Supplementary Materials

Supplementary Materials S1: Sentence stimuli

Each of the sentences listed below (along with their content category) was presented once per fMRI run. A randomly chose half of the sentences was presented in neutral tone during neutral blocks, the other half was presented in emotional (happy, sad, or angry) tone during the emotion blocks. To render the emotions, a team member with acting training (C.E.C.) imagined a situation in which a child might be happy, sad, or angry while uttering each sentence. For example, "We are visiting grandma until Sunday" could make a child very happy (if grandma is a beloved and rarely seen family member), sad (if the visit means that they child will miss out on another weekend activity they would otherwise have been able to attend) or even angry (if that alternative activity would have been much preferred and visiting grandma is not something the child particularly likes to do).
For the control conditions, the same sentences were played in reverse, rendering them incomprehensible.

1)	My sister bought me a shirt with green stripes.	gift
2)	Uncle Ted sent me a chemistry set.	gift
3)	Dad always gives me books for my birthday.	gift
4)	Maybe Santa will bring me a new bike.	gift
5)	The tooth fairy left me two dollars last night.	gift
6)	Mom will buy me red rainboots.	gift
7)	I got three cards on Valentine's day.	gift
8)	Grandma is knitting me a pair of socks.	gift
9)	Mom poured me a huge glass of milk.	food
10)	Grandpa put green olives in the salad.	food
11)	Dad cooked potroast for dinner last night.	food
12)	On Mondays, dad makes broccoli casserole.	food
13)	Grandma serves spinach whenever we visit.	food
14)	We eat fish for dinner every Friday.	food
15)	She cut the crusts off of my sandwich.	food
16)	We had oatmeal this morning.	food
17)	My uncle showed us around his farm.	trip
18)	Grandpa took me to the library.	trip
19)	We camp in the mountains every weekend.	trip
20)	We spent the weekend at our beach house.	trip
21)	We are visiting grandma until Sunday.	trip
22)	Last summer my family drove to Vermont.	trip
23)	Mom takes me to the museum on Sundays.	trip
24)	We'll be at the zoo all day tomorrow.	trip

To investigate whether the target emotion was recognizable in the forward version of the sentence, and whether it was still recognizable in the reverse version, we ran a validation study

with N=30 young adults. This sample was not identical with the participants of the MRI study, but drawn from the same cohort.
Across the three fMRI runs, we presented 36 emotional (12 “happy”, 12 “sad”, 12 ”angry”) and 36 neutral (12 “food”, 12 “gift” and 12 “trip”) renderings, and their reverse versions. In the validation study, we presented these same stimuli in random order and asked participants whether they perceived them as happy, sad, angry, or neutral.

For each emotional forward speech sentence, the majority of the validation sample chose the target emotion. In fact, the lowest agreement for any forward stimulus was at 67%, far above the chance level of 25%, identification was perfect across all subjects and sentences for happy and sad stimuli, and average accuracy for all emotional forward sentences was at 95%. This indicates that our prosody manipulation worked as intended.

At 79% correct, average accuracy was significantly lower for reverse than for forward emotional speech stimuli, but still exceeded chance level significantly. This indicates that playing the stimuli in reverse degraded, but did not fully remove recognizable emotional prosody cues. Thus, implicit emotion recognition might have happened to some extent in the EmotRev condition, even though the task did not require participants to pay attention to emotional prosody in that condition.

Table S1: Recognizability of emotions in forward and reverse emotional stimuli

	Forward stimuli		Reverse stimuli		Paired t-test comparing forward and reverse stimuli
	Mean (SD) accuracy	t-test against chance (25%)	Mean (SD) accuracy	t-test against chance (25%)	
overall emotional	95 (4)	t(29)=87, p<0.00001	79 (9)	t(29)=31, p<0.00001	t(29)=8.4, p<0.00001
neutral	94 (9)	t(29)=43 , p<0.00001	76 (18)	t(29)=15 , p<0.00001	t(29)=5.7, p<0.00001

A Confusion matrix for forward stimuli:

observed response \ target response	happy	sad	angry	neutral
happy	100	0	13	1
sad	0	100	1	1
angry	0	0	85	5
neutral	0	0	1	93

B Confusion matrix for reverse stimuli:

observed response \ target response	happy	sad	angry	neutral
happy	64	0	10	4
sad	9	100	1	2
angry	11	0	72	15
neutral	15	0	17	79

Figure S1: Confusion matrices for forward (A) and reverse (B) stimuli. For all emotions, the majority of the responses coincided with the intended target emotion. In the forward stimuli, the most frequent error was that stimuli intended to sound angry were mistaken for happy. In the reverse stimuli, intended angry stimuli were most frequently misclassified as neutral or happy. Note, however, that accuracy differences between different emotions are irrelevant with regard to the fMRI results, as all emotions were presented interleaved in each block of each condition.

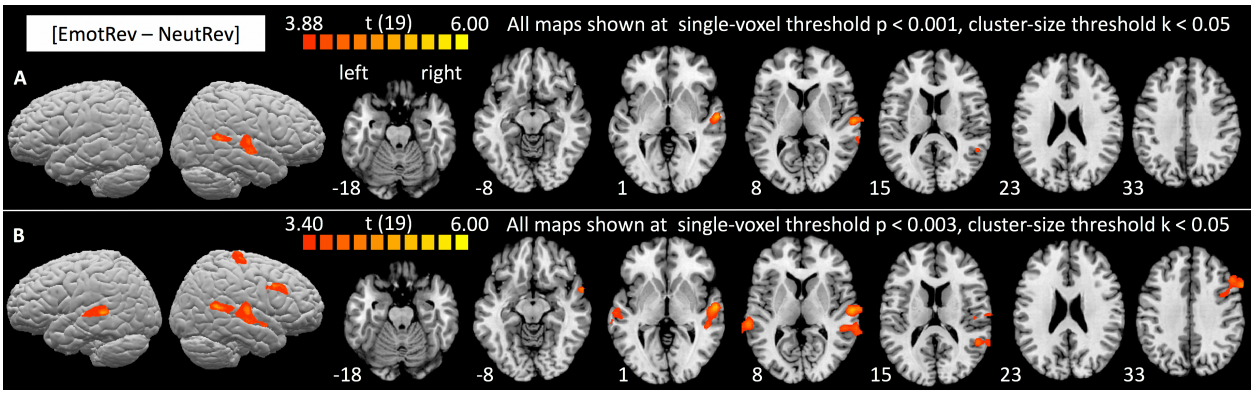

Figure S2. Areas of stronger activation for emotional than neutral reversed speech.
The strongest effects were in right temporal cortex (A), in line with the notion of a right-hemisphere preference for the suprasegmental acoustic features carrying the prosody signal. A slightly more lenient threshold (B) also revealed an effect in left temporal and right inferior frontal cortex, similar to the activation pattern observed for contrast [Emot – Neut] (see Figure 2B). We postulate that these activations reflect automatic processing of emotion in condition EmotRev, in line with the results of the stimulus validation study (see supplementary materials S1 above) showing that the emotion signal was present and recognizable in the emotional reverse speech stimuli. See supplementary Table S2 for cluster labels and peaks.

Table S2. Areas of stronger activation for emotional than neutral reversed speech.

Contrast	Cluster location	Peak Talairach coordinates x,y,z *	Peak t-value
EmotRev – NeutRev (at p < 0.001)	Right middle and anterior STS, BAs 41, 42, 22	54, -13, -2	6.35
	Right posterior superior temporal sulcus, BA 22	63, -43, 10	5.34
additionally at p < 0.003	Left posterior superior temporal sulcus, BA 22 (extending into 21)	-60, -31, 7	4.78
	Right inferior frontal cortex, BA 44 (extending into BA 9)	57, 17, 37	5.61

It may come as a surprise that this contrast did not result in strong activation in primary auditory cortex, considering that other studies have found such activation when contrasting emotional and neutral prosody stimuli (see Witteman et al., 2012, Table 1 and Figure 1, for a meta-analysis). We suspect that this is because reverse neutral speech is a stronger driver of right auditory cortex activation than forward neutral speech, possibly because its incomprehensibility and "strangeness" (a human voice producing sounds that are impossible to generate naturally) make it more salient. In line with this speculation, we did indeed observe a significantly higher signal change relative to baseline in right BA 41 for NeutRev than for Neut.

In this context, it is also important to note that contrast [EmotRev – NeutRev] does not capture stimulus-driven emotional prosody processing. While it does capture the higher amplitude and pitch modulations associated with emotional speech, which are present in forward (Emot) and reverse emotional stimuli (EmotRev) alike, it does not fully capture activations associated with identification and interpretation of natural-sounding prosodic features (e.g., specific pitch contours), as those are distorted in the reverse speech stimuli.

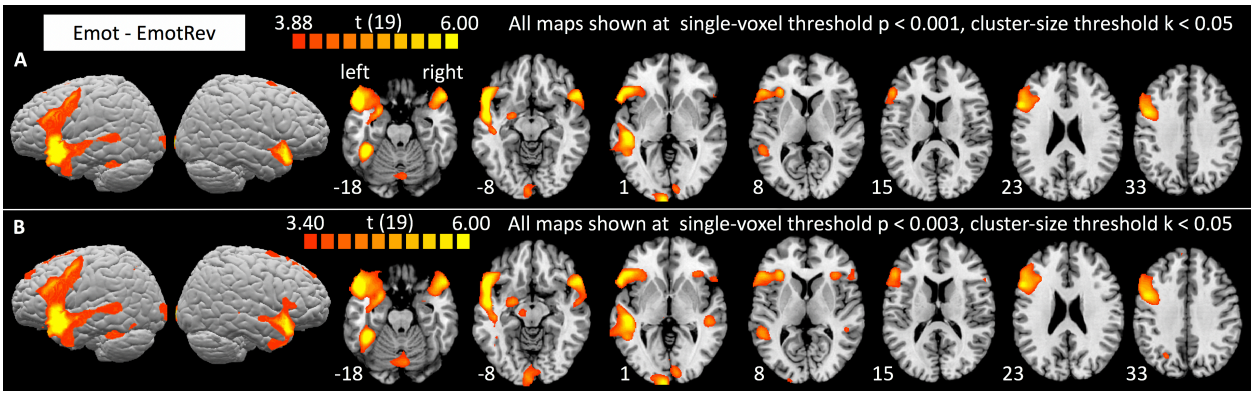

Figure S3. Areas of stronger activation for forward compared to reverse emotional speech. This contrast captures activations due to automatic sentence comprehension and intentional emotion processing, but not activations driven by the stronger pitch and amplitude modulations of emotional speech (as those are also present in the reverse speech stimuli). (A) At a single-voxel threshold of $p < 0.001$, this contrast activated the same left-hemisphere areas as sentence comprehension contrast (Neut – NeutRev), indicating that sentence comprehension took place automatically even if the task did not require it. In addition, it activated right anterior temporal cortex (possibly the site of a "prosody contour lexicon" used to interpret the intact pitch contours that are present in condition Emot, but not EmotRev.) (B) Interestingly, the right superior temporal and inferior frontal activations observed for contrast (Emot – Neut) were only observed for this contrast at a slightly more lenient single-voxel threshold, and at that threshold, they were also observed for contrast (EmotRev – NeutRev) (see supplementary Figure S2 above). This suggests that automatic emotion processing (as captured by [EmotRev – NeutRev]) contributed as much or more to right fronto-temporal activations as intentional, task-driven emotion processing (as captured here). See supplementary Table S3 for cluster labels and peaks.

Table S3. Areas of stronger activation for emotional than neutral reversed speech.

Contrast	Cluster location	Peak Talairach coordinates x,y,z *	Peak t-value
Emot – EmotRev (at $p < 0.001$)	Left superior/middle temporal cortex, BAs 22, 21, 37	-48, -27, 1	7.36
	Left fusiform gyrus, BAs 37, 36	-39, -43, -17	7.78
	Left anterior temporal cortex, BA 38	-48, 14, -17	11.11
	Left inferior frontal cortex, BAs 6, 9, 44, 45, 46, 47	-51, 17, 26	7.05
	Left amygdala	-27, -3, -13	5.70
	Left anterior insula	-31, 25, 7	5.89
	Midline occipital cortex, BA 18	-6, -101, 1	6.52
	Midline frontal cortex, BA 9, 6	3, 5, 58	6.10
	Right anterior temporal cortex, BA 38 extending into inferior frontal BA 47	54, 17, -8	6.35
	Right cerebellum	15, -67, -29	5.85
additionally at $p < 0.003$	Left interior parietal cortex, BAs 39, 7	-27, -58, 42	5.42
	Right amygdala	25, -1, -14	4.69
	Right anterior insula	32, 24, 7	4.72
	Right BAs 45, 47	52, 20, 10	4.09
	Right STS, BAs 22, 21	45, -31, 2	4.48

* When activations merged into one big cluster at $p < 0.001$, we determined these peaks by increasing the threshold until the clusters separated, allowing identification of local maxima.

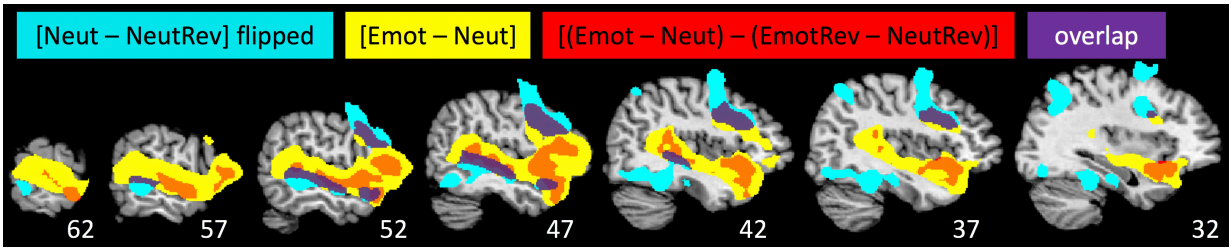

Figure S4. Spatial similarity of activation maps.
When flipped into the right hemisphere, the sentence comprehension activation map (blue) overlapped with the activation maps for emotional prosody (yellow = including acoustic difference effects; red = excluding acoustic difference effects) in lateral temporal and inferior frontal cortex. Maps were binarized to facilitate identification of overlapping areas (purple).

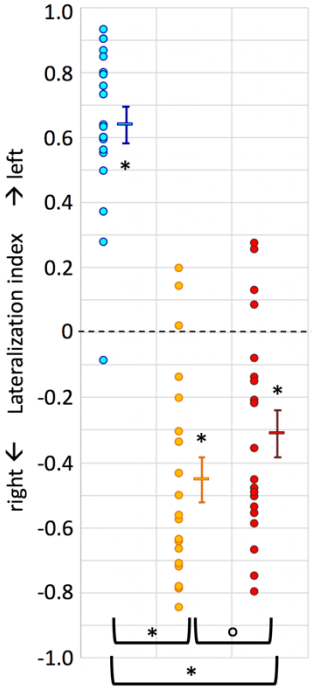

Figure S5. Lateralization in whole-hemisphere ROIs.
Lateralization indices comparing LH and RH activation across the entire hemispheres for sentence comprehension (blue) and emotional prosody processing including (yellow) and excluding (red) the influence of acoustic differences between emotional and neutral stimuli. Conventions as in Figure 3B.