0.49$ ). Analysis of the accuracy revealed a subtle but significant difference between the tasks (one-way ANOVA, 2 tasks) ( $F_{(1,68)} = 5.01$ ,  $p < 0.05$ ), which indicates that although the participants responded more accurately in the lexicon task (mean, 98.2%) compared with the prosody task (mean, 97.5%), the stimuli in both tasks were easily recognized as IDS or ADS. There was no significant difference among the subject groups ( $F_{(3,65)} = 2.16$ ,  $p = 0.10$ ).

Next, we investigated subject-group differences in the reaction time and accuracy by considering the stimulus types. We conducted ANOVAs that treated stimulus types of *Lexicon* (IDS, ADS) and *Prosody* (IDS, ADS) as the repeated measure factors and *Experience* as the between-subject factor (i.e., 2 lexicon-types  $\times$  2 prosody-types  $\times$  4 groups) in the lexicon-judgment and prosody-judgment tasks (the lexicon-judgment task is shown in Table 1, and the prosody-judgment task shown in Table 2).

In the lexicon-judgment task, the reaction time did not show a significant main effect of subject group ( $F_{(3,65)} = 1.73$ ,  $p = 0.17$ ), but the accuracy did show a significant main effect ( $F_{(3,65)} = 3.59$ ,  $p < 0.05$ ). Slightly lower accuracy on the lexicon-judgment task in the non-mothers (mean, 96.8%) compared with the three groups of mothers (mean, 98.6%) may reflect the relatively little experience non-mothers have with IDS. There were no significant interactions between stimulus types and groups (i.e., 2 lexicon-types  $\times$  4 groups; 2 prosody-types  $\times$  4 groups; 2 lexicon-types  $\times$  2 prosody-types  $\times$  4 groups) for either reaction time or accuracy ( $F_{(3,65)} \leq 2.70$ ,  $p > 0.05$  for all cases).

In the prosody-judgment task, neither reaction time ( $F_{(3,65)} = 0.79$ ,  $p = 0.50$ ) nor accuracy ( $F_{(3,65)} = 0.58$ ,  $p = 0.63$ ) showed a significant main effect of subject group. There were no significant interactions between stimulus types and groups



**Table 1 | Performance in the lexicon task for each subject group.**

	Mean reaction time (SD) (msec)		Mean accuracy (SD) (%)	
	Lexical ADS	Lexical IDS	Lexical ADS	Lexical IDS
<b>Non-mothers (n = 15)</b>				
Prosodic ADS	652.8 (146.2)	652.8 (146.1)	97.9 (3.7)	93.5 (5.8)
Prosodic IDS	604.7 (200.0)	615.4 (147.7)	97.9 (3.3)	97.9 (3.7)
<b>Mothers with preverbal infants (n = 20)</b>				
	758.3 (195.5)	747.8 (203.2)	98.9 (1.8)	96.6 (4.2)
	704.2 (248.9)	724.0 (243.2)	98.5 (2.9)	98.6 (2.1)
<b>Mothers with toddlers in the two-word utterance stage (n = 16)</b>				
	788.3 (174.6)	787.8 (231.3)	99.2 (1.4)	98.2 (2.0)
	743.0 (200.9)	737.3 (217.9)	98.2 (3.2)	99.2 (1.8)
<b>Mothers with children in the first year of elementary school (n = 18)</b>				
	727.4 (130.7)	729.4 (137.2)	100 (0)	99.0 (2.1)
	682.8 (142.3)	689.4 (144.3)	99.3 (1.7)	98.3 (3.1)

Note: Reaction time was measured from the end point of each stimulus presentation to compensate for stimulus time variations. ADS: adult-directed speech, IDS: infant-directed speech, SD: standard deviation.

**Table 2 | Performance in the prosody task for each subject group.**

	Mean reaction time (SD) (msec)		Mean accuracy (SD) (%)	
	Prosodic ADS	Prosodic IDS	Prosodic ADS	Prosodic IDS
<b>Non-mothers (n = 15)</b>				
Lexical ADS	609.2 (146.2)	557.9 (150.8)	99.2 (1.9)	98.1 (2.0)
Lexical IDS	584.0 (185.7)	520.1 (141.3)	94.8 (3.0)	96.9 (4.3)
<b>Mothers with preverbal infants (n = 20)</b>				
	677.0 (225.9)	652.6 (224.5)	99.7 (1.0)	98.0 (3.1)
	659.7 (244.0)	650.8 (254.2)	96.1 (4.2)	97.8 (3.8)
<b>Mothers with toddlers in the two-word utterance stage (n = 16)</b>				
	670.8 (208.1)	634.2 (235.5)	99.0 (3.2)	98.8 (1.9)
	684.2 (226.1)	583.7 (227.6)	94.7 (5.3)	97.4 (3.6)
<b>Mothers with children in the first year of elementary school (n = 18)</b>				
	607.3 (143.6)	628.0 (168.5)	99.0 (2.1)	98.6 (2.4)
	597.2 (141.0)	590.7 (148.0)	95.0 (4.4)	97.0 (5.0)

Note: Reaction time was measured from the end point of each stimulus presentation to compensate for stimulus time variations. ADS: adult-directed speech, IDS: infant-directed speech, SD: standard deviation.

(i.e., 2 lexicon-types  $\times$  4 groups; 2 prosody-types  $\times$  4 groups; 2 lexicon-types  $\times$  2 prosody-types  $\times$  4 groups) for either reaction time or accuracy ( $F_{(3,65)} \leq 1.49$ ,  $p > 0.22$  for all cases) with the exception of an interaction between prosody-types and subject groups in reaction time (2 prosody-types  $\times$  4 groups, ( $F_{(3,65)} = 3.21$ ,  $p < 0.05$ )). Slightly different reaction times between IDS-vs. ADS-prosody judgment were observed in the prosody task in the mothers with school-age children (609 ms to IDS prosody; 602 ms to ADS prosody) compared with the other three groups (600 ms to IDS prosody; 647 ms to ADS prosody), which showed the mothers' heightened sensitivity to ADS prosody. These differences may reflect the age of the mothers, who were the eldest group, rather than the experience with their children.

Behavioral results of the reaction time were essentially unchanged after removal of outliers beyond 2  $\times$  standard deviation (SD) of the individual participant's reaction-time distributions in each group (2 participants; one of mothers of preverbal infants and one of mothers with school-age children) (Ratcliff,

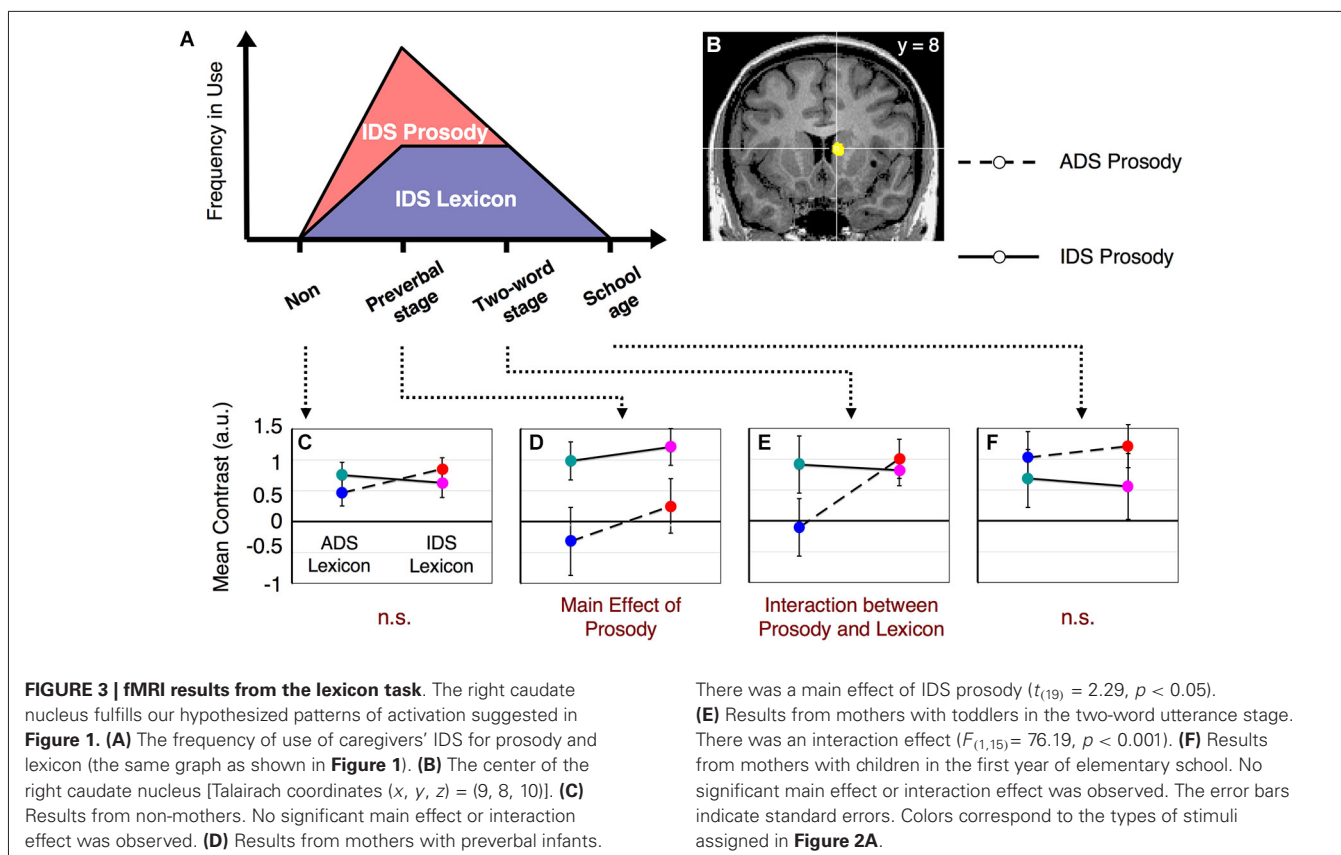
1993). There were no outliers in the behavioral data of the accuracy.

Overall, our behavioral results indicate that attentional levels were approximately the same among the four subject groups.

## FUNCTIONAL IMAGING RESULTS

To determine the brain sites that are specifically involved in the IDS vs. ADS evaluation of word meaning and prosody, we contrasted the lexicon and prosody tasks directly. This contrast failed to reveal significant differences, which suggests that both tasks recruit topographically similar brain structures (data not shown).

Main effects and interactions of both the lexicon and prosody tasks in each subject group have been previously reported (Matsuda et al., 2011). In this study, we sought to identify the neural substrates that showed a dynamic change corresponding to the interaction between prosody-type and lexicon-type stimuli depending on the child's developmental stage (Figure 1). By conducting a 2 (lexicon type)  $\times$  2 (prosody type)  $\times$  4 (group

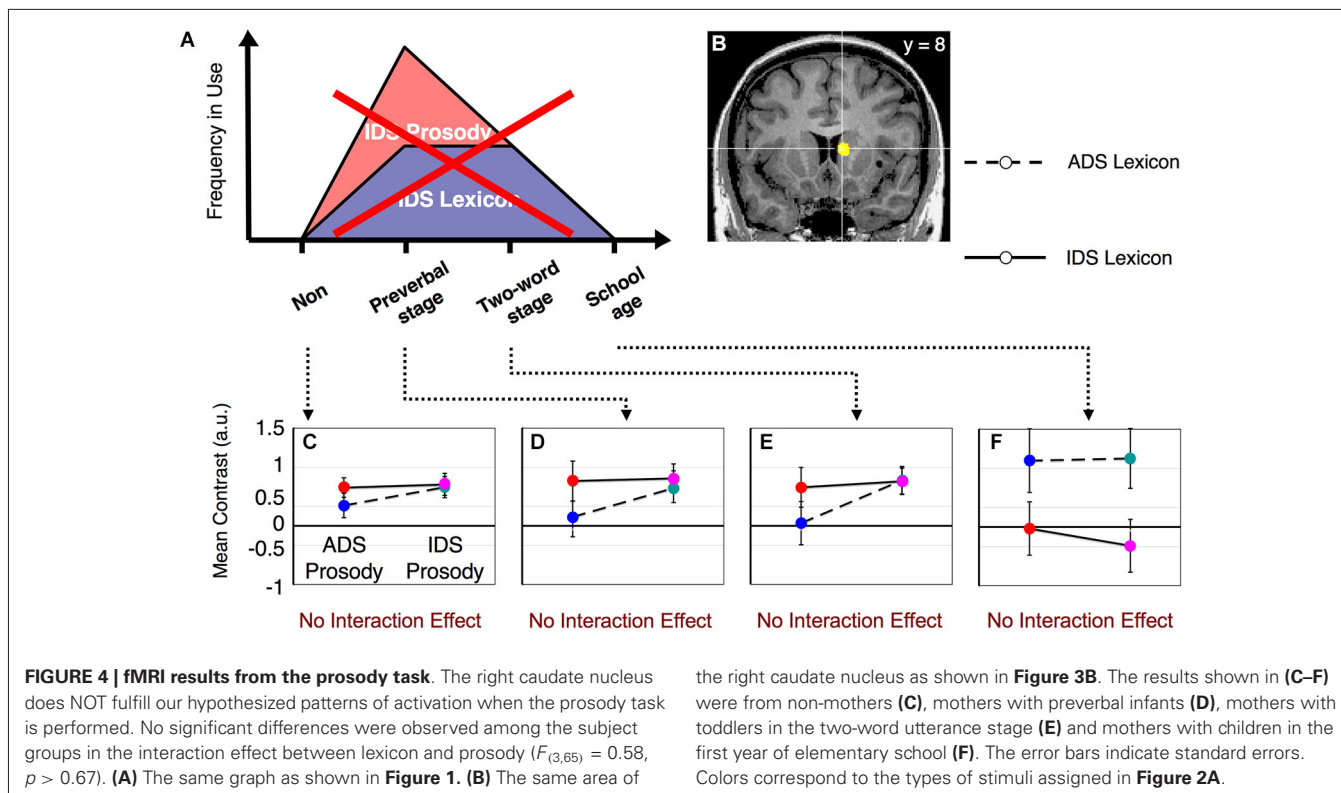


type) three-way ANOVA in the lexicon-judgment and prosody-judgment tasks, we identified three brain areas with significant interactions ( $q < 0.05$ , FDR) in the lexicon-judgment task, but none in the prosody-judgment task. These areas included the right caudate nucleus [Talairach coordinates (x, y, z): 9, 8, 10], the medial prefrontal cortex (mPFC) [5, 37, 9], and the right dorsolateral prefrontal cortex [35, 46, 17]. Within the three different areas, only the right caudate nucleus satisfied our hypothesized patterns of the experience-dependent behavior (Figure 3B). *Post hoc* comparisons among subject groups showed that the right caudate nucleus met the following four conditions in the lexicon-judgment task, i.e., a 2 (lexicon type)  $\times$  2 (prosody type) ANOVA for each subject group (Figures 3C–F): (1) non-mothers showed neither significant main effects (lexicon-type (IDS vs. ADS) and prosody-type (IDS vs. ADS)) nor interaction (2 (lexicon type)  $\times$  2 (prosody type)) ( $F_{(1,14)} < 3.19, p > 0.05$  for all cases); (2) mothers with preverbal infants showed main effects of prosody type ( $F_{(1,19)} = 4.39, p < 0.05$ ; mean of IDS prosody = 1.10, mean of ADS prosody =  $-0.03$ ), and there were no significant main effects of lexicon type or interaction (2 (lexicon type)  $\times$  2 (prosody type)); (3) mothers with toddlers at the two-word stage showed interaction effects between lexicon and prosody types, where enhanced activities were evoked by either the lexical or prosodic IDS or by the combination of the two ( $F_{(1,15)} = 76.19, p < 0.001$ ; mean of IDS lexicon and IDS prosody = 0.82, mean of IDS lexicon and ADS prosody = 1.01, mean of ADS lexicon and IDS prosody = 0.91, mean of ADS lexicon and ADS prosody =  $-0.11$ ); and

(4) mothers with school-age children showed no significant main effects (lexicon-type (IDS vs. ADS) and prosody-type (IDS vs. ADS)) nor interaction (2 (lexicon type)  $\times$  2 (prosody type)) ( $F_{(1,17)} < 0.77, p > 0.39$  for all cases). This experience-dependent behavior in the caudate nucleus was observed in the lexicon task (explicit-lexicon and implicit-prosody processing), but not in the prosody task (explicit-prosody and implicit-lexicon processing). Among the subject groups, the right caudate nucleus did not show significant differences in the interaction effect between lexicon and prosody types in the prosody task (a 2 (lexicon type)  $\times$  2 (prosody type)  $\times$  4 (group type) ANOVA,  $F_{(3,65)} = 0.58, p = 0.67$ ) (Figure 4). The mothers with school-age children showed a tendency toward different activity of lexicon type (IDS lexicon vs. ADS lexicon) compared with the other subject groups, but this finding was not significant among the subject groups (a 2 (lexicon type)  $\times$  4 (group type) ANOVA,  $F_{(3,65)} = 0.58, p = 0.67$ ).

Although the medial and dorsolateral prefrontal cortices showed a significant interaction between lexicon, prosody and subject groups (a 2 (lexicon type)  $\times$  2 (prosody type)  $\times$  4 (group type) ANOVA) in the lexicon task, these areas did not meet our hypothesized patterns of activation (Figure 1). The main effect and the interaction were not significant in each subject group ( $F_{(3,65)} < 2.09, p > 0.1$  for all cases).

To summarize, these results indicate that the ongoing use of IDS is reflected in the explicit lexicon and implicit prosody processing of IDS in the right caudate nucleus.



## DISCUSSION

We determined that the right caudate nucleus is a neural substrate for the experience-dependent changes in the interaction between language and emotion in IDS processing. The task, which is close to the naturalistic environment of IDS usage, activated the right caudate nucleus in an experience-dependent manner as opposed to the control, less naturalistic condition, i.e., the right caudate nucleus is sensitive to the ongoing use of IDS components (prosody, lexicon or both), and provide neurophysiological evidence supporting IDS as a flexible, goal-directed behavior in social communication.

It is of interest to note that even though all participants performed the tasks in similar fashions (Tables 1, 2), the blood oxygen level dependent (BOLD) signals differed dramatically among the subject groups. This observation suggests that signal differences were not the result of difficulties in performing the tasks (e.g., delayed response or low accuracy) (Binder et al., 2004), but rather they were caused by the increase in sensitivity in the caudate nucleus.

While simple auditory perception and subsequent IDS or ADS judgment were required as experimental tasks, we could indirectly evoke the participants' motor-related response to IDS by auditory observation of the IDS stimuli, which were uttered by another caregiver. We previously showed that a mother's speech-related motor circuit is recruited by the perception of IDS in an experience-dependent manner (Matsuda et al., 2011). The observation of specific actions performed by other individuals appears to be linked to a covert simulation of well-trained motor activation. This fine-tuned sensorimotor "resonance"

system develops as a consequence of extensive motor practice (Aglioti et al., 2008). Recent studies using transcranial magnetic stimulation demonstrated that stimulation of the production-related motor cortex resulted in an altered sensory-perception performance (Gough et al., 2005; Meister et al., 2007; D'Ausilio et al., 2009). This finding confirms the idea that the neural circuitry in motor areas is functionally and directly connected to perception processing. Auditory perception of IDS was expected to implicate the mothers' motor system for IDS. Therefore we employed this use-dependent resonance effect as a functional probe of the motor response to IDS.

### IDS PROCESSING AS A GOAL-DIRECTED BEHAVIOR

IDS is the speech-related motor response to an infant, and its goal is to encourage and communicate with infants. Caregivers flexibly regulate their own IDS to obtain a sufficient reaction (outcome) from developing infants, and hence, the contents of IDS dynamically change through the modification of the ratio between the linguistic and emotional components. This flexibility makes IDS distinct from a habit system of stimulus–response learning or Pavlovian conditioning, which involves passive learning of stimulus–outcome associations. In the scheme of a goal-directed action with stimulus–response–outcome association learning in the mother–infant interplay, we revealed experience-dependent changes in IDS processing in the interaction between prosody and lexicon in the right caudate nucleus. The neural representations of the ongoing use and regulation of IDS are consistent with the known functions in this area, i.e., flexible goal-directed actions (Yin et al., 2005a,b; Yin and Knowlton, 2006). Beyond

the arbitrary association learning of stimulus–response–outcome conditions identified in human (Valentin et al., 2007; Xue et al., 2008; de Wit et al., 2009) and animal studies (Miyachi et al., 1997, 2002; Yin et al., 2005a,b, 2009), our results demonstrate that the caudate nucleus is involved in naturalistic goal-directed behavior in social communication, such as the mother–infant interaction.

Previous studies have shown that during motor and procedural learning, changes in neural activity occur in the striatum, which is the major input center of the basal ganglia (Jenkins et al., 1994; Doyon et al., 1996; Carelli et al., 1997; Ungerleider et al., 2002; Brasted and Wise, 2004; Barnes et al., 2005). Several studies have also shown that different striatal circuits and processes are engaged during the early and late phases of skill learning (Miyachi et al., 1997, 2002; Costa et al., 2004; Yin et al., 2009). The dorsomedial or associative striatum (the caudate nucleus in primates), which receives input primarily from the association cortices, such as the PFC (McGeorge and Faull, 1989; Voorn et al., 2004), appears to be preferentially involved in the initial stages of visuomotor learning and during the rapid acquisition of action–outcome contingencies (Miyachi et al., 1997, 2002; Yin et al., 2005b). However, the dorsolateral or sensorimotor striatum (the putamen in primates), which receives inputs from the sensorimotor cortex (McGeorge and Faull, 1989; Voorn et al., 2004), is critical for the more gradual acquisition of habitual and automatic behaviors (Miyachi et al., 1997, 2002; Yin et al., 2004). Our results observed in the caudate nucleus are consistent with its function in the early phases of skill learning, i.e., flexible action–outcome contingencies in goal-directed behavior.

Our auditory observation paradigm of IDS may enable caregivers to recall the outcome (infant's reaction) indirectly rather than having a direct and immediate outcome. The anticipation of peak emotion (or “chills”) activates the right caudate nucleus when human subjects listen to their own favorite music, while the experience of the peak activates the nucleus accumbens (Salimpoor et al., 2011). In our study, the experience-dependent activation in the right caudate nucleus was observed only in the lexicon task that required the explicit processing of lexicon and implicit processing of prosody of IDS. This lexicon task is more comparable to the naturalistic environment of IDS perception compared with the prosody task (explicit-prosody and implicit-lexicon processing of IDS). As caregivers are known to address infants with a conscious conversational and sub-conscious emotional speech style in IDS (Snow and Ferguson, 1977), the resemblance of the lexicon task to naturalistic IDS may easily induce the anticipation of the outcome, i.e., the infant's reaction. The results further suggest that the right caudate nucleus represents experience-dependent changes in the outcome anticipation for caregivers, which co-varies with the speech style of IDS along with child development.

However, there might be alternative possibilities other than goal-directed behavior and/or prolonged sensorimotor tuning, but simply a different strategy for performing the tasks. One alternative possibility is that mothers with IDS experience automatically anticipated the infant's reaction. The right caudate activation might be attributable to some subject groups imagining infants' reactions and others not doing so. However, our

fMRI results of group comparisons did not show our hypothesized pattern of activity (**Figure 1**) in either the visual areas or memory-related areas (e.g., hippocampus), which indicate that there were no differences among subject groups in imaging and/or remembering of infants' reaction by the auditory stimuli. One may further claim that even non-mothers would also activate the right caudate nucleus if they were asked to imagine an infant's reaction. We exclude the possibility of the readiness of imaging an infant's reaction by auditory stimuli through parental experience. This conclusion follows from our finding that non-mothers and mothers with school-age children showed similar responses in the right caudate nucleus (**Figure 3**). Another possibility is a group difference in threshold to suppress the prosody information, which was spoken by another mother. Mothers with preverbal infants or toddlers frequently use IDS prosody in their own way including individual differences, and thus the auditory observation of other's IDS prosody might interfere and by necessity suppress the processing of their own IDS to properly conduct the lexicon-judgment task with emotionally charged stimuli. If the suppression threshold was lower, the reaction time should be slower in these subject groups, but this was not the case. As we mentioned in the results section for the behavioral data, the between-subject comparison did not show a significant interaction with stimulus-type differences in the lexicon-judgment task.

#### ASYMMETRICAL INTERACTION BETWEEN LANGUAGE AND EMOTION IN IDS

Notably, the right caudate nucleus showed distinct activity that depended on the tasks that the subjects performed. The caudate nucleus exhibited experience-*dependent* processing in the lexicon task, while it displayed experience-*independent* processing in the prosody task. These differences between the two tasks suggest an asymmetrical interaction between lexical and prosodic processing in IDS, i.e., prosodic interference to lexical processing differs from lexical interference to prosodic processing.

Our results may lead us to reconsider the three generally accepted functions of IDS (Grieser and Kuhl, 1988; Cooper et al., 1997; Singh et al., 2002). First, IDS may attract and maintain infants' attention. Second, it may communicate positive emotion between a caregiver and an infant. Third, it may facilitate language acquisition. These three functions may coexist during infant development, with the attentional and emotional functions dominating during early infancy, and the linguistic function gaining importance as the infant progresses in the language learning process (Song et al., 2010). Our study suggests that this dynamic shift in IDS functions results from explicit processing of language communication rather than the intended emotional communication. As revealed in this study, only explicit language processing (i.e., the lexicon task) enabled this shift in the course of the infant's development. We contend that the emotional function during early infancy is an implicit by-product induced by the intention to produce language communication with preverbal infants, which is, in effect, an explicit goal-directed action of caregivers.

#### ACKNOWLEDGMENTS

This work was supported, in part, by funding from the Japan Science and Technology Agency, Exploratory Research for Advanced



Technology, Okanoya Emotional Information Project, by a Grant-in-Aid for Young Scientists (B) to Y.M. (KAKENHI 21720156), and by MEXT KAKENHI “Constructive Developmental Science” (24119004). We thank T. Asamizuya and C. Suzuki for their help in data acquisition and M. Hasegawa and Y. Sogabe for subject recruitment.

## REFERENCES

- Aglioti, S. M., Cesari, P., Romani, M., and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.* 11, 1109–1116. doi: 10.1038/nn.2182
- Amano, S., Nakatani, T., and Kondo, T. (2006). Fundamental frequency of infants’ and parents’ utterances in longitudinal recordings. *J. Acoust. Soc. Am.* 119, 1636–1647. doi: 10.1121/1.2161443
- Barker, B. A., and Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition* 94, B45–B53. doi: 10.1016/j.cognition.2004.06.001
- Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., and Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437, 1158–1161. doi: 10.1038/nature04053
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., and Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301. doi: 10.1038/nn1198
- Brasted, P. J., and Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur. J. Neurosci.* 19, 721–740. doi: 10.1111/j.0953-816x.2003.03181.x
- Burnham, D., Kitamura, C., and Vollmer-Conna, U. (2002). What’s new, pussycat? On talking to babies and animals. *Science* 296, 1435. doi: 10.1126/science.1069587
- Carelli, R. M., Wolske, M., and West, M. O. (1997). Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *J. Neurosci.* 17, 1804–1814.
- Cheng, K., Waggoner, R. A., and Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* 32, 359–374. doi: 10.1016/s0896-6273(01)00477-9
- Cooper, R. P., Abraham, J., Berman, S., and Staska, M. (1997). The development of infants’ preference for motherese. *Infant Behav. Dev.* 20, 477–488. doi: 10.1016/s0163-6383(97)90037-0
- Cooper, R. P., and Aslin, R. N. (1990). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Dev.* 65, 1663–1677. doi: 10.1111/j.1467-8624.1994.tb00841.x
- Costa, R. M., Cohen, D., and Nicolelis, M. A. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. *Curr. Biol.* 14, 1124–1134. doi: 10.1016/j.cub.2004.06.053
- D’Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., and Fadiga, L. (2009). The motor somatotopy of speech perception. *Curr. Biol.* 19, 381–385. doi: 10.1016/j.cub.2009.01.017
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015. doi: 10.1126/science.1077066
- de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., and Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 29, 11330–11338. doi: 10.1523/jneurosci.1639-09.2009
- Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., and Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *Eur. J. Neurosci.* 8, 637–648. doi: 10.1111/j.1460-9568.1996.tb01249.x
- Fernald, A. (1991). Prosody in speech to children: Prelinguistic and linguistic functions. *Ann. Child Dev.* 8, 43–80.
- Garnica, O. K. (1977). “Some prosodic and paralinguistic features of speech to young children,” in *Talking to Children: Language Input and Acquisition*, eds C. E. Snow and C. Ferguson (New York: Cambridge University Press), 63–88.
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878. doi: 10.1006/nimg.2001.1037
- Gough, P. M., Nobre, A. C., and Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *J. Neurosci.* 25, 8010–8016. doi: 10.1523/jneurosci.2307-05.2005
- Grieser, D. L., and Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Dev. Psychol.* 24, 14–20. doi: 10.1037//0012-1649.24.1.14
- Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., and Passingham, R. E. (1994). Motorsequence learning: a study with positron emission tomography. *J. Neurosci.* 14, 3775–3790.
- Kajikawa, S., Amano, S., and Kondo, T. (2004). Speech overlap in Japanese mother-child conversations. *J. Child Lang.* 31, 215–230. doi: 10.1017/S0305000903005968
- Kitamura, C., and Burnham, D. (2003). Pitch and communicative intent in mother’s speech: adjustments for age and sex in the first year. *Infancy* 4, 85–110. doi: 10.1207/s15327078in0401\_5
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686. doi: 10.1126/science.277.5326.684
- Kuhl, P. K., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Ann. Rev. Neurosci.* 31, 511–534. doi: 10.1146/annurev.neuro.30.051606.094321
- Matsuda, Y.-T., Ueno, K., Waggoner, R. A., Erickson, D., Shimura, Y., Tanaka, K., et al. (2011). Processing of infant-directed speech by adults. *Neuroimage* 54, 611–621. doi: 10.1016/j.neuroimage.2010.07.072
- Mazuka, R., Kondo, T., and Hayashi, A. (2008). “Japanese mothers’ use of specialized vocabulary in infant-directed speech: infant-directed vocabulary in Japanese,” in *The Origins of Language*, ed N. Masataka (Tokyo: Springer-Verlag), 39–58.
- McGeorge, A. J., and Faull, R. L. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29, 503–537. doi: 10.1016/0306-4522(89)90128-0
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr. Biol.* 17, 1692–1696. doi: 10.1016/j.cub.2007.08.064
- Minagawa-Kawai, Y., Matsuoka, S., Dan, I., Naoi, N., Nakamura, K., and Kojima, S. (2009). Prefrontal activation associated with social attachment: facial-emotion recognition in mothers and infants. *Cereb. Cortex* 19, 284–292. doi: 10.1093/cercor/bhn081
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321. doi: 10.1523/jneurosci.1984-06.2007
- Miyachi, S., Hikosaka, O., and Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp. Brain Res.* 146, 122–126. doi: 10.1007/s00221-002-1213-7
- Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., and Rand, M. K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.* 115, 1–5. doi: 10.1007/pl00005669
- Ogura, T. (2006). “How the use of ‘non-adult words’ varies as a function of context and children’s linguistic development,” in *Studies in Language Science* (5), ed M. Nakayama (Tokyo: Kuroshio Shuppan), 103–120.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, E., et al. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Nat. Acad. Sci. U S A* 100, 11702–11705. doi: 10.1073/pnas.1934290100
- Penman, R., Cross, T., Milgrom-Friedman, J., and Meares, R. (1983). Mothers’ speech to prelingual infants: a pragmatic analysis. *J. Child Lang.* 10, 17–34. doi: 10.1017/s0305000900005109
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532. doi: 10.1037/0033-2909.114.3.510
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., et al. (2013). Motherese in interaction: at the cross-road of emotion and cognition? (A systematic review). *PLoS One* 8:e78103. doi: 10.1371/journal.pone.0078103
- Saito, Y., Aoyama, S., Kondo, T., Fukumoto, R., Konishi, N., Nakamura, K., et al. (2007). Frontal cerebral blood flow change associated with infant-directed

- speech. *Arch. Dis. Child. Fetal Neonatal Ed.* 92, F113–F116. doi: 10.1136/adc.2006.097949
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14, 257–262. doi: 10.1038/nn.2726
- Singh, L., Morgan, J. L., and Best, C. T. (2002). Infants' listening preferences: baby talk or happy talk? *Infancy* 3, 365–394. doi: 10.1207/s15327078in0303\_5
- Snow, C. E., and Ferguson, C. A. (1977). *Talking to Children: Language Input and Acquisition*. New York: Cambridge UP.
- Soderstrom, M. (2007). Beyond babytalk: re-evaluating the nature and content of speech input to preverbal infants. *Dev. Rev.* 27, 501–532. doi: 10.1016/j.dr.2007.06.002
- Song, J. Y., Demuth, K., and Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *J. Acoust. Soc. Am.* 128, 389–400. doi: 10.1121/1.3419786
- Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Med.
- Taoumepeau, M., and Ruffman, T. (2008). Stepping stones to others' minds: maternal talk relates to child mental state language and emotion understanding at 15, 24 and 33 months. *Child Dev.* 79, 284–302. doi: 10.1111/j.1467-8624.2007.01126.x
- Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71. doi: 10.1207/s15327078in0701\_5
- Trainor, L. J., Austin, C. M., and Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11, 188–195. doi: 10.1111/1467-9280.00240
- Ungerleider, L. G., Doyon, J., and Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiol. Learn. Mem.* 78, 553–564. doi: 10.1006/nlme.2002.4091
- Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026. doi: 10.1523/jneurosci.0564-07.2007
- Vallabha, G. K., McClelland, J. L., Pons, E., Werker, J. F., and Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proc. Natl. Acad. Sci. U S A* 104, 13273–13278. doi: 10.1523/jneurosci.0564-07.2007
- Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., and Pennartz, C. M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* 27, 468–474. doi: 10.1016/j.tins.2004.06.006
- Watamaki, T., and Ogura, T. (2004). *The Japanese MacArthur Communicative Development Inventory: Words and Grammar*. Kyoto: Kyoto International Social Welfare Exchange Center.
- Werker, J. F., and McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43, 230–246. doi: 10.1037/h0084224
- Xue, G., Ghahremani, D. G., and Poldrack, R. A. (2008). Neural substrates for reversing stimulus-outcome and stimulus-response associations. *J. Neurosci.* 28, 11196–11204. doi: 10.1523/jneurosci.4001-08.2008
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nnr1919
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy, but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189. doi: 10.1111/j.1460-9568.2004.03095.x
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512. doi: 10.1111/j.1460-9568.2005.04219.x
- Yin, H. H., Mulcare, S. P., Hilário, M. R., Clouse, E., Holloway, T., Davis, M. I., et al. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat. Neurosci.* 12, 333–341. doi: 10.1038/nn.2261
- Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523. doi: 10.1111/j.1460-9568.2005.04218.x
- Zangl, R., and Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy* 11, 31–62. doi: 10.1207/s15327078in1101\_2
- Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., et al. (2011). Neural coding of formant-exaggerated speech in the infant brain. *Dev. Sci.* 14, 566–581. doi: 10.1111/j.1467-7687.2010.01004.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 June 2014; accepted: 23 October 2014; published online: 10 November 2014.

Citation: Matsuda Y-T, Ueno K, Cheng K, Konishi Y, Mazuka R and Okanoya K (2014) Auditory observation of infant-directed speech by mothers: experience-dependent interaction between language and emotion in the basal ganglia. *Front. Hum. Neurosci.* 8:907. doi: 10.3389/fnhum.2014.00907

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Matsuda, Ueno, Cheng, Konishi, Mazuka and Okanoya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.