Lateralization in emotional speech perception following transcranial direct current stimulation

Alex Francois-Nienaber[1], *Jed A. Meltzer*[2], *Frank Rudzicz*[1,3]

[1]Department of Computer Science, University of Toronto, Toronto, Canada
[2]Rotman Research Institute, Toronto, Canada
[3]Toronto Rehabilitation Institute-UHN, Toronto, Canada

lex@cs.toronto.edu, jmeltzer@research.baycrest.org, frank@cs.toronto.edu

Abstract

The degree to which the perception of spoken emotion is lateralized in the brain remains a controversial topic. This work examines hemispheric differences in the perception of emotion in speech by applying tDCS, a neurostimulation protocol, to the T-RES speech emotion rating paradigm. We find several significant effects, including a strong interaction of prosody and neurostimulation for perceptual ratings when considering only lexical content, and that the perception of happiness does not appear to be affected by tDCS, but anger and (to a large extent) fear appear less intense after stimulation.

Index Terms: emotional perception, prosody, tDCS

1. Introduction

Human communication does not rely solely on *what* is said – the lexical content of speech – but also on *how* it is said – its prosody. Recent studies of hemispheric specialization in language have mitigated theories of "dominance" of the left hemisphere and shown involvement of both hemispheres in certain language tasks. Specifically, while the left hemisphere has consistently demonstrated a superior role in processing linguistic content, the right hemisphere has been established as a centre for emotional processing, especially of paralinguistic emotional content [1]. Lesion studies have shown that a particular aspect of prosody, specifically the encoding of received auditory emotions, strongly relies on the right hemisphere. In contrast, the predominant role of the right hemisphere in emotional prosody perception has been challenged by neuroimaging studies [2][3][4], and more recent efforts with transcranial magnetic stimulation (TMS) have shown bilateral involvement of brain sites peripheral and anterior to both auditory cortices [5][6]. In either case, how our brains *combine* prosody and lexical information to determine the perceived emotion of a spoken sentence remains an open question.

Emotional prosody (EP) in speech perception involves bilateral superior temporal gyri (STG) and inferior frontal gyri (IFG) [4][7][8]. Although there appears to be a stronger involvement of right STG [9][10] in EP tasks, the precise role of those sites in each hemisphere is still disputed. For instance, a neuroimaging study showing that right IFG had a predominant role in EP perception [10] was later mitigated by results of a TMS study showing little lateralization of IFG in EP tasks [6].

The goal of this study was to alter the processing of emotional prosody in received speech through transcranial direct current stimulation (tDCS). TDCS is a neurostimulation protocol in which a weak electrical current is sent across the scalp. Some of this current flows through the cerebral cortex and leads to a change in transmembrane neuronal potential. Anodal tDCS tends to increase cortical excitability, lowering the threshold needed to initiate an action potential, while cathodal stimulation has an opposite, inhibitory effect [11][12]. Unlike TMS, it does not make neurons fire directly, rather it modulates their spontaneous firing. All such changes are temporary and may last up to 1 hour after stimulation has ended.

2. Background

It is generally accepted that the processing of EP by the right auditory cortex is a multi-stage process [1][4][13][14], recruiting sites on the left hemisphere for parts of the processing, which explains some *seemingly* contradictory findings involving lateralization. For example, Hoekert *et al.* first showed a definite involvement of the right hemisphere for the identification of the emotion fear (and sadness) in prosody [5], and later showed a stronger left hemispheric bias for that particular emotion [6]. However, inhibitory stimulation can actually have the reverse effect by preventing competition with other areas, as was shown with TMS [15] and tDCS [16]. In the TMS study, the authors note that their subjects' reaction times decreased on an emotional semantic task after TMS over the right fronto-parietal operculum (FPO), and postulate that this is caused by facilitated processing by the ventrolateral prefrontal cortex. In the tDCS study, perhaps the first one involving EP, the authors used a dichotic listening paradigm and showed that cathodal stimulation over the right IFG facilitated the recognition of EP. Their main argument was that cathodal stimulation helped select the right prosodic category by inhibiting signals from the distractors.

Given the multiple adjacent regions involved, tDCS might lend itself well to investigate cortical EP processing. Indeed, increasingly many studies show the involvement of a range of sites along the 'what' and 'where' pathways of audition rather than a few localised areas. Thus, even though earlier studies showed a stronger involvement of the right anterior STG than the right posterior STG (and left anterior STG) [10], recent studies have shed light on an auditory complex for EP computation [13], involving right FPO, bilateral IFG, and bilateral STG, which are activated at different stages of emotional processing [1][4] or even in parallel (e.g., right anterior STG and FPO [5]). Efforts to understand the lateralization of EP in the brain have developed to target these areas individually, however we believe that valuable insights may be gained by treating each auditory complex as a unit. First, our results would be more general, and second, the affordability and availability - not to mention safety and compatibility with pharmacotherapy plans - of tDCS devices make it a powerful medium for the deployment of new neuro/cognitive rehabilitation protocols.

A 2011 meta-analysis [4] discussed four different hypotheses of lateralization, and proceeded by elimination to settle on two non-mutually exclusive theories: the functional lateralization hypothesis (also known as the acoustic lateralization hypothesis [8]), which postulates that the left hemisphere processes the linguistic prosody category and the right hemisphere processes the EP category, and the cue-dependent lateralization hypothesis, which postulates that which hemisphere is doing most of the EP processing depends on the acoustic properties of the prosody. The analysis was not able to directly compare the functional versus cue-dependent hypotheses.

3. Methods

3.1. Experimental design

If the functional hypothesis is true, then we would expect excitation of the left hemisphere and inhibition of the right to facilitate recognition of the emotion in the lexical content despite the effect of prosody. We thus decided to contrast 2 opposite stimulation conditions. In the **A-C** condition, we stimulate the left hemisphere with anodal tDCS (excitatory), and place the cathode on the right hemisphere to inhibit it [11][12]. In the **C-A** condition, the electrode montage is reversed (i.e., Cathode on the left and Anode on the right). Since cortical effects of tDCS last from 30 minutes to 1 hour after stimulation, we timed all the tasks in the study and separated both stimulation conditions by at least a 24-hour interval. As a result, each participant underwent two sessions, each involving 20 minutes of tDCS. We also included a control condition before any stimulation and, to reduce practice effects, half of the control tasks were performed in the first session, and the other half in the second.

Data collection took place in a sound-attenuated room at the University of Toronto. Participants were seated in front of a monitor and asked to perform some practice tests at the start of each session. Once comfortable with the software used for testing, they took a short break before performing two consecutive emotional processing tasks. In each, they had to gauge the emotional intensity of emotionally charged sentences presented via headphones. The control part of the session was followed by 20 minutes of stimulation with tDCS, after which participants performed two more rating tasks.

Fourteen students (12 male, 2 female) from the University of Toronto volunteered for this study. The participants were all right handed to control for lateralization effects, and aged between 18 and 30 years (mean 23.5, s.d. 3.46). Data collection was conducted in accordance with established safety guidelines for tDCS [11][17][18].

3.2. tDCS procedure

The tDCS was administered using a medical-grade stimulator from TCT technologies. We used intensity and duration settings typical for the tDCS protocol, i.e., 20 minutes of stimulation under a current intensity of 2 mA [12]. We used two rectangular anode and cathode sponge electrodes, each 5 cm $\times$ 7 cm (i.e., a surface area of 35 cm^2), for a current density of .057 mA/cm^2 under each electrode. Electrodes were soaked in a saline solution at a concentration of 0.9%, and held in place using the neoprene strap included with the device. To minimize discomfort at the stimulation sites, we used the first 30 seconds of stimulation to ramp up the current [19]. We monitored for potential side-effects of stimulation by asking participants to complete a tDCS adverse effects questionnaire at the end of their second session. Apart from typical tDCS side-effects (e.g., skin irritation, tingling under the electrodes, mild headache), no serious issues were reported. One participant dropped out of the study because the tingling caused discomfort.

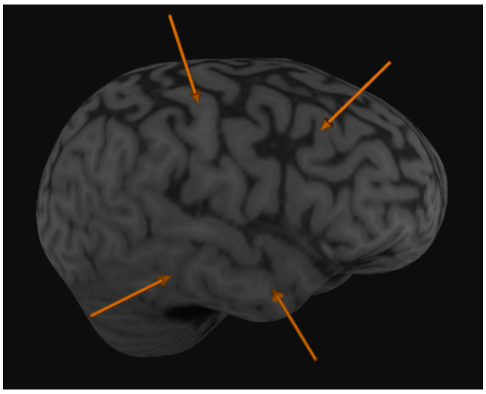

Figure 1: The region of interest: the anterior superior temporal gyrus and fronto-parietal operculum (shown on the right hemisphere). The orange arrows indicate the corners of the sponge.

The electrodes were placed on the scalp using the 10-20 electroencephalographic electrode positioning system. To determine landmarks as precisely as possible, we used neuronavigation software and equipment originally intended for TMS, and located the desired region of stimulation on the scalp of a right-handed individual not participating in the study. This helped us establish that good coverage of our region of interest would be ensured by placing the top left corner of the sponge electrode on C4 (respectively the top right corner of the contralateral sponge on C3). How the location of the electrode on the scalp maps to the right hemisphere is presented in figure 1.

To induce a more consistent state of relaxation prior to the experiment, we preceded the control condition with 3 to 5 minutes of 'play time' consisting of 4 simple puzzle games the participant could play with, none involving language.

3.3. Speech stimuli

All stimuli originate from the Test of Rating of Emotions in Speech (T-RES) [20]. We used 76 distinct audio prompts consisting of linguistically equated sentences that were recorded by a native English actress. Both the lexical and prosodic content of each sentence independently expresses one of 5 cardinal emotions (happiness, anger, sadness, fear, neutral). Hence, in certain cases, the emotions conveyed by the lexical content and the prosody match (e.g., *'I am so happy'* spoken in a happy prosody), but not in others (e.g., *'I am so happy'* spoken in an angry prosody). Audio files were on average 4 seconds long.

The stimuli were organized into 3 blocks of 24 prompts, 2 for the A-C and C-A stimulation conditions and 1 for the control. In each, every combination of emotions in the two dimensions considered (lexical and prosodic) were represented, except that in which a lexically neutral sentence is spoken in a neutral prosody. This allowed the tasks to be slightly shorter without compromising our objectives.

We should note that due to stimuli availability, we had to reuse 4 prompts. Thus, the control block shares 2 prompts with each of the stimulation blocks. To mitigate experimental bias, we designed the overlaps so that each contain an equal number of *withdrawal* (sadness, fear) and *approach* emotions (happiness, anger) in the prosody (the lexical content was neutral for all 4 overlapping prompts).

Given that the control condition is split across both sessions, we divided the 24-prompt control block in half. In order to maintain good coverage across both dimensions, we selected two subsets of 12 prompts from the control set in a latin square fashion. The 8 remaining prompts were used for the practice tests performed to familiarize participants with the software.

3.4. Language tasks

The two linguistic tasks used in the experiment are inspired by the original T-RES [20]. In each, the participant listens to a speech utterance via headphones, then is asked to rate it on four 6-point Likert scales, indicating the degree to which the speaker was either happy, angry, sad, or fearful (6 being the greatest degree). As the effects of tDCS wear off gradually, we fixed a maximum duration of 15 seconds to rate each prompt. We had initially determined 30 seconds would be a sufficient amount of time for the ratings to be made however, after a pilot session, we noticed that once comfortable with the interface, the participant was able to provide all the ratings at an average of 5 seconds per prompt. We also randomized the order in which the stimuli were presented in each block.

Concretely, what this entails is that the participant first clicks a 'play' button, which loads an audio file at random from the appropriate block. Once it has finished playing, the participant clicks a 'rate' button, upon which four Likert scales and a 15-second non-numerical timer appear on the screen. When the counter reaches 0, the scales disappear from the screen and this is repeated until all the stimuli in a task have been rated. The four emotional ratings, as well as the total time it took to complete all ratings, are recorded.

To contrast the effects of bilateral tDCS over the auditory cortices on emotional prosody perception, it is sufficient for us to replicate two of the three tasks in T-RES. In the **General task**, the participants are asked to rate the sentence as a whole, using information from both the lexical content and the prosody; while in the **Lexical-only task**, participants are asked to focus exclusively on the lexical content, ignoring the prosody. Ben-David *et al.* noted that typical subjects have trouble with selectively attending to only a single dimension, and that this difficulty was more pronounced when attempting to ignore the prosody [20]. By having participants attend to the lexical dimension rather than the prosody, we increase the likelihood of detecting a clear effect of tDCS given the inherently stronger bias towards attending to prosodic information.

4. Results

We analyze both the distribution of subjective ratings and the response times under the various conditions.

4.1. Subjective ratings

Our first goal is to determine the extent to which tDCS affects the ability to ignore prosody. Table 1 shows, for each of the three experimental conditions (i.e., **CTRL**, **A-C**, and **C-A**) and tasks (i.e., **General** and **Lexical-only** ratings), the average subjective rating given to the degree of the emotion encoded in the *lexical* dimension, regardless of prosody, for each combination of lexical and prosodic emotions. This allows a clear visual distinction between tasks, and tests the extent to which listeners are able to attend to the lexical dimension – if listeners ignore prosody, then ratings should be uniform across prosody for each lexical dimension.

		General				Lexical-only			
	lexical \ prosody	A	F	H	S	A	F	H	S
CTRL	A	5.21	4.07	4.14	2.14	5.00	5.29	4.93	4.21
	F	2.71	5.50	3.21	4.36	4.86	5.57	5.21	4.93
	H	1.79	2.93	5.43	2.14	5.00	5.57	5.50	4.57
	S	2.79	4.14	1.93	5.86	5.50	3.29	3.29	5.29
A-C	A	5.79	3.93	2.71	2.43	5.50	5.36	4.79	4.64
	F	2.86	4.64	2.86	5.07	4.00	4.71	5.14	5.43
	H	1.71	3.71	5.50	3.36	4.79	5.29	5.21	5.14
	S	2.64	3.29	3.07	5.64	4.93	4.29	5.64	5.36
C-A	A	5.79	3.79	2.86	2.14	5.29	5.36	5.07	4.21
	F	3.14	5.07	2.50	5.14	4.86	4.86	5.00	5.43
	H	1.57	3.43	5.36	3.00	4.71	5.07	5.14	5.50
	S	2.64	3.07	3.07	5.79	4.07	4.50	5.36	5.29

Table 1: Average subjective ratings of the emotion encoded in the indicated lexical dimension, across all lexical/prosody permutations, tasks, and neurostimulation conditions. For example, in the General task in the CTRL condition, participants on average rated the *fearfulness* of lexically fearful and prosodically angry utterances 2.71, and on average they rated the *happiness* of lexically happy and prosodically angry utterances 1.79. Black = 6.0, white = 0.0.

The ratings in the CTRL condition are similar to those in the original T-RES [20]. To analyze the effects of stimulation on the ratings, we performed a repeated-measures ANOVA to test the influence of the prosody and partition out inter-subject variability. We found a strong interaction of the prosody and stimulation condition for the ratings on the Lexical-only task ($F_{6,78} = 2.77, p = 0.017$) but not on the General task ($F_{6,78} = 1.99, p = 0.078$), with a significant 3-way interaction showing stronger tDCS by prosody correlations in the Lexical-only task ($F_{6,78} = 2.69, p = 0.020$). This suggests that, after tDCS, it is harder to ignore the prosody. Although we did not notice significant differences between the two stimulation conditions when taking into account all the emotions, we speculate that this could be caused by emotions being processed differently across both hemispheres. If the functional lateralization hypothesis is true, then we would expect to observe more influence of the prosody in C-A relative to A-C, which is not the case.

The CTRL condition in the General task is clearly diagonally dominant but, in the A-C and C-A conditions of the General task, sentences that are lexically fearful but prosodically sad are considered to be *more fearful* than sentences with both fearful prosody and lexical content. This implies that, in the General task after stimulation, individuals attend more to the lexical content (when fearful) than to prosody; this is significant across stimulation conditions (right-tailed homoscedastic $t(110) = 1.73, p < 0.05, CI = [0.02, \infty]$). Secondly, when asked to measure the sadness of lexically sad sentences, in the CTRL condition and Lexical-only task, subjects gave much lower scores when the prosody was either fearful (3.29) or happy (3.29) than the other two prosodic scenarios, on average, which is significant (two-tailed homoscedastic $t(110) = 5.51, p < 0.001, CI = [1.19, \infty]$); this case vanishes during both types of stimulation.

We also observe that with an angry prosody, A-C tended to lower the ratings of fear in the Lexical-only task (though not significantly, $F_{1,13} = 2.586, p = 0.132$) compared to C-A. Along similar lines, C-A tended to decrease sadness ratings when compared to A-C (again, not significantly, $F_{1,13} = 4.216, p = 0.061$).

If we conflate tasks, we can focus on our second goal, which is to determine how tDCS effects the perception of different emotions in speech generally. Figure 2 shows average scores given to each emotion, across all speech stimuli, by brain stimulation condition. Interestingly, the perception of happiness does not appear to be in any way affected by tDCS (homoscedastic two-tailed $t(2014) = 0.019, p = 0.99$), while the ratings of other emotions do drop during any kind of brain stimulation. Neither changes to anger (homoscedastic two-tailed $t(2014) = 1.825, p = 0.07$) nor fear (homoscedastic two-tailed $t(2014) = 1.620, p = 0.11$) is significant.

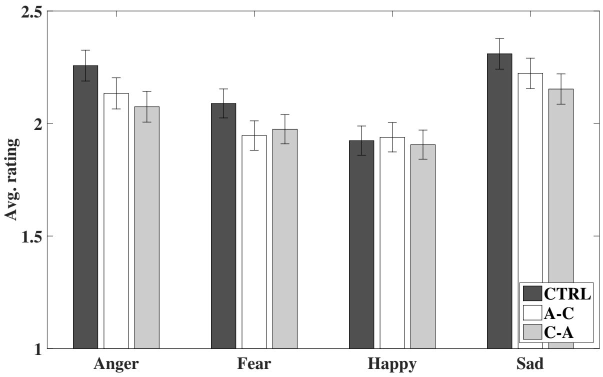

Figure 2: Average scores (and std.err.) given to each emotion by stimulation condition.

4.2. Response times

Table 2 shows the average (and std.dev.) of the response times per stimulus for each task and condition. A full n-way ANOVA across all stimuli shows that there are significant main effects of stimulation condition ($F_{2,2} = 13.81, p < 0.001$) and rating task ($F_{1,2} = 78.82, p < 0.001$), but no interaction between condition and task ($F_{2,2} = 0.75, p = 0.47$). That is, on average subjects are both uniformly slower in rating during the General task than the Lexical-only task, and uniformly slower in the CTRL condition than in either stimulation condition. We hypothesize that the slowness of the General task is due to the need to integrate multiple modalities of input. That participants are faster in the tDCS condition than in CTRL is interesting since we see no other increases in response times over time. This is demonstrated across emotions in figure 3, which shows the average time (and std. dev.) taken to rate each speech stimulus along the indicated emotional dimension, by stimulation condition.

	Condition		
	CTRL	A-C	C-A
General	8.69 (2.95)	8.13 (2.87)	8.07 (2.96)
Lexical-only	7.71 (2.91)	6.78 (2.50)	7.00 (2.92)

Table 2: Average (and std.dev) response times, in seconds, per stimulus for each task and condition.

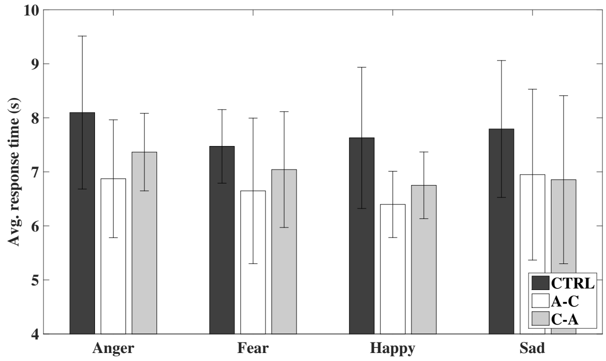

Figure 3: Average times (and std.dev.) taken to rate each speech stimulus for each emotion, by stimulation condition.

5. Discussion

We present the first application of transcranial direct current stimulation (tDCS) to an emotional rating paradigm using speech stimuli of this type (i.e., T-RES). Our primary hypothesis - that A-C would bias responses in favour of lexical content, and C-A in favour of prosody - was not confirmed. Rather, we found more nuanced effects of stimulation that varied across emotions. Specifically, 1) individuals attend more to fearful lexical content than to prosody during tDCS when also considering prosody, 2) there is a strong interaction of prosody and stimulation condition for perceptual ratings when considering Lexical-only content, 3) the perception of happiness does not appear to be affected by tDCS, but anger and (to a large extent) fear appear less intense after stimulation. We also find that rating perceived emotions is significantly faster after stimulation; no significant differences are observed over time within the control condition, so this does not appear to be related to practice effects.

Future work includes expanding this paradigm to a larger body of participants, including a Prosody-only task, and exploring perceptual patterns across specific emotions. While these results indicate how tDCS can impact the perception of emotions in speech in healthy individuals, we are also interested in expanding this work to pathological populations.

6. Acknowledgements

We thank the T-RES authors for allowing us to use their audio prompts [20].

7. References

[1] A. Schirmer and S. A. Kotz, "Beyond the right hemisphere brain mechanisms mediating vocal emotional processing," *Trends in Cognitive Sciences*, vol. 10, 2006.

[2] S. A. Kotz, M. Meyer, K. Alter, M. Besson, D. Y. von Cramon, and A. D. Friederici, "On the lateralization of emotional prosody: An event-related functional MR investigation," *Brain and Language*, vol. 86, pp. 366–376, 2003.

[3] J. Obleser, F. Eisner, and S. A. Kotz, "Bilateral Speech Comprehension Reflects Differential Sensitivity to Spectral and Temporal Features," *The Journal of Neuroscience*, vol. 28, pp. 8116–8124, 2008.

[4] J. Witteman, M.H. van IJzendoorn, D. van de Velde, V.J.J.P. van Heuven, and N.O. Schiller, "The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature," *Neuropsychologia*, vol. 49, pp. 3722–3738, 2011.

[5] M. Hoekert, L. Bais, R.S. Kahn, and A. Aleman, "Time Course of the Involvement of the Right Anterior Superior Temporal Gyrus and the Right Fronto-Parietal Operculum in Emotional Prosody Perception," *PLoS ONE*, vol. 3, no. 5, e2244, 2008.

[6] M. Hoekert, G. Vingerhoets, and A. Aleman, "Results of a pilot study on the involvement of bilateral inferior frontal gyri in emotional prosody perception: an rTMS study," *BMC Neuroscience*, vol. 11, no. 93, 2010.

[7] G.Rota, R. Sitaram, R. Veit, M. Erb, N. Weiskopf, G. Dogil, and N. Birbaumer, "Self-Regulation of Regional Cortical Activity Using Real-Time fMRI: The Right Inferior Frontal Gyrus and Linguistic Processing," *Human Brain Mapping*, vol. 30, pp. 1605–1614, 2009.

[8] D. Wildgruber, H. Ackermann, B. Kreifelts, and D. Ethofer, "Cerebral processing of linguistic and emotional prosody: fMRI studies," *Progress in Brain Research*, vol. 156, Chap. 13, 2006.

[9] L. Alba-Ferrara, A. Ellison, and R. L. C. Mitchell, "Decoding emotional prosody: Resolving differences in functional neuroanatomy from fMRI and lesion studies using TMS," *Brain Stimulation*, vol. 5, pp. 347–353, 2012.

[10] T. W. Buchanan, K. Lutz, S. Mirzazade, K. Specht, N. J. Shah, K. Zilles, and L. Jäncke, "Recognition of emotional prosody and verbal components of spoken language: an fMRI study," *Cognitive Brain Research*, vol. 9, pp. 227–238, 2000.

[11] A. P. Arul-Anandam, C. Loo, and P. Sachdev, "Transcranial direct current stimulation - what is the evidence for its efficacy and safety?," *F1000 Medicine Report*, vol. 1, 2009.

[12] E. Dayan, N. Censor, E. R. Buch, M. Sandrini, and L. G. Cohen, "Noninvasive brain stimulation: from physiology to network dynamics and back," *Nature Neuroscience*, vol. 16, pp. 838–844, 2013.

[13] J. Witteman, V.J.J.P. van Heuven, and N.O. Schiller, "Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception," *Neuropsychologia*, vol. 50, pp. 2752–2763, 2012.

[14] S. A. Kotz, M. Meyer, and S. Paulmann, "Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design," *Progress in Brain Research*, vol. 156, chap. 15, 2006.

[15] S.van Rijn, A. Aleman, E. van Diessen, C. Berckmoes, G. Vingerhoets, and R.S. Kahn, "What is said or how it is said makes a difference: role of the right Fronto-Parietal Operculum in emotional prosody as revealed by rTMS," *European Journal of Neuroscience*, vol. 21, pp. 3195–3200, 2005.

[16] T. Alexander, K. Avirame, and M. Lavidor, "Improving emotional prosody detection in the attending ear by cathodal tDCS suppression of the competing channel," *Neuroscience Letters*, vol. 508, pp. 52–55, 2012.

[17] M.A. Nitsche, D. Liebetanz, N. Lang, A. Antal, F. Tergau, and W. Paulus, "Safety criteria for transcranial direct current stimulation (tDCS) in humans," *Clinical Neurophysiology*, vol. 114, pp. 2220–2223, 2003.

[18] C. Poreisz, K. Boros, A. Antal, and W. Paulus, "Safety aspects of transcranial directcurrent stimulation concerning healthy subjects and patients," *Brain Research Bulletin*, vol. 72, pp. 208–214, 2007.

[19] A. F. DaSilva, M. S. Volz, M. Bikson, and F. Fregni, "Electrode Positioning and Montage in Transcranial Direct Current Stimulation," *Journal of Visualized Experiments*, vol. 51, e2744, 2011.

[20] B. M. Ben-David, N. Multani, N. A-M. Durham, F. Rudzicz, and P. H. H. M. Van Lieshout, "T-RES: test of rating of emotions in speech: interaction of affective cues expressed in lexical content and prosody of spoken sentences," *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics (Fechner Day 2011)*, pp. 391–396, 2011.