



A Neurolinguistic Investigation of Emotional Prosody and Verbal Components of Speech

Ali Rıza Sonkaya^{1*}, Zeynep Zeliha Bayazit²

ABSTRACT

Emotional prosody has a key role implication of timbre component, mood sense, and prosodic content. For this reason, it serves a highly important function for sense, the meaning to be reflected, and ability to provide effective communication. Due to its crucial role in verbal communication, it has a critical relationship within many disciplines such as linguistics, computer sciences, medicine, etc. Consequently, the knowledge of how the prosody sequence works in the brain will contribute to both language development and foreign language teaching as well as clinical evaluation of individuals with verbal communication difficulty. From this point of view, the current study takes an interdisciplinary perspective to address the investigation of brain localization of emotional prosody and verbal components of speech. In accordance with this purpose, the fNIRS technique was used. fNIRS has recently become popular as an emerging optical brain imaging technique for studying human brain. However, it is still not widespread compared to other neuroimaging techniques. This study was conducted on both 20 healthy native speakers of Turkish and English. Participants were recorded by using fNIRS while performing emotional prosody production and auditory stimulus tasks to measure the brain activation. Our results showed superior temporal gyrus, middle temporal gyrus, which includes primary and secondary auditory cortex, and superior temporal cortex, which comprises of temporal sulcus, were strongly activated by prosodies irrespective of emotional valence. Our findings also demonstrated left inferior frontal gyrus which comprises of pars triangularis; (Brodmann Area 45) and the frontal eye field (Brodmann Area 8) were significantly activated for happy, angry, and fearful prosodies.

Key Words: Emotional prosody, fNIRS Technique, Neurolinguistics, Speech, Brain Localization

DOI Number: 10.14704/nq.2018.16.12.1877

NeuroQuantology 2018; 16(12):50-56

Introduction

Increased interest in the relationships between the brain and behavior over the past several decades has made prosody a topic of study in disciplines like neurology, psychiatry, and neuroscience. This is because prosody is an important communicative element, pointed as one of the main extralinguistic attributes present in oral communication. It consists of the melodic line of discourse, produced by the variation in frequency, rhythm, and emission emphasis. It plays a crucial role during social

communicative interaction and also second language acquisition.

Emotional prosody which is one of the communicative elements of prosody and transferring emotion to spoken language, has a key role implication of timbre component, mood sense, and prosodic content. For this reason, it serves a highly important function for sense, the meaning to be reflected, and ability to provide effective communication. Due to its crucial role in verbal communication, it has a critical relationship within many disciplines such

Corresponding author: Ali Rıza Sonkaya, M.D.

Address: ¹Neurologist, Okmeydani Training and Research Hospital, Department of Neurology, 34384 Okmeydani-Sisli, İstanbul, Turkey; ²Neurolinguistic Researcher, Ankara University, Department of Linguistics, 06540 Cebeci, Ankara, Turkey

e-mail ✉ drsonkay@gmail.com

Relevant conflicts of interest/financial disclosures: 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: 30 August 2018; **Accepted:** 22 November 2018



as linguistics, medicine, computer sciences, etc. Consequently, the knowledge of how the prosody sequence works in the brain will contribute to both language development and foreign language teaching as well as clinical evaluation of individuals with verbal communication difficulty. For this reason, the current study takes an interdisciplinary perspective to address the investigation of brain localization of emotional prosody and verbal components of spoken language. In accordance with this purpose, we used the fNIRS, which has recently become popular as an emerging optical imaging technique for studying human brain function. It is non-invasive, portable, affordable, safe, easy-to-use and does not expose people to radiation. These advantages make it especially useful in the field of language researches. However, it is still not widespread compared to other neuroimaging techniques despite being an easier to use research tool.

Methods

Participants

This study was conducted on 20 healthy native speakers of English (7 female+13 male, mean age 21.67) between 18 and 50 years of age who are trained on intensive Turkish as a second language, from initial exposure to high proficiency and 20 native speakers of Turkish (6 female+14 male, mean age 32,24). Both groups were constituted same age and education level. All participants were right handed. They were reported to have no history of neurological and psychiatric disorder, known hearing impairment, history of drug or alcohol abuse, long periods of unconsciousness and head injuries. Self-report screening was used to assess the exclusion criteria.

The participants were recruited giving written consent form before the experiment. Prior to scheduling their participation, each step of experiment was explained verbally and also it was given the information about fNIRS to ensure that they qualified for the study. Each participant had a study information sheet, which allows them to be informed about the study. Participants were allowed to withdraw from the study at any point. The experimental protocol was approved by the Ethics Committee of Indiana University (IRB Protocol No: 1712586128) and this study was performed strictly in accordance with the approved guidelines.

Stimuli

Auditory stimuli consisted of semantically neutral four words (*sinif, alarm, doktor, gazete*), spoken in different emotional intonation with irrespective of emotional valence. The words were recorded by a native Turkish-speaking male phonetician in angry, happy, neutral, and fearful tones of voices. Among the four words, the phonetic pronunciations “alarm” and “doktor” have meaning in both Turkish and English, although they are spelled differently in English. The other two words, ‘sinif’ and ‘gazete’, although they are not part of the standard English lexicon, are widely used in the Turkish language.

The stimuli were edited to a common length of approximately 650 to 700 ms and equalized in intensity. They were then stored in 16-bit, digital format on a computer. After the fNIRS recording, all the participants were required to classify each prosodic word into one of four emotion categories. The mean recognition rate was 0.96 ± 0.03 , 0.89 ± 0.05 , 0.91 ± 0.06 , 0.98 ± 0.07 for anger, fear, happy and neutral prosodies.

Procedure

The study was conducted in a quiet room. Each participant was sat in front of the computer and experimental stimuli was shown. Emotional prosodic sounds were listened via computer more or less 60 cm from the participants. This entire study lasted approximately 30–40 minutes for each participant, including instructions and interviews. Resting-state fNIRS data were first recorded for 5 min (eyes opened), followed by a 20-min passive listening task.

The stimuli consisted of each of the four words spoken in each of the four emotions resulting in 16 separate word/emotion combinations. In each of the four experimental conditions, subjects were asked to listen for a different target and press a computer mouse when that target was detected. Resting-state NIRS data were first recorded for 5 min (eyes opened), followed by a passive listening task.

fNIRS data recording

fNIRS is designed to employ near-infrared light to non-invasively measure changes in the concentration of oxygenated (oxyHb), deoxygenated (deoxyHb) and total (tHb) hemoglobin in the brain, readily penetrating the skull and reaching cortical tissue (Wolf *et al.*, 2007, Dieler *et al.*, 2012). The NIRS data were recorded in a continuous-wave mode with the NIRScout 1624 system (NIRx Medical Technologies, LLC. Los Angeles, USA), which composed of 16 LED



emitters (intensity = 5 mW/wavelength) and 23 detectors at two wavelengths (750 and 850 nm). Based on previous studies (Zhang *et al.*, 2018, Frühholz *et al.*, 2016), optodes were calibrated in the frontal and temporal areas of the brain, using a NIRS-EEG compatible cap which designed according to the international 10/5 system. 54 useful channels, where source and detector were at a mean distance of 3 cm (range = 2.8 to 3,8 cm) from each other, were evaluated. The data were continuously sampled with 4 Hz. Detector saturation never taken place during the recording.

To assess the cortical activities underlying fNIRS channels (Zhang *et al.*, 2018, Frühholz *et al.*, 2016), a Matlab toolbox NFRI (<http://brain.job.affrc.go.jp/tools/>) was used to forecast the NMI coordinates of optodes according to the EEG 10/5 systems. The locations of fNIRS channels were determined at the central zone of the light path between each adjacent source-detector pair (Zhang *et al.*, 2018).

Statistical analyses

Statistical significance of concentration changes was adjusted a general linear model of the standard hemodynamic response function (parameters in nirsLAB = [6 16 1 1 6 0 32]), with a separated cosine transformation used for temporal filtering (high-pass frequency cutoff = 128 sec) (Zhang *et al.*, 2018). However, both oxyHB and deoxyHB signals were derived from data, in this study, only the oxyHB was chosen to perform statistical analyses because of its superior signal-to-noise ratio relative to deoxyHB. When forecasting beta, nirsLAB used a SPM-based algorithm to calculate a least-squares solution to an over described system of linear equations.

To statistically analyze the data, one-way ANOVA was first performed on the beta values related with oxyHB (four levels: neutral, fearful, happy and angry prosody), concluding in a thresholded (corrected $p < 0.05$) F -statistic map. The statistical results in individual channels were corrected for multiple comparisons across channels by the false discovery rate (FDR), following the Benjamin and Hochberg procedure performed in Matlab (v2015b, the Mathworks, Inc., Natick, USA) (Zhang *et al.*, 2018).

Results

According to the one-way ANOVA results, our findings showed 7 fNIRS channels (3, 8, 15, 30, 35, 36 and 51) had divergent activation patterns across the four experimental conditions which consist of neutral, angry, fearful and happy prosodies.

Emotional prosody > neutral prosody comparison

First of all, we investigated the brain areas which were more activated for emotional compared to neutral prosodies. The t -test demonstrated when compared to neutral prosodies, two channels had significantly augmented activations in response to emotional prosodies, corresponding to brain regions of right posterior STG (BA 22, Channel 51; $t(17) = 3.07$, $p < 0.001$, corrected $p = 0.047$) and right primary/secondary AC (BA 42, Channel 48; $t(17) = 3.57$, $p = 0.004$, corrected $p = 0.033$). It is significant that while the main effect of prosodies had leftward lateralization in the posterior STG (paired-samples t -test: $t(17) = 3.66$, $p = 0.021$) and primary/secondary AC, the contrast of emotional and neutral prosodies within these areas showed rightward lateralization (AC: $t(17) = -4.57$, $p = 0.003$; STG: $t(17) = -4.27$, $p = 0.001$; Fig. 1).

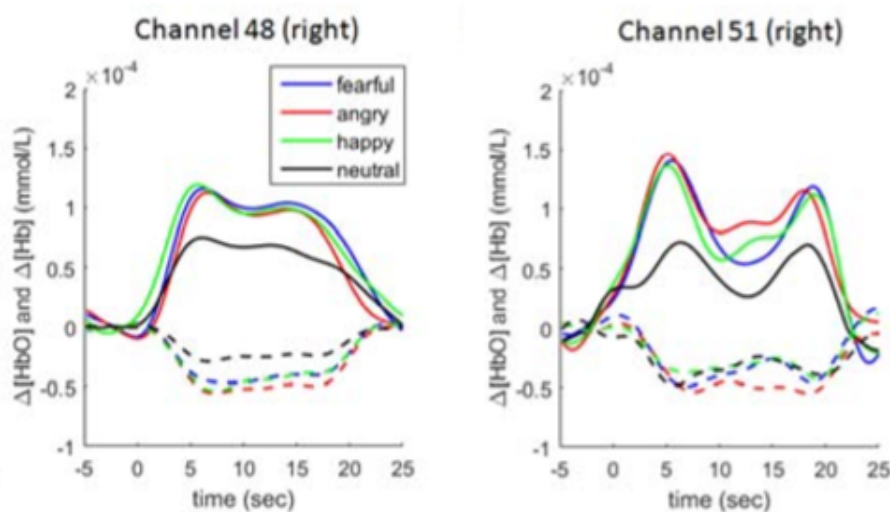


Figure 1. The time course of oxyHB and deoxyHB in respond to the four prosodies. The four subplots display the waveforms at A, right primary auditory cortex (Channel 48); B, right posterior superior temporal gyrus (Channel 51)



Positive > negative prosody comparison

Second, we investigated the brain areas which were more activated for happy compared to angry and fearful prosody. The *t*-test demonstrated when compared to negative prosody, two channels had significantly augmented activations in response to positive (happy) prosody. The related brain areas were left pars triangularis (BA 45, Channel 15; $t(17) = 2.87, p = 0.003$, corrected $p = 0.048$) and frontal eye fields (BA 8, Channel 34; $t(17) = 3.71, p = 0.003$, corrected $p = 0.047$). It is important that while the main effect of prosody had rightward lateralization in the middle inferior frontal gyrus (paired-samples *t*-test: $t(17) = -3.25, p = 0.020$), the contrast of happy and fearful/angry prosody revealed leftward lateralization ($t(17) = 3.16, p = 0.021$; Fig. 2).

Happy > neutral prosody comparison

Third, we investigated the brain regions which were more activated for happy compared to neutral prosody. The *t*-test demonstrated that Channel 15 had significantly augmented activations in response to happy prosody ($t(17) = 3.15, p < 0.001$, corrected $p = 0.047$). The related brain areas were left pars triangularis which corresponds to BA 45 (See Fig. 2).

Angry > neutral prosody comparison

Fourth, we investigated the brain areas which were more activated for angry compared to neutral prosody. The *t*-test demonstrated that two channels, which were symmetrical, had significantly augmented activations in response to angry prosodies. These channels are corresponding to frontopolar and

orbito-frontal areas which are the part of orbito-frontal cortex and represented BA 10/11. But the activation was not statically significant when multiple comparison correction was performed (Channel 3: $t(17) = 4.56, p = 0.003$, corrected $p = 0.069$; Channel 30: $t(17) = 4.68, p = 0.004$, corrected $p = 0.072$; Fig. 3).

Fearful > neutral prosody comparison

Finally, we investigated the brain areas which were more activated for fearful compared to neutral prosody. There were no channels significantly activated even before multiple comparison correction.

Discussion

In this study, the superior temporal cortex has been shown to take a critical part in decoding vocal statements of emotions. While the lower-level structures of superior temporal cortex analyze acoustic features in auditory statements, the higher-level structures of superior temporal cortex compound the decoded auditory features and build up percepts of vocal statements (Zhang *et al.*, 2018). Consistent with this consideration, the current study evidenced that while speech prosodies activated the left primary auditory cortex (BA 42) most significantly but emotional prosodies activated the right superior temporal cortex (BA 22/42) when compared to neutral prosodies.

Our results ensure further evidence to clears up the lateralization of emotional prosody processing in the superior temporal cortex. We observed that presentation of speech stimuli demonstrated significant leftward lateralization in

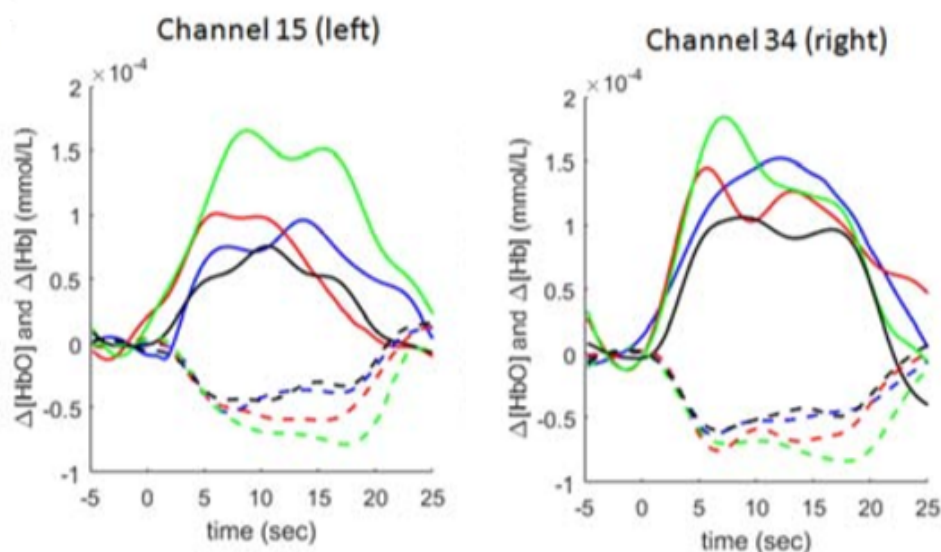


Figure 2. The time course of oxyHB and deoxyHB in respond to the four prosodies. The four subplots display the waveforms at bilateral middle inferior frontal gyrus (Channel 15 and 34)



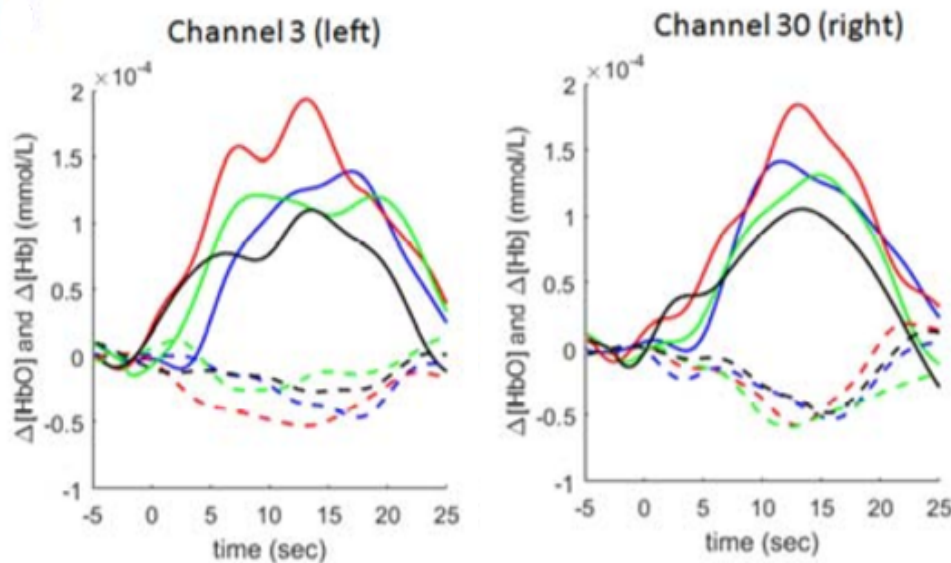


Figure 3. The time course of oxyHB and deoxyHB in respond to the four prosodies. The four subplots display the waveforms at bilateral frontopolar and orbito-frontal area (Channel 3 and 30).

the both primary and secondary auditory cortices and posterior superior temporal gyrus, which is parallel with the consideration that the left hemisphere is better accoutered for the analysis of swiftly changing phonetic representations in spoken language (Kotz *et al.*, 2013, Wildgruber *et al.*, 2005). However, our findings demonstrated a significant right lateralization for emotional prosody perception within the superior temporal cortex, which is consistent with the previous findings that the right hemisphere is more susceptible to slow-varying acoustic profiles of emotions (Beaucousin *et al.*, 2006, Vigneau *et al.*, 2011, Zhang *et al.*, 2018).

It is also highlighting that although the cortex responses were explored within five contrasts, the superior temporal cortex demonstrated significant activations only within the first contrasts which taken place the emotional and neutral prosody comparison. This finding suggests that the superior temporal cortex may be shadowed out in general response to emotional prosodies regardless of valence or emotional categories, which is compatibly many previous studies investigating a U-shaped dependency between brain activation and valence of prosodies in the superior temporal cortex (Ethofer *et al.*, 2006, Frühholz *et al.*, 2011).

One authentic finding of this study is that the pars triangularis (BA 45) and the frontal eye field (BA8), which are located in the left inferior frontal gyrus, were significantly demonstrated activation for happy relative to angry/ fearful prosodies. It has mostly been reported that the pars triangularis

of the inferior frontal gyrus plays a crucial role in semantic comprehension (Goucha and Friederici, 2015, Schirmer and Kotz, 2006). In this study, the finding of the higher tendency to semantically process happy relative to fearful and angry prosodies may be due to the positivity offset (Zhang *et al.*, 2018). The participants were more energized to comprehend happy prosodies though they were only required to passively listen. Since both Turkish and English sentences were used in the study, this potential semantic procedure may also activate the BA 8 (English sentences for Turkish native speakers and Turkish sentences for English native speakers), which is contained in the management of uncertainty.

Neural bases of happy prosody process were examined in previous studies. In line with this study, Zhang *et al.* (2018) observed happy relative to neutral prosodies activated left inferior frontal gyrus (Kotz *et al.*, 2013). Another study which was conducted by Johnstone *et al.* (2006) found enhanced activation in right inferior frontal gyrus for happy relative to angry prosodies (Johnstone *et al.*, 2006). The incongruent lateralization of inferior frontal gyrus activation may be due to the differences in stimuli, i.e., the participants in this study and in Zhang *et al.* (2018) only listened to speech prosodies but the participants listened to prosodies and watched congruent or incongruent facial expressions at the same time in Johnstone *et al.* (2006) (Zhang *et al.*, 2018). The contrast of happy to neutral prosody in this study is overlapping the results of Zhang *et al.* (2018).

One important finding is the significant activation in bilateral orbito-frontal cortex (BA 10/11) for angry compared to neutral prosody, which is almost overlapping the finding of Kotz *et al.* (2013). The orbito-frontal cortex, which is a key neural correlate of anger (Lindquist *et al.*, 2012), plays an important role in conflict resolution and suppression of inappropriate behavior such as aggression (Zhang *et al.*, 2018). Hornak *et al.* (2003) indicated the impairment of patients with bilateral damages of the orbito-frontal cortex while the identification of voice expression. Also authors had observed the significant changes in patients' subjective emotional state (Hornak *et al.*, 2003). Previous fMRI studies comparing angry to neutral prosodies have demonstrated different results: while some researchers claim that the bilateral frontal regions such as the orbito-frontal cortex are generally recruited regardless of implicit and explicit tasks (Quadflieg *et al.*, 2008, Zhang *et al.*, 2018), some others believed that only in explicit tasks the bilateral orbito-frontal cortex be affected to angry prosodies (Ethofer *et al.*, 2006, Sander *et al.*, 2005). Taking in consideration the passive listening task in this study, authors think the current finding supports the former notion.

Interestingly, no significant brain activations were observed for fearful compared to neutral prosody. The result seems inconsistent with the consideration of "the negativity bias" that thinks fit the processing of fearful faces/pictures/ words (Cacioppo *et al.*, 1999, Ito *et al.*, 1998). It may be related with the quickly processing of visual emotional stimuli, which contributes individuals to start a timely fight-or-flight behavior, but communication of emotional prosodies have no biologically salient cues because of their fine-grained characteristics (e.g. pitch, loudness contour, and rhythm) evolve on a long time scale (Liebenthal *et al.*, 2016).

Conclusion

In this study, fNIRS technique was used to investigate how speech prosodies of different emotional categories are processed in the cortex. Taken together, the present findings suggest that while emotional prosody processing within the superior temporal cortex primarily plays a role to distinguish between emotional and neutral stimuli, categorization of emotions might take place within a high-level brain region—the frontal cortex. The results confirmed and overlapped previous fMRI findings in adult brain and

also ensured a "developed version" of brain activation for the following neonatal research.

Acknowledgements

This study was funded by "The Scientific and Technological Research Council of Turkey (TUBITAK)" (1059B141600335).

References

- Beaucousin V, Lacheret A, Turbelin MR, Morel M, Mazoyer B, Tzourio-Mazoyer N. FMRI study of emotional speech comprehension. *Cerebral cortex*, 2006; 17(2): 339-52.
- Cacioppo JT, Gardner WL, Berntson GG. The affect system has parallel and integrative processing components: Form follows function. *Journal of personality and Social Psychology*, 1999; 76(5): 839.
- Dieler AC, Tupak SV, Fallgatter AJ. Functional near-infrared spectroscopy for the assessment of speech related tasks. *Brain and language*, 2012; 121(2): 90-109.
- Ethofer T, Anders S, Wiethoff S, Erb M, Herbert C, Saur R, Grodd W, Wildgruber D. Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport*, 2006; 17(3): 249-53.
- Frühholz S, Ceravolo L, Grandjean D. Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*, 2011; 22(5): 1107-17.
- Frühholz S, Trost W, Kotz SA. The sound of emotions—Towards a unifying neural network perspective of affective sound processing. *Neuroscience & Biobehavioral Reviews*, 2016; 68: 96-110.
- Goucha T, Friederici AD. The language skeleton after dissecting meaning: a functional segregation within Broca's Area. *Neuroimage*, 2015; 114: 294-302.
- Hornak J, Bramham J, Rolls ET, Morris RG, O'Doherty J, Bullock PR, Polkey CE. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, 2003; 126(7): 1691-712.
- Ito TA, Larsen JT, Smith NK, Cacioppo JT. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of personality and social psychology*, 1998; 75(4): 887.
- Johnstone T, Van Reekum CM, Oakes TR, Davidson RJ. The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience*, 2006; 1(3): 242-9.
- Kotz SA, Kalberlah C, Bahlmann J, Friederici AD, Haynes JD. Predicting vocal emotion expressions from the human brain. *Human Brain Mapping*, 2013; 34(8): 1971-81.
- Liebenthal E, Silbersweig DA, Stern E. The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in neuroscience*, 2016; 10: 506.
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF. The brain basis of emotion: a meta-analytic review. *Behavioral and brain sciences*, 2012; 35(3): 121-43.



- Quadflieg S, Mohr A, Mentzel HJ, Miltner WH, Straube T. Modulation of the neural network involved in the processing of anger prosody: the role of task-relevance and social phobia. *Biological psychology*, 2008; 78(2): 129-37.
- Sander D, Grandjean D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P. Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage*, 2005; 28(4): 848-58.
- Schirmer A, Kotz SA. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in cognitive sciences*, 2006; 10(1): 24-30.
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, 2011; 54(1): 577-93.
- Wildgruber D, Riecker A, Hertrich I, Erb M, Grodd W, Ethofer T, Ackermann H. Identification of emotional intonation evaluated by fMRI. *Neuroimage*, 2005; 24(4): 1233-41.
- Wolf M, Ferrari M, Quaresima V. Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *Journal of biomedical optics*, 2007; 12(6): 062104.
- Zhang D, Zhou Y, Yuan J. Speech Prosodies of Different Emotional Categories Activate Different Brain Regions in Adult Cortex: an fNIRS Study. *Scientific reports*, 2018; 8(1): 218.

