Decoding emotional prosody: resolving differences in functional neuroanatomy from fMRI and lesion studies using TMS

L. Alba-Ferrara, A. Ellison, R. L. C. Mitchell

Department of Psychology, Durham University, United Kingdom

Background
Prosody conveys information about the emotional state and intention of others. Lesion studies have shown that damage to the right posterior temporal region is associated with prosody decoding deficits. Dissimilarly to findings from lesion studies, neuroimaging data show substantial bilateral peri-Sylvian activation.

Objective
This study aimed to investigate the involvement of the left and right superior temporal gyrus (STG) in prosodic and semantic processing using transcranial magnetic stimulation (TMS). These two regions of interest were chosen for their correspondence to Wernicke’s area in the left hemisphere and its analog in the right.

Methods
Offline TMS with a stimulation frequency of 1 Hz and intensity of 60% of stimulator output (approximately 1.1 Tesla) with one pulse applied per second for 10 minutes (600 pulses) was performed. Directly after TMS on the right STG, the left STG or sham-stimulation, participants completed a prosody decoding or a semantic judgment task (whether the tone/meaning was happy or sad).

Results
Reaction times (RT) for the prosodic task were significantly slower when TMS was applied in the right STG in comparison to left STG and sham conditions. TMS over both right and left STG delayed RT in the semantic task, significantly when the tone of voice was incongruent with the meaning.

Conclusions
Our data strongly suggests that left temporal regions are not crucial to the basic task of prosody decoding per se; however, the analogous region on the right is. Hence, involvement of the left STG in prosodic decoding revealed in previous imaging data is incidental.

© 2011 Elsevier Inc. All rights reserved.

This study was supported by the Department of Psychology at Durham University. The authors declare they have no competing financial interests. Correspondence: L. Alba-Ferrara, Department of Psychology, Durham University, Science Site, South I Road, Durham, Co Durham, DH1 3LE, UK. E-mail address: l.m.alba-ferrara@durham.ac.uk

Submitted April 7, 2011. Accepted for publication June 15, 2011.

1935-861X/S - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.brs.2011.06.004
Emotional prosody is a crucial higher-order language function that encompasses nonverbal aspects of speech necessary for recognizing and conveying emotions in speech. In addition to prosody, speech also conveys linguistic or semantic content. The neural underpinnings of semantic aspects of language have been extensively assessed. It is accepted that lexicosemantic decoding predominantly relies on the processing that occurs in the left superior temporal gyrus (STG). Dissociations between emotional prosody decoding and semantic processing were found in patients with left hemisphere lesions. Taking into account the impairment in the expression of emotional prosody in the comprehension of emotional prosody that patients with right hemisphere brain damage suffer from, Ross proposed that emotional prosody relies on the integrity of the right hemisphere. Further neuropsychologic investigations have determined that patients with specifically right posterior lateral temporal lobe lesions have deficits in the comprehension (rather than the production or repetition) of emotional prosody with patients with analogous lesions on the left producing near normal performance.

Even though lesion studies have indicated the importance of the right posterior hemisphere for decoding emotional prosody, the neural underpinning of this function is not without controversy. In fact, most functional imaging experiments of speech prosody, show substantial bilateral peri-Sylvian activations. For instance, bilateral activation in the posterior STG in affective prosodic aspects of language has led to the interpretation of the left hemisphere as a contributor toward phonetic segmental processing of the vocal stimuli. Alternatively, it has been claimed that bilateral peri-Sylvian activation indicates an increase in task demand. In other words, when the task demand exceeds the processing capacity of the right peri-Sylvian area, its homotopic region needs to be recruited. Controversy exists as to the exact causal substrate of prosody decoding, particularly in relation to the lateralization of the superior temporal cortex in this task. According to some studies, bilateral activation during prosody decoding is circumscribed to the middle superior temporal cortex adjacent to the Sylvian fissure, as well as bilateral anterior and posterior regions. However, with respect to more posterior aspects of the superior temporal cortex, some imaging studies in this field report highly lateralized effects for posterior STG and identify a contribution to prosody decoding particularly for the right posterior superior temporal cortex. Thus, we consider that the fMRI literature on prosody decoding cannot entirely determine the causal involvement of the posterior right STG disentangling it from posterior left STG. The controversy between neuroimaging and lesion studies in this domain is thus still ongoing.

To examine this apparent disagreement between neuropsychologic and neuroimaging findings, we propose to use a neurodisruptive technique (TMS) to determine the absolute involvement of the posterior right and left STG to investigate dissociations in involvement in semantic and prosodic judgements. Functional imaging is correlative in nature and when comparing imaging and neuropsychologic data, imaging data indicate brain regions that may plausibly be involved (either directly or coincidently) in a task, but not those that are "necessary." Also, inferences about normal neural function based on lesion studies are not robust to problems either as lesions often lead to compensatory reorganisation, and their foci may spread across more than one region.

In some previous neuroimaging studies, prosodic processing has not been totally disentangled from semantic processing because of speech conveys both kinds of information. Therefore, the nature of the speech stimuli, in addition to the correlational nature of fMRI findings, make it difficult to interpret whether the obtained brain responses reflect linguistic or nonlinguistic processes. Our study uses the same stimuli for both conditions, but different instructions (i.e., judging the semantic content ignoring the tone of voice and vice versa) creating a prosodic and a semantic judgment task, allowing us to test for a double dissociation between the neural correlates of semantic and prosodic processing. Moreover, our study includes incongruent sentences that are necessary to identify the source from which participants derived the emotion cues. In other words, if semantics and prosody were consistently congruent we would not be able to disentangle whether the participant's responses are due to the meaning of the sentence or the tone of voice.

Our study uses the same experimental stimuli applied by Mitchell and colleagues. By doing so, we aimed to make our results directly comparable with this prior fMRI research. One problem in the literature consists in the diversity of tasks and stimuli applied to measure prosody decoding. For example, some studies used in single words whether others used whole sentences. It is believed that auditory stimuli length modulates activity in the primary and secondary auditory cortex, thus the difference in stimuli length might have contribute toward contradictory findings. We used TMS to target the posterior part of the STG in both hemispheres. The right STG is the contralateral equivalent to Wernicke’s area (which is located in the left STG and is known for its role in semantic processing) and as previously mentioned has been implicated in prosodic processing in both neuropsychologic and neuroimaging studies. Moreover, it is currently believed that prosodic processing in the right hemisphere mirrors semantic processing in the left. For that reason we decided to target the posterior STG. The selection of this region of
interest (ROI) for TMS as well as the design of a paradigm
to evaluate the interhemispheric neuroanatomy of emotional
prosody decoding by targeting homotopic regions in both
hemispheres are the novel aspects of our study.

TMS methods

Materials and methods

Participants
Eleven healthy native English speakers (age Mean 34
Standard Deviation [SD] 10, 4 females) participated in
the study. All participants reported to be right handed and
having normal hearing. Participants gave their signed
informed consent in accordance with the Declaration of
Helsinki and with the approval of Durham University
Ethics Advisory Committee, and could leave the experi-
ment at any point. Subject selection complied with current
guidelines for rTMS research.\textsuperscript{28,29}

TMS
TMS was applied over two regions of interest, left and right
posterior STG, chosen for their correspondence to
Wernicke’s area in the left hemisphere and its homologue
in the right (Figure 1). Positions were located before and
maintained during each experimental session, using frame-
less stereotaxy (Brainsight, Rogue Research, Montreal,
Canada) on each subject’s anatomic magnetic resonance
imaging (MRI) scan previously collected at the Newcastle
Magnetic Resonance Centre.

Offline rTMS with a stimulation frequency of 1 Hz was
performed using a 70-mm diameter figure-of-eight coil
connected to a Magstim Super Rapid magnetic stimulator
(Magstim, Whitland, Dyffed, Wales, UK). The coil was
held tangentially on the skull over the ROI in a constant
position with the handle pointing medially parallel to the
horizontal and midsagittal plane by the experimenter at all
times. The stimulation intensity was set at 60% of
stimulator output (approx 1.1 Tesla) with one pulse applied
per second for 10 minutes (600 pulses). According to the
literature, this procedure should affect the neural activity of
the ROI for approximately 6 minutes\textsuperscript{30}

TMS procedure
Participants were seated in a comfortable chair in front of
a computer screen and fitted with a swimming cap to allow
marking of the stimulation sites. A chin rest was used to

Figure 1 Stimulated areas were localized using each subject’s MPrages coregistered to their skull coordinates using Brainsight software.
minimize head movements during the experimental blocks. Directly after rTMS on the right STG, the left STG, or sham-stimulation (in which a nondischarging coil was held to one or other ROI whereas a discharging coil was in close proximity resulting in similar conditions to TMS application but without the magnetic pulse), participants completed one block of the task for approximately 4 minutes. Subsequently, there was a 30-minute break before the next block so that neural activity returned to baseline because it has been demonstrated that the effects of rTMS in neural excitability that outlast the period of stimulation may last for several minutes. The order of the experimental tasks and the stimulation blocks were counterbalanced over subjects and sessions. The entire study was completed in two sessions separated by 7 days interval. In each session, TMS on the right STG, left STG, and sham-stimulation was applied and the three blocks that comprise the prosodic or the semantic tasks were performed. Sessions lasted approximately 1.5 hours per subject.

**Emotional prosody tasks**

Sentences of happy or sad semantic content (i.e., “she was delighted to be pregnant,” “the dog had to be put down”), pronounced in a happy or sad tone of voice by a male native phonetician of British English were used as stimuli. Half of the sentences had congruency between prosodic and semantic valences (i.e., sentences with happy meaning spoken in happy tone of voice) whether the other half were incongruent. The sentences were approximately the same length (+ 100 milliseconds) to avoid variability in decision time, and were of a consistent style and format (duration Mean = 2.1511, SD = 0.2558). Three stimuli lists were produced. Each of the stimuli list contained equal numbers of happy and sad content sentences spoken in happy and sad tones of voice. Stimuli list and stimulation side (left, right, and sham) were counterbalanced. The stimuli were the same as those used previously in the fMRI study of Mitchell et al. The tasks were developed and presented using E-prime software (Psychology Software Tools; Sharpsburg, PA). Each trial lasted 4 seconds including the sentence and the intertrial interval. Each experimental block consisted of 60 trials. The stimuli were presented through two loudspeakers, located on each side of the PC screen. Participants had to respond via a key press on a PST serial response box. The index finger (left button) was used to respond for “sad,” and the middle finger (right button) was used to indicate “happy.” For both tasks the same combined semantic-prosodic stimuli were used. In the semantic tasks, participants were asked to focus in the meaning of the sentence, ignoring the tone of voice and answer whether the content of the sentence was happy or sad. In the prosodic task, participants were asked to focus on tone of voice and ignore the meaning, and indicate whether the intonation was happy or sad. Participants were instructed to respond as fast as they could, but without sacrificing accuracy.

**Statistical analysis**

Our main analysis comprised a 2 × 2 (task [semantic, prosodic] × TMS site [right STG, left STG]) repeated measures analysis of variance (ANOVA) using normalized reaction times (normalized RT = (RT [TMS] − RT [sham]) / RT [sham]) for comparison across task and stimulation site to investigate relative differences in TMS effect across conditions. This is a standard analysis that takes into account each participant’s performance with respect to their relative control and so the effect of TMS across site and task can be easily compared. This analysis was also completed to investigate trends in error rates.

Performance within each task was further analyzed by four one-factor repeated measures (TMS [sham RT vs left TMS RT vs right TMS RT]) ANOVAs in which the effect of congruency was also taken into account (i.e., congruent prosody, incongruent prosody, congruent semantic, incongruent semantic).

**Results**

**Error data analysis**

During both tasks and all stimulation sides (including sham) the accuracy was almost perfect (Prosody task [mean accuracy, SE]: left STG TMS: 93 %, 0.02%; right STG TMS 92 %, 0.01%; Sham TMS: 92%. 0.02%.) Semantic task: left STG TMS 98%, 0.01%; right STG TMS 97%; 0.01%; Sham TMS 98%; 0.01%. Normalized accuracy data was analyzed through a 2 × 2 ANOVA. There was no main effect of task (F(1,10) = 0.083, P = 0.779, η² = 0.008) or TMS (F(1,10) = 1.558, P = 0.240, η² = 0.135). There was no interaction between TMS side and task (F(1,10) = 0.094, P = 0.765, η² = 0.009). Therefore, no further analysis on error rates was carried out. All errors were removed from the dataset for analysis of reaction times.

**Reaction time analysis**

This analysis reveal that there was no main effect for task (F(1,10) = 0.366, P = 0.559, η² = 0.04). However, there was a main effect for TMS indicating that TMS had a greater effect (F(1,10) = 11.81, P = 0.006, η² = 0.54) on reaction times when applied over the right STG in comparison to the left. There was a significant interaction between TMS side and task (F(1,10) = 5.35, P = 0.043, η² = 0.35) (Figure 1). A post hoc Bonferroni test comparing the effect of TMS on right and left STG for the prosodic task showed that TMS had a smaller effect over the left STG in comparison to the right (t(10) = −4.18, P < 0.001). However, there was no significant difference between TMS effect over the right or left STG in the semantic task (t(10) = −124, P < 0.904) (Figure 2).

The TMS effect on right STG in the prosodic task was found to be significant using a one-tailed comparison to
significant effect of TMS in congruent trials ($F_{(1,10)} = 5.91, p = 0.035$ and $t_{(10)} = 2.068, P = 0.033, \text{respectively}$).

Our further within task analysis revealed there was no significant effect of TMS in congruent trials ($F_{(2,20)} = 2.18, p = 0.139 \eta^2 = .17$) in the semantic task. However, TMS did significantly affect reaction times for incongruent trials ($F_{(2,22)} = 5.91, p = 0.009 \eta^2 = .34$) in this task. As can be seen in Figure 3, post hoc Fisher’s least significant difference (LSD) pair-wise comparisons revealed that TMS significantly increased reaction times over right STG ($P = 0.015$) and also over left TMS ($p = 0.018$).

In the prosodic task, there was no significant effect of TMS in congruent trials ($F_{(2,20)} = 2.196, p < 0.137 \eta^2 = 0.19$). TMS did affect incongruent trials however with a main effect of TMS ($F_{(2,20)} = 4.095, p < 0.032 \eta^2 = 0.22$) with a post hoc significant difference (LSD) only between reaction times when TMS is applied over the right STG and sham TMS.

### Discussion

The purpose of the current study was to clarify the interhemispheric neural correlates of emotional prosody decoding. We have found a critical involvement of right STG for the emotional prosody task, unlike its contralateral side. We also found both right and left STG were involved in the semantic task. On further investigation, however, it would seem that there is bilateral involvement of STG for trials in which the tone and meaning are incongruent when participants are asked to make a purely semantic judgement. Such investigation shed light on how incongruency may increase semantic demand resulting in bilateral involvement.

Our main finding was a prominent distinction in the involvement of the right and left STG in emotional prosody decoding. The right STG seems to have a causal contribution in the prosodic task, as TMS over this region showed a disruptive effect, unlike TMS over its contralateral homologue. Partial eta squared showed that 35% of the variance that was found in the analysis was associated with the effect that TMS had on the task, and this large effect size strongly suggests a dissociation between the right and left STG in prosody decoding. Our finding sheds light onto the interhemispheric localization of emotional prosody decoding, and is in line with lesion studies demonstrating a causal involvement of the right STG in emotional prosody decoding.\(^7,35\)

According to the fMRI literature, not only the right but also the left STG appears to be associated with emotional prosody decoding tasks.\(^12,16,25,36\) The current study used the same stimuli as were used in a previous fMRI study that uncovered bilateral activations (Mitchell et al.\(^24\)); however, TMS has now shown that left STG does not have a causal role. The fMRI indicated left STG involvement in emotional prosody decoding has been interpreted as related to explicit labelling of emotional valences during prosody tasks or to automatic linguistic processing depending on the semantic load of the stimuli.\(^12\) Lesion studies have shown that the more complex the linguistic information embedded in the emotional prosody stimuli, the more frequent emotional prosody decoding deficits were amongst patients with left hemisphere lesions.\(^12,37-39\) In contrast, patients with right hemisphere lesions experienced difficulty independently of the stimuli linguistic load. One interpretation for this result lies in the confounds associated with lesions studies. Functional and structural changes in homotopic regions in the cortex contralateral to a lesion have been reported in the literature, mainly linked to neural connections between the areas.\(^40\) Thus effects due to brain reorganization of a cognitive function cannot be dismissed. We consider that lesion studies claiming an involvement of the left STG in emotional prosody decoding should be interpreted taking into account brain plasticity phenomena specifically differentiation of regions homotopic to the lesion.

Due to its role in facial processing, Van Rijn et al.\(^41\) used TMS to investigate the role of the right frontal operculum in prosodic processing. TMS has previously been used to investigate the role of the right frontal operculum
in prosodic and the time course of involvement of this region and the right STG was further investigated by Hoekert et al., who found that stimulation of both ROIs resulted in longer RT in comparison to a control condition. Moreover, a recent study found increased RT in the emotional prosody decoding task during TMS over both left and right inferior frontal gyrus as compared with sham condition. They interpreted this finding as a demonstration of the critical involvement of both right and left inferior frontal gyri in emotional prosody decoding. The current study builds on these findings by investigating language processing in the temporal lobe and also including a control measure to rule out nonspecific effects of TMS. In contrast to activity in the frontal cortex, prosodic processing would seem to be lateralized to the right STG.

Our second finding shows that stimulation of either hemisphere (left STG and right STG) delayed processing of the incongruent trials of the semantic task in comparison to the baseline, and this difference was not found for congruent trials alone. In addition, there was no difference between RTs between right and left STG. Thus, our finding should be interpreted taking into account that our stimuli contained high emotional load (emotional meaning) from which participants judged the emotional content. We interpret the bilateral involvement of the STG as related to the semantic emotional salience of the stimuli, in which the left STG would be in charge of lexicosemantic processing in general and the right STG would contribute to the processing of the emotional load conveyed by the meaning of the sentences. In line with our results, a considerable number of lesion studies reported that patients with right brain damage at various loci present difficulties in the perception of lexically based emotional stimuli. It may be possible that higher task demands can increase the recruitment of brain regions therefore the judgement of incongruent trials may engage a different a neural network than the congruent trials. In other words, greater task demands drive the involvement of STG in a semantic task therefore resulting in a greater effect of neurodisruption in incongruent trials. In our study, the stimuli used concurrently conveyed semantic and prosodic information. By manipulating the instructions and asking participants to judge prosody while ignoring semantics (and vice versa), we intended to bias their attention toward the processing of prosodic features. However, we acknowledge the possibility of implicit background effects of the prosodic information. Further research including sentences with emotionally neutral meaning is urgently needed. We propose future studies investigating the neural underpinnings of semantic decoding of utterances with and without emotional content.

Conclusion

Our study strongly suggested that left temporal regions are not crucial to the basic task of prosody decoding in the absence of semantic processing, however, the analogous region on the right is. Hence, previous imaging data indicate incidental involvement of the left STG in prosodic decoding but our TMS data show it is not necessary for pure prosody. Furthermore, it has also shown that left as well as right STG is involved in the semantic judgement of sentences with emotional meaning. This may be because right STG was involved on emotional grounds (as all sentences had emotional meaning) and left STG on semantic grounds. We anticipate further research addressing the neural correlates of semantic judgement with and without emotional meaning.

Acknowledgment

We would like to thank Prof. Elliott D. Ross (Department of Neurology, University of Oklahoma College of Medicine) for the motivating discussions for the inception of this study, and his valuable comments on its design.

References

12. Mitchell RLC, Ross ED. fMRI evidence for the effect of verbal complexity on lateralisation of the neural response associated with


