DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

The building blocks of social communication

Citation
Niznikiewicz, Margaret A. 2013. "The building blocks of social communication." Advances in Cognitive Psychology 9 (4): 173-183. doi:10.2478/v10053-008-0145-6. http://dx.doi.org/10.2478/v10053-008-0145-6.

Published Version
doi:10.2478/v10053-008-0145-6

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:12064515

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

ADVANCES IN COGNITIVE PSYCHOLOGY

The building blocks of social communication

Margaret A. Niznikiewicz

Boston VA Healthcare Center, Harvard Medical School, Boston, MA, USA

ABSTRACT

In the present review, social communication will be discussed in the context of social cognition, and cold and hot cognition. The review presents research on prosody, processing of faces, multimodal processing of voice and face, and the impact of emotion on constructing semantic meaning. Since the focus of this mini review is on brain processes involved in these cognitive functions, the bulk of evidence presented will be from event related potential (ERP) studies as this methodology offers the best temporal resolution of cognitive events under study. The argument is made that social communication is accomplished via fast acting sensory processes and later, top down processes. Future directions both in terms of methodology and research questions are also discussed.

KEYWORDS

social cognition, ERP, social communication

INTRODUCTION

After years of intense interest in different aspects of sensory and cognitive processes, more recently dubbed "cold cognition," that included sensory analyses, attention, memory, language, and conflict resolution, to name a few, researchers realized that people are first and foremost social beings. Thus, one of the most important functions which these processes serve is social communication (Brück, Kreifelts, & Wildgruber, 2011; Decety & Svetlova, 2012). Neuroscience has contributed enormously to our understanding of how cold cognition processes are supported by the brain's neural architecture and function (e.g., D'Ardenne et al., 2012; Newman, Carpenter, & Varma, 2003; Peelle & Davis, 2012; Poeppel, Emmorey, Hickok, & Pylkkänen, 2012). Likewise, technological advances and novel designs within the field of neuroscience allowed studying elements of social cognition at the brain level. They also significantly contributed to the broadening of a scope of inquiry into its constituents (Arlinger et al., 2009; Barraclough & Perrett, 2011). Accordingly, recent years have seen an enormous interest in the topic of *social cognition* defined as the ability to understand and interpret the intentions and emotions of others and adaptively react to these signals (Blakemore, 2012; Frith, 2012; Lieberman, 2012). In parallel to developments in research on social cognition, a line of research developed on the role of emotion and its interactions with other cognitive functions. Dubbed research on "hot cognition," it examined different aspects of emotion processing. Like research on cold cognition and on social cognition, it has been enormously enriched by a neuroscience perspective. In Figure 1, I describe briefly, and non-exhaustively, research subsumed under the different rubrics of cold cognition, hot cognition, and social cognition, with an area "targeted" in this brief review marked with a star (see Figure 1). I will argue that in order to fully characterize human interactions with others, we should be focusing not only on the social cognition but, more broadly, on social communication. Social communication includes not only social cognition but also language and its interactions with emotion.

The building blocks of social communication are sensory processes that provide rapid analyses of incoming external stimuli, attentional processes that select sensory data for further processing as relevant for a task at hand, working memory processes that enable maintaining relevant information, and long-term memory structure and processes within it that allow rapid comparisons between oncoming information and existing semantic and semiotic knowledge (see Figures 1 and 2). A formal language system, both written and spoken, with the latter including syntactic and emotional prosody, is an important tool of social communication. It is only rivaled by facial and body gestures and expressions in conveying a full spectrum of human information,

Corresponding author: Margaret A. Niznikiewicz, Boston VA Healthcare Center, Harvard Medical School, Psychiatry 116A, 940 Belmont Street, Brockton, MA 02301, USA. E-mail: margaret_niznikiewicz@hms.harvard.edu

2013 • volume 9(4) • 173-183

http://www.ac-psych.org

DOI • 10.2478/v10053-008-0145-6

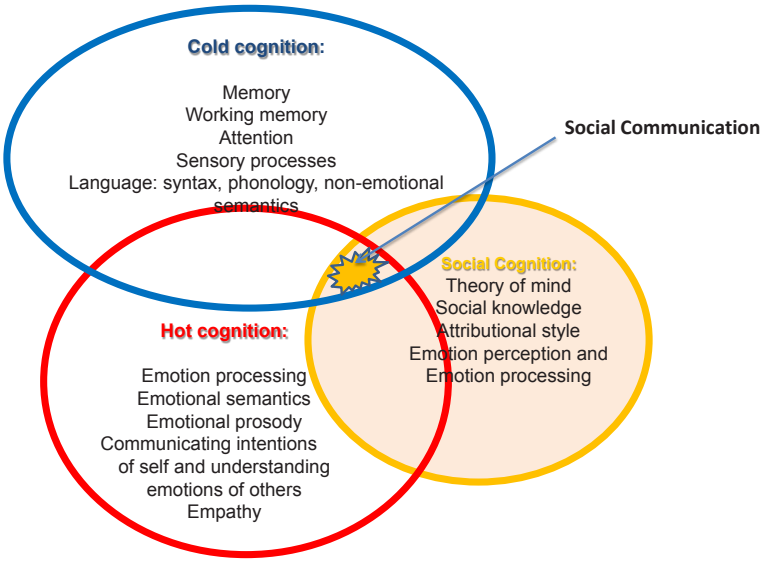

FIGURE 1.

A model of interdependencies between "cold cognition," "hot cognition," and "social cognition." The building blocks of social communication belong to all three domains.

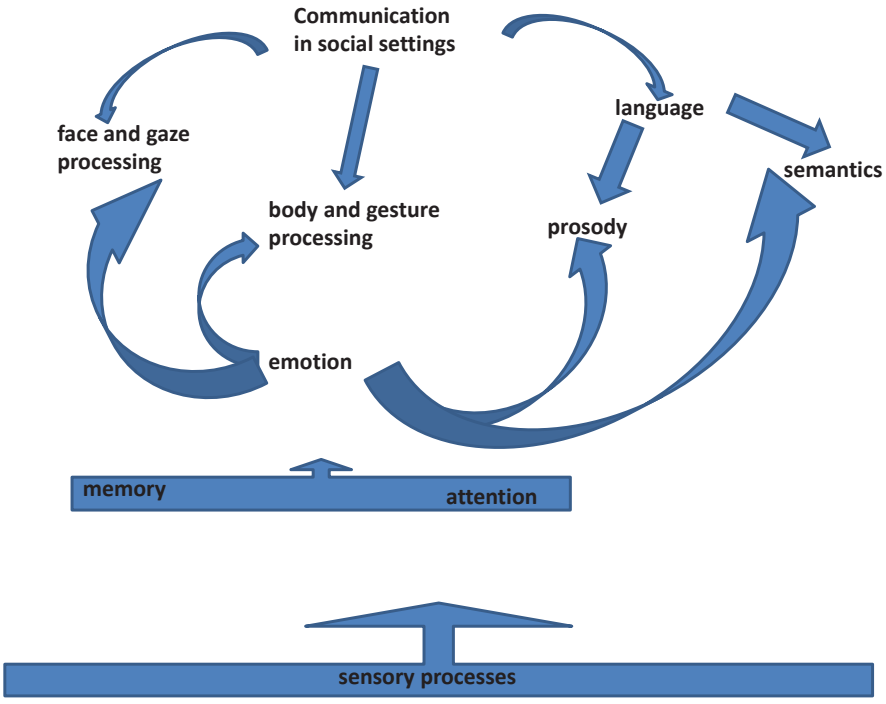

FIGURE 2.

Social communication relies on language and its constituents, on gaze and facial expressions, as well as on body gestures. Sensory processes modulated by attention and memory processes provide raw data from which social meaning is constructed across all domains.

intentions, attitudes, and emotional states (see Figure 2). It is impossible to cover a topic of this complexity in a single paper.

Therefore, I will sample from different domains of research on social communication and focus on how neurophysiological processes contribute to its different instances.

As described above, the devices that people use to communicate with each other include their voice and its modulations: The voice modulations mark both syntactic structure of a sentence and emotional attitudes of a speaker (by prosody). They also include facial expressions and language. More recently, science has started to ask questions about brain processes involved in prosody and face processing as well as about how these two sources of affective information, voice and face, are processed together by the brain (see e.g., Ethofer et al., 2013; Jacob, Brück, Domin, Lotze, & Wildgruber, 2013). Another domain recently examined is the influence of affective states on cognitive processing, especially on language processing (e.g., Federmeier et al., 2001).

As indicated above, in this review, the emphasis will be on identifying neurophysiological processes that underpin these phenomena. Therefore, the methodology that I will focus on will be event related potentials (ERPs). ERPs remain the only existing technology that allows for capturing neurocognitive events with millisecond resolution as they unfold in real time. The overarching question in this enquiry into the different aspects of social communication is when the neurocognitive system distinguishes between different classes of sensory data such that, at a later stage, they become meaningful "building blocks" of social communication.

Accordingly, I will discuss studies examining prosody processing, including processing of non-verbal emotional vocalizations, face processing, simultaneous voice and face processing, and effects of emotion on semantic representations.

PROSODY PROCESSING

With a tone of voice we can express that we are happy, sad, angry, disappointed, or sarcastic. *Emotional prosody* refers to a tone of voice with which we speak. Prosodic information in a voice signal is primarily carried by fundamental frequency (F0) with contributions from other phonetic devices such as voice timbre, intensity, and speech rhythm (Schirmer & Kotz, 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). Recent functional studies established a network of brain regions that seem to be involved in prosody processing. They include superior and middle temporal gyrus (STG and MTG), parietal-temporal juncture, insula, as well as inferior frontal gyrus and orbitofrontal gyrus (Leitman et al., 2011; Schirmer & Kotz, 2006). ERP studies, on the other hand, pointed to the speed with which salient sensory features differentiating between different prosody types can be detected. Early prosody studies focused on exploring the question whether a neural system is sensitive to changes from one type of prosody to another. Most of these studies used a splicing technique that cuts an utterance into two different parts where the first one is said, for example, with a happy voice, and the second one with a sad voice (Paulmann & Kotz, 2008). This approach yields findings that point to the brain's sensitivity to a shift of voice. Another class of ERP prosody studies focused on natural prosodic utterances (Kotz & Paulmann, 2007; Paulmann & Kotz, 2008; Pinheiro, Del Re, Mezin, et al., 2012). These studies point directly to the speed with which prosodic information is decoded and hint at the processing stages involved. In addition, they provide information on the functional architecture of the auditory system as applied to prosody processing.

All ERP studies of naturally occurring prosody report the N100 and P200 components of the waveform. The N100 is sensitive to sensory, physical properties of an acoustic signal, while the P200 has been associated with assigning sensory data different categorical values (Rosburg, Boutros, & Ford, 2008). Both components are believed to be modulated by attention. Since, by definition, the N100 peaks about 100 ms, and P200 about 200 ms after the onset of the stimulus, it appears that the first operations that allow assigning emotional valence to an auditory signal happen within the 200 ms from the onset of that signal. Most studies that focused on the analyses of the ERP correlates of natural prosody processing formally examined only the P200 component. Most of them indicated that, at the level of the P200, the prosodic signal is categorized into emotional and non-emotional rather than into discrete emotional categories such as happy, angry, or sad (Kotz & Paulmann, 2007; Paulmann & Kotz, 2008; but see Pinheiro, Del Re, Mezin, et al. 2012, where the distinction was found in the P200 amplitude between neutral relative to happy, relative to angry prosody).

The Pinheiro, Del Re, Mezin, et al. study (2012) was one of the first to formally analyze the N100 in terms of its sensitivity to prosodic information. We used simple neutral sentences that either carried semantic information intact or removed by an acoustic manipulation. In semantically intact sentences, the N100 was more negative to neutral prosody relative to angry prosody (see Figure 3). For the same stimuli, when they were stripped off their semantic content, the N100 was more negative to the neutral relative to the emotional prosody (see Figure 3). These results lead to the conclusion that emotional meaning is constructed from the sensory data whose physical properties are processed differentially from the early stages of analyses. In a series of analytic steps, they are endowed, or tagged, with increasingly rich "meaning." These data also suggest that physical properties of a signal matter, such as in a differential N100 response to intelligible versus non-intelligible speech. Since in normal populations the size of the N100 is associated with the ease of processing a given stimulus, we can further speculate that the physical properties of neutral undistorted speech are easier to process than the same speech when it is distorted. Interestingly, the pattern for the emotional prosody for the distorted and undistorted speech was opposite to that observed for the neutral prosody. It remains an issue of further research to examine whether these differences have to do solely with the physical properties of the speech signal or whether they are mediated via attentional and, perhaps, memory mechanisms.

The P200 in this study was also sensitive to prosodic information similarly to previous ERP prosody studies. Similarly to the N100 results, the pattern of amplitude differences was different in sentences with and without semantic content. Interestingly, unlike in earlier

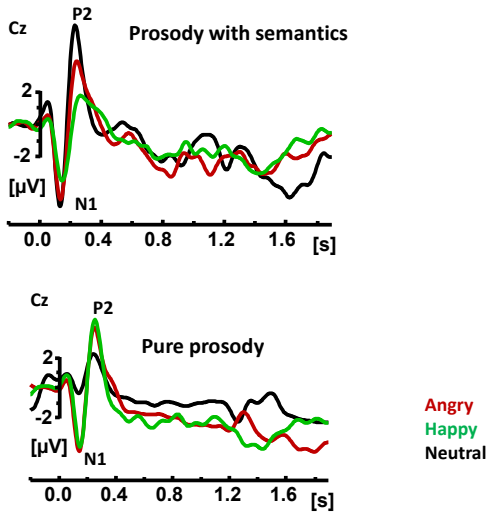

FIGURE 3.

Grand average waveforms timelocked to the onset of a sentence to sentences with prosodic and semantic content and to pure prosody sentences in 15 individuals. The prosody types were neutral, happy, and angry. Adopted from "Sensory-Based and Higher-Order Operations Contribute to Abnormal Emotional Prosody Processing in Schizophrenia: An Electrophysiological Investigation," by A. P. Pinheiro, E. Del Re, J. Mezin, P. G. Nestor, A. Rauber, R. W. McCarley, et al., 2012, *Psychological Medicine*, *43*, 603-618.[1]

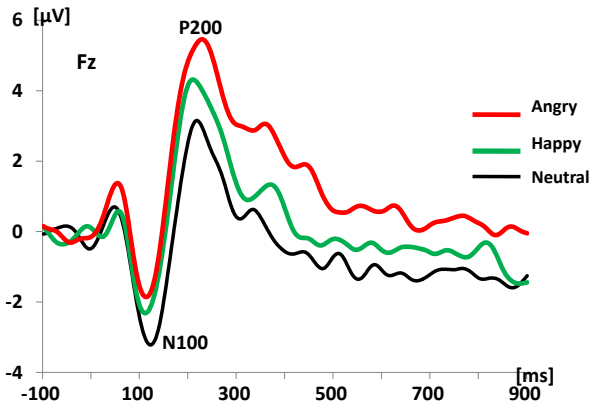

FIGURE 4.

Grand average waveforms to non-semantic vocalizations: neutral, angry, and happy in 19 individuals. Adopted from "Electrophysiological Insights Into Processing Non-Verbal Emotional Vocalizations," by T. Liu, A. P. Pinheiro, G. Deng, P. G. Nestor, R.W. McCarley, and M. A. Niznikiewicz, 2012, *NeuroReport*, *23*, p. 108-112 (see Footnote 1).

studies, the P200 amplitude did distinguish between different types of prosody: neutral, happy, and angry (see Figure 3).

The speed with which prosodic information is processed and a role of sensory and early categorization processes were further explored in a study of non-verbal human sounds that connoted neutral, happy, and angry emotional states (Liu, Pinheiro, Deng, et al., 2012). In that study, we used neutral "mm" sounds, laughs, and angry sounds to examine whether such non-semantic emotional sounds will be associated with similar brain responses to those observed for sentences. There were notable similarities and differences between the ERP results recorded to non-semantic emotional vocalizations relative to emotional prosody in sentences. As in the study of prosody processing in sentences, both N100 and P200 were sensitive to different emotional voices (see Figure 4). However, for these non-semantic sounds, the first differentiation between different vocal emotions was observed at the level of P50 which was more positive to angry sounds relative to either happy or neutral vocalizations (see Figure 4). One can only speculate that perhaps angry sounds have the most ecologically valid salience relative to the two other emotions: happy and neutral. The N100 effects were similar to those observed in the sentence prosody study in that the N100 was more negative to neutral relative to emotional vocalizations. However, the P200 effects were similar to those observed in the sentence prosody study for distorted (i.e., non-semantic) sentences in that the P200 was more positive to emotional relative to neutral vocalizations (see Figure 4).

Together, these results suggest that vocal emotion, whether it is a patterned prosody found in sentences or emotional non-semantic vocalizations, is processed by the brain fairly rapidly within the first 200 ms after a stimulus onset. However, the specific processes as indexed by the N100 and P200, and possibly by earlier components, will depend on the particular sensory properties of an acoustic signal, and perhaps, on interactions of these sensory processes with modulations from attentional and memory processes.

FACE PROCESSING

The face is another important source of social information. Even without using words we can express a whole range of emotions ranging from neutral expressions through boredom, happiness, to fear, and anger. Functional imaging studies point to fusiform gyrus, and especially fusiform face area as well as occipital face area as involved in processing information from the face (e.g., Rhodes, Byatt, Michie, & Puce, 2004) with contributions from inferior frontal gyrus, amygdala, and superior temporal sulcus (Cohen Kadosh, Cohen Kadosh, Dick, & Johnson, 2011; Cohen Kadosh, Johnson, Henson, Dick, & Blakemore, 2012; Skelly & Decety, 2012). It takes the first decade of a person's life to develop efficient strategies to process faces.

The first reports of an ERP response specific to faces came from McCarthy and Puce who identified N170 (Allison, Puce, Spencer, & McCarthy, 1999; Puce, Allison, & McCarthy, 1999) as the potential sensitive to face processing, especially structural face processing (McCarthy, Puce, Belger, & Allison, 1999), with numerous studies following and examining different aspects of face analysis (for a

review, see Boehm & Paller, 2006). The N250 was also reported to be sensitive to emotional face processing (for a review, see Eimer & Holmes, 2007).

As with prosody processing, it is noteworthy that the first differentiation of sensory data related to face falls within the first 200 ms from the onset of a stimulus.

MULTIMODAL PROCESSING OF FACE AND VOICE

As reviewed above, the first wave of studies on the brain correlates of emotion processing focused separately on the processing of voice and of face information. However, in real life, voice and face information is rarely processed separately and more recent studies addressed the issue of simultaneous processing of socially relevant information from these two modalities. The questions addressed by a handful of studies on multimodal face and voice processing focused on two major issues: (a) Which brain regions/networks uniquely support multimodal face and voice processing (functional magnetic resonance imaging [fMRI] studies), and (b) what cognitive processes are associated with multimodal face and voice processing. Thus far, most studies have been conducted using fMRI methodology (e.g., Klasen et al., 2012). In spite of the fact that many details of the brain architecture involved in different aspects of multimodal face and voice processing are missing, the network of main regions involved has been delineated. Kreifelts et al.'s (2007) fMRI study of several emotions (neutral emotion included) presented in auditory, visual, and audiovisual modalities identified posterior superior temporal gyrus (pSTG) and right thalamus as regions which showed more activation in the audiovisual relative to auditory or visual stimuli. In addition, better accuracy of emotion identification was observed for audiovisual stimuli. In the follow-up study using the same type of stimuli, Kreitfelts et al. (2009) identified different parts of the superior temporal sulcus (STS) as more sensitive to information from voice and face; in that study, the area sensitive to audiovisual information was at the interface between the regions mostly sensitive to voice or to face. Joassin et al. (2011) examined cross modal interactions associated with the process of a person's identity recognition. Using voices, static faces, and face plus voice stimuli, the authors identified both unimodal regions supporting voice and face processing, and multimodal regions including the left angular gyrus and the right hippocampus as sensitive to the processing of face-voice pairings.

The two ERP studies examining temporal dynamics of information from face and voice (Jessen & Kotz, 2011; Latinus, VanRullen, & Taylor, 2010) emphasized early sensitivity to multimodal information. In the Latinus et al. study, the early effects were found within 30-100 ms and at later stages around 180-320 ms, post-stimulus. In the Jessen and Kotz study, the effects were found on the N100 that was reduced to audiovisual stimuli relative to auditory stimuli.

In our inquiry into how multimodal emotional cues are processed by the brain (Liu, Pinheiro, Zhao, et al., 2012), we presented neutral, happy, and angry faces whose presentation was time-locked to the onset of neutral, happy, or angry vocalizations. A complex set of components was recorded in the audiovisual condition that was different both from a pattern of components recorded in either auditory or visual modality (see Figure 5). We observed two distinct patterns of ERP components: parietally distributed N100, N170, and P270, and fronto-centrally distributed N100, P200, N250, and P300. Of note, the parietal components, even though they were clearly sensitive to face stimuli, did not distinguish between different emotional states. Only fronto-central components showed differential sensitivity to emotional states. The components that clearly distinguished between neutral and emotional face/voice pairings were P200, N250, and P300 associated with initial categorization processes, assigning emotional valence to stimuli, and attentional processes, respectively (see Figure 5). Overall, these results suggest that a processing stream takes actually longer and is more complex for face-voice stimuli than for either voice (indexed with N100 and P200) or face alone (indexed with N170 and in some designs with P300).

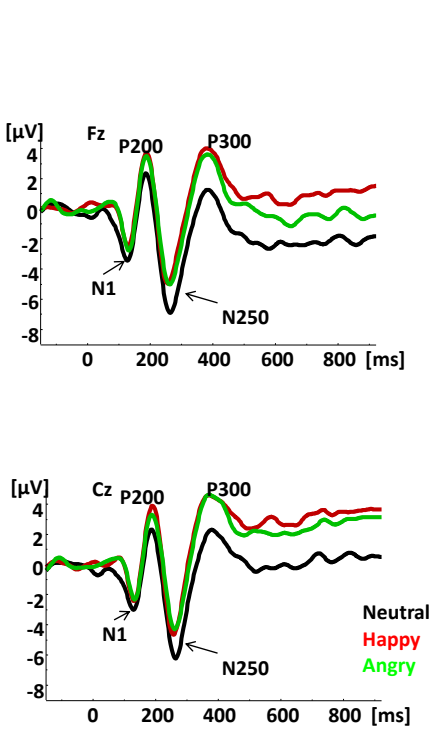

FIGURE 5.

Grand average waveforms to face and voice (multimodal) stimuli. Panel A. Fronto-central components sensitive to the distinction between neutral and emotional face/voice pairings. Adopted from "Emotional Cues During Simultaneous Face and Voice Processing: Electrophysiological Insights," by T. Liu, A. P. Pinheiro, Z. Zhao, P. G. Nestor, R. W. McCarley, and M. A. Niznikiewicz, 2012, *PLOS ONE*, *7*(2), e31001 (see Footnote 1).

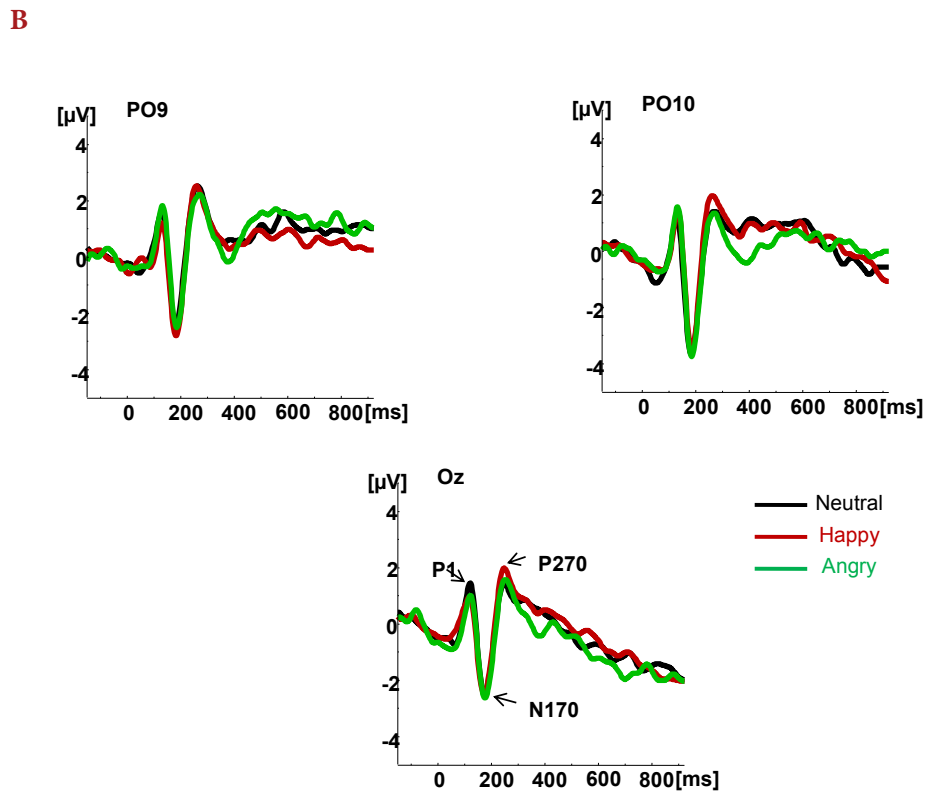

FIGURE 5 (CONTINUED).

Grand average waveforms to face and voice (multimodal) stimuli. Panel B. Parietal-occipital components indexing structural face processing. Adopted from "Emotional Cues During Simultaneous Face and Voice Processing: Electrophysiological Insights," by T. Liu, A. P. Pinheiro, Z. Zhao, P. G. Nestor, R. W. McCarley, and M. A. Niznikiewicz, 2012, *PLOS ONE*, *7*(2), e31001 (see Footnote 1).

THE INTERACTIONS BETWEEN MOOD AND SEMANTIC MEMORY CONTENTS

The component that is sensitive to semantic incongruity is the N400 that peaks around 400 ms after a target word was presented (Deacon, Hewitt, Yang, & Nagata, 2000; Holcomb, 1993; Kutas & Federmeier, 2000, 2011). The N400 is least negative if a word fits into the preceding context well, and is significantly more negative if it does not (Kutas & Federmeier, 2011). The N400 has been used successfully in numerous studies of language to probe processes and structure of semantic memory, as well as the way in which meaning is constructed by the human brain (e.g., Federmeier & Kutas, 1999a, 1999b; Kutas & Federmeier, 2011).

More recently, it has been demonstrated that the N400 can be a useful probe of the way in which emotional states influence the processing of meaning. Behavioral studies suggest that positive mood helps generate more associations and results in greater cognitive flexibility, but also contributes to making more errors. In contrast, negative mood narrows a range of associations but is also related to making fewer errors (e.g., Bolte, Goschke, & Kuhl, 2003; Dreisbach, 2006; Dreisbach & Goschke, 2004; Fiedler, 1988, 2001; Isen & Daubman, 1984; Isen, Niedenthal, & Cantor, 1992). However, behavioral data do not allow inferences about neurophysiological processes that underpin these effects.

A handful of ERP studies examined an interaction between semantics and affect from several related angles. Overall, their results strongly suggest interactions between semantics and emotion with specific studies providing evidence for different facets of these interactions. Goerlich et al. (2012) demonstrated that meaning carried by single words is processed faster and is associated with a more reduced N400 if the target word is preceded by an emotional stimulus congruent with the valence of the target word. An interaction between information structure and emotion was examined in the Wang, Bastiaansen, Yang, and Hagoort (2013) study. In that study, an interaction was found be-

tween emotionally salient words and information structure at the level of the N400: Semantic integration was influenced by information structure only for neutral words but not for emotional words. It was suggested that the greater emotional salience of emotional words overrode the influence of information structure. Chwilla, Virgillito, and Vissers (2011) took a slightly different approach by examining an interaction between N400 and mood where subjects' mood was manipulated with a mood induction procedure. The study was designed to test two competing theories of language: Embodied theories of language use suggest that symbols used in language are grounded in perception, action, and emotion, while abstract symbol theories suggest that meaning is constructed from syntactic combinations of abstract symbols (Chwilla et al., 2011). Within this context, interactions between semantics and emotion are treated as support for the embodied theories of language use. In this study, the N400 was differentially affected by induced mood lending support for the embodied theories of language.

Even though the Federmeier, Kirson, Moreno, and Kutas (2001) study can be interpreted within the same framework as the Chwilla et al. (2011) study, the authors emphasized interactions between the contents and structure of semantic memory and emotional states of language users. They tested directly the premise that positive mood broadens a pool of semantic associations and thus makes it easier to accept words that fit context less well. The authors used a series of two-sentence mini-paragraphs in which the last word in a second sentence was manipulated such that it either formed a well-form ending (expected ending), an ending that was incongruent with a sentence previous context but belonged to the same category as the correct sentence ending (within category violation), or was an incorrect sentence ending from a different semantic category (between category violation). For example, expected ending: "Paul loved to watch all TV shows that were set in hospitals. He has always wanted to be a *doctor*"; within category violation: "Louise was suffering from a toothache for several days, but she still refused to do anything about it. She has always been afraid of going to the *doctor*"; between category violation: "Michaela loved to read and she couldn't wait to check out a new set of books. She knew they were set aside for her by the friendly *doctor*" (examples from Pinheiro, Del Re, Nestor, et al., 2012). Accordingly, under positive mood, the N400 to within category violations (i.e., to words that did not fit the context well but belonged to the same category as the targets) was not different from the N400 to expected endings. This result confirmed that positive mood indeed broadens a pool of semantic associations such that words poorly fitting the context but associated with a good fit are treated by the cognitive system as acceptable – as indexed by a reduced N400.

We have replicated and extended this finding by testing the effects of both positive and negative moods using a new set of stimuli

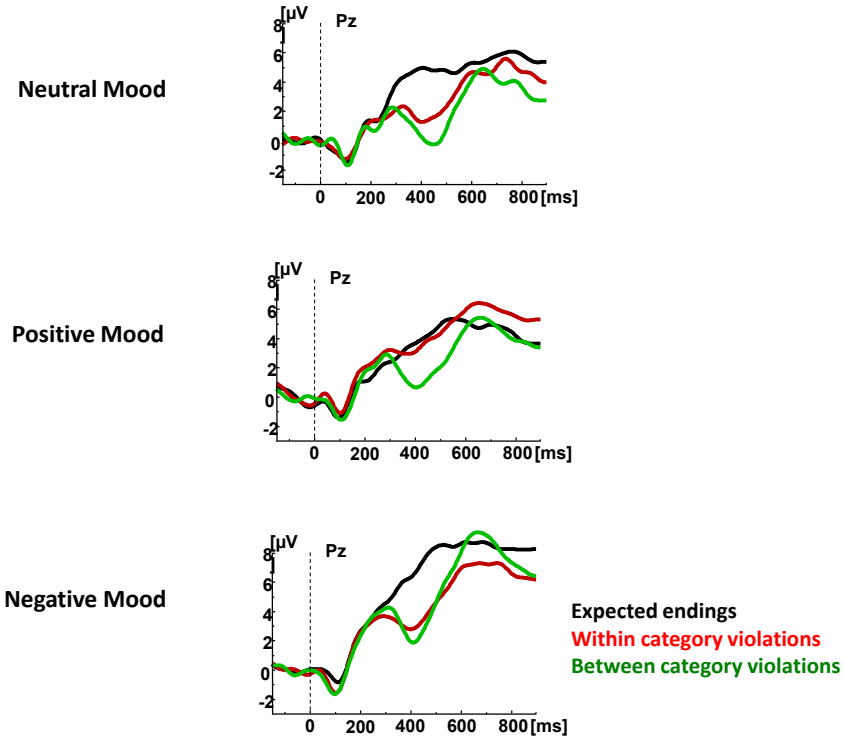

FIGURE 6.

Grand average waveforms to final words in the two-sentence paragraphs that constituted expected endings, within category or between category violations in 15 individuals. Adopted from "Interactions Between Mood and the Structure of Semantic Memory: Event-Related Potentials Evidence," by A. P. Pinheiro, E. Del Re, P. G. Nestor, R. W. McCarley, O. F. Gonçalves, and M. A. Niznikiewicz, 2012, *Social Cognitive and Affective Neuroscience*, *8*, 579-594 (see Footnote 1).

(Pinheiro, Del Re, Nestor, et al., 2012). We used the two-sentence paragraph paradigm described above and presented it to participants under three different moods: neutral, positive, and negative, with the order of presentation counter-balanced across participants. The N400 response to within category violations changed in the mood-specific manner. Under the positive mood, the N400 to within category violations became not significantly different from the N400 to correct targets suggesting that under the positive mood, within category violations were accepted as correct sentence endings (see Figure 6). Conversely, under the negative mood, the prediction mechanisms were altered such that the within category violations were treated similarly to the between category violations, that is, items that shared few semantic features with the target word (see Figure 6). Thus, it appears that mood modulates the interactions between contents of semantic memory and meaning construction. It is not entirely clear how such influence might be exerted but speculations include a role of several neurotransmitters including serotonin, norepinephrin, and dopamine that play a role both in emotion regulation and in the context use (Ruhe, Mason, & Schene, 2007).

SUMMARY

This quite limited overview of research on social communication that includes different aspects of emotion processing from voice and from voice and face, as well as the influence of emotion on a cognitive processing style, exemplifies well how these behaviors are instantiated in the brain. First, it is apparent that the neurophysiological processes involved are relatively fast and span from 50 ms to 450 ms after the stimulus onset. Second, emotional information from voice and face is first tagged and segregated into processing streams at the level of physical features that have ultimately emotional valence assigned. For example, while similar ERP components, and thus, presumably, similar neurophysiological events were associated with the processing of semantic and non-semantic prosody, there were also differences between these two types of processes. The non-semantic affect seems to be processed into neutral and emotional types without the distinction of the type of semantic emotion, while semantic prosody seems to be segregated according to emotion type. As one might expect, the joint processing of face and voice is associated with a more complex set of the ERP components, and thus more complex processes, than processing of the face or the voice alone.

In terms of both methodological and theoretical considerations, it is important to keep in mind that the ERP effects observed between 50 to 300 ms for voice, and voice and face processing, do not represent all processes that are involved in affect processing from voices and faces. For example, it has been demonstrated that orbito-frontal gyrus (OFG) is involved in assigning emotional valence to a voice signal, and that ERPs are not sensitive to activity found in the orbito-frontal gyrus (Paulmann, Seifert, & Kotz, 2009). The activity in the OFG comes relatively late in the processing stream. Similarly, it is apparent that emotion is processed both on cortical and sub-cortical levels and at least some of the subcortical processes are not reflected in the surface ERPs. Finally, as is evident from the work of LeDoux and others, emotion networks, including cortical and sub-cortical regions, are both complex and emotion specific (LeDoux, 2000; Johansen, Wolff, Lüthi, & LeDoux, 2012). The full account of emotional cues processing from both voice and face will have to take into consideration evidence from multiple methodologies and experimental approaches.

The interaction between mood and cognitive processing style is even more complicated. Mood elicitation is associated with a unique set of processes not discussed here. Their result is a shift in mood that, at the brain level, is likely associated with a differential set of specific brain regions' sensitivities to signals mediated by changes in neurotransmitter levels. These brain changes result in cascade reactions where the same semantic information is processed somewhat differently depending on an emotional framework provided by the mood.

This line of research demonstrates how elements of social cognition emerge from sensory data and hints at interactions with higher order processes. Some of the findings in this domain have been highlighted. However, even from this limited review, it is apparent that our understanding of both sensory processes that operate on physical data, and of higher order cognitive processes that modulate them, is still quite limited. It is even less understood what is the nature of the interactions between sensory and higher cognitive processes that ultimately leads to a rich experience of socially relevant events.

In terms of methodology, this review points to the limitations of just one imaging methodology, here ERPs. It is suggested that multimodal assessment of these complex behaviors will bring us closest to their understanding. One technique that is gaining popularity is multimodal fusion of brain imaging data where information from several (at least two) imaging methodologies is used to describe cognitive phenomena under study (Sui, Adali, Yu, Chen, & Calhoun, 2012). It is hoped that the use of multimodal data fusion techniques will not only enhance the sophistication of the tools used to ask important questions but will also contribute to a deeper understanding of how an act of social communication is accomplished across brain regions and brain systems.

Finally, several aspects of social communication are still not very well understood. To name a few, the function of body language and its brain correlates is not well delineated. With the exception of a few studies, relative contributions of eye gaze and facial expressions to convey socially relevant information are also not well described. While not discussed in this review, the role of a theory of mind and empathy in shaping social communication and brain processes that support these processes are also just beginning to be unraveled.

While these questions present challenges, they also present exciting opportunities to better understand the human brain as an organ of social cognition. In so doing we may be able to help these individuals for whom effective social communication is difficult to achieve.

FOOTNOTES

[1] While the figures bear resemblance to already published figures, they were generated from our own data and no part of the published figures was used in their creation.

REFERENCES

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, *9*, 415-430.

Arlinger, S., Lunner, T., Lyxell, B., & Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scandinavian Journal of Psychology*, *50*, 371-384. doi: 10.1111/j.1467-9450.2009.00753.x

Barraclough, N. E, & Perrett, D. I. (2011). From single cells to social perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1739-1752. doi: 10.1098/rstb.2010.0352

Blakemore, S. J. (2012). Development of the social brain in adolescence. *Journal of the Royal Society of Medicine*, *105*, 111-116. doi: 10.1258/jrsm.2011.110221

Boehm, S. G., & Paller, K. A. (2006). Do I know you? Insights into memory for faces from brain potentials. *Clinical EEG and Neuroscience*, *37*, 322-329.

Bolte, A., Goschke, T., & Kuhl, J. (2003). Emotion and intuition: Effects of positive and negative mood on implicit judgments of semantic coherence. *Psychological Science*, *14*, 416-421. doi: 10.1111/1467-9280.01456

Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processiog. *Physics of Life Reviews*, *8*, 383-403. doi: 10.1016/j.plrev.2011.10.002

Chwilla, D. J., Virgillito, D., & Vissers, C. T. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of Cognitive Neuroscience*, *23*, 2400-2414. doi: 10.1162/jocn.2010.21578

Cohen Kadosh, K., Cohen Kadosh, R., Dick, F., & Johnson, M. H. (2011). Developmental changes in effective connectivity in the emerging core face network. *Cerebral Cortex*, *21*, 1389-1394. doi: 10.1093/cercor/bhq215

Cohen Kadosh, K., Johnson, M. H., Henson, R. N., Dick, F., & Blakemore, S. J. (2012). Differential face-network adaptation in children, adolescents, and adults. *NeuroImage*, *69*, 11-20. doi:10.1016/j.neuroimage.11.060

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 19900-19909. doi: 10.1073/pnas.1116727109

Deacon, D., Hewitt, S., Yang, C., & Nagata, M. (2000). Event-related potential indices of semantic priming using masked and unmasked words: Evidence that the N400 does not reflect a post-lexical process. *Cognitive Brain Research*, *9*, 137-146.

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience*, *2*, 1-24. doi: 10.1016/j.dcn.2011.05.003

Dreisbach, G. (2006). How positive affect modulates cognitive control: The costs and benefits of reduced maintenance capability. *Brain and Cognition*, *60*, 11-19. doi: 10.1016/j.bandc.2005.08.003

Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 343-353. doi: 10.37/0278-7393.30.2.343

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, *45*, 15-31.

Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., & Kreifelts, B. (2013). Functional responses and structrural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *NeuroImage*, *76*, 45-56.

Federmeier, K. D., Kirson, D. A., Moreno, E. M., & Kutas, M. (2001). Effects of transient, mild mood states on semantic memory organization and use: An event-related potential investigation in humans. *Neuroscience Letters*, *305*, 149-152. doi: 10.1016/S0304-3940(01)01843-2

Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469-495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*, 373-392.

Fiedler, K. (1988). Emotional mood, cognitive style, and behavioral regulation. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 100-119). Toronto: Hogrefe.

Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 85-98). Mahwah, NJ: Lawrence Erlbaum Associates.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 2213-2223. doi: 10.1098/rstb.2012.0123

Goerlich, K. S., Witteman, J., Schiller, N. O., Van Heuven, V. J., Aleman, A., & Martens, S. (2012). The nature of affective priming in music and speech. *Journal of Cognitive Neuroscience*, *24*, 1725-1741. doi: 10.1162/jocn_a_00213

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*, 47-61.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology*, *47*, 1206-1217. doi: 10.1037/0022-3514.47.6.1206

Isen, A. M., Niedenthal, P., & Cantor, N. (1992). The influence of positive affect on social categorization. *Motivation and Emotion*, *16*, 65-78.

Jacob, H., Brück, C., Domin, M., Lotze, M., & Wildgruber, D. (2013). I can't keep your face and voice out of my head: Neural corre-

lates of an attentional bias toward nonverbal emotional cues. *Cerebral Cortex*. Advance online publication.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage*, *58*, 665-674. doi: 10.1016/j.neuroimage.2011.06.035

Joassin, F., Pesenti, M., Maurage, P., Verreckt, E., Bruyer, R., & Campanella, S. (2011). Cross-modal interactions between human faces and voices involved in person recognition. *Cortex*, *47*, 367-376. doi: 10.1016/j.cortex.2010.03.003

Johansen, J. P., Wolff, S. B., Lüthi, A., & LeDoux, J. E. (2012). Controlling the elements: An optogenetic approach to understanding the neural circuits of fear. *Biological Psychiatry*, *71*, 1053-1060. doi: 10.1016/j.biopsych.2011.10.023

Klasen, M., Chen, Y. H., & Mathiak, K. (2012). Multisensory emotions: Perception, combination, and underlying neural processes. *Reviews in the Neurosciences*, *23*, 381-392. doi: 10.1515/revneuro-2012-0040

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, *1151*, 107-118.

Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: An event-related fMRI study. *NeuroImage*, *37*, 1445-1456.

Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., & Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia*, *47*, 3059-3066. doi: 10.1016/j.neuropsychologia.2009.07.001

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Latinus, M., VanRullen, R., & Taylor, M. J. (2010). Top-down and bottom-up modulation in processing bimodal face/voice stimuli. *BMC Neuroscience*, *11*, 36. doi: 10.1186/1471-2202-11-36

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155-184.

Leitman, D. I., Wolf, D. H., Laukka, P., Ragland, J. D., Valdez, J. N., Turetsky, B. I., et al. (2011). Not pitch perfect: Sensory contributions to affective communication impairment in schizophrenia. *Biological Psychiatry*, *70*, 611-618. doi: 10.1016/j.biopsych.2011.05

Lieberman, M. D. (2012). A geographical history of social cognitive neuroscience. *NeuroImage*, *61*, 432-436. doi: 10.1016/j.neuroimage.2011.12.089

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G, McCarley, R. W., & Niznikiewicz, M. A. (2012). Electrophysiological insights into processing non-verbal emotional vocalizations. *NeuroReport*, *23*, 108-112. doi: 10.1097/WNR.0b013e32834ea757

Liu, T., Pinheiro, A., Zhao, Z., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Emotional cues during simultaneous face and voice processing: Electrophysiological insights. *PLOS ONE*, *7*(2), e31001. doi: 10.1371/journal.pone.0031001

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431-444.

Newman, S. D., Carpenter, P. A., Varma, S., & Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia*, *41*, 1668-1682.

Paulmann, S., & Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, *105*, 59-69. doi: 10.1016/j.bandl.2007.11.005

Paulmann, S., Seifert, S., & Kotz, S. A. (2009). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social Neuroscience*, *5*, 59-75. doi: 10.1080/17470910903135668

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 320. doi: 10.3389/fpsyg.2012.00320

Pinheiro, A. P., Del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., et al. (2012). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: An electrophysiological investigation. *Psychological Medicine*, *43*, 603-618. doi: 10.1017/S003329171200133X

Pinheiro, A. P., Del Re, E., Nestor, P. G., McCarley, R. W., Gonçalves, O. F., & Niznikiewicz, M. (2012). Interactions between mood and the structure of semantic memory: Event-related potentials evidence. *Social Cognitive and Affective Neuroscience*, *8*, 579-594. doi: 10.1093/scan/nss035

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, *32*, 14125-14131. doi: 10.1523/JNEUROSCI.3244-12.2012

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, *9*, 445-458.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, *16*, 189-203.

Rosburg, T., Boutros, N. N., & Ford, J. M. (2008). Reduced auditory evoked potential component N100 in schizophrenia: A critical review. *Psychiatry Research*, *161*, 259-274. doi: 10.1016/j.psychres.2008.03.017

Ruhe, H. G., Mason, N. S., & Schene, A. H. (2007). Mood is indirectly related to serotonin, norepinephrine, and dopamine levels in humans: A meta-analysis of monoamine depletion studies. *Molecular Psychiatry*, *12*, 331-359.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*, 24-30.

Skelly, L. R., & Decety, J. (2012). Passive and motivated perception of emotional faces: Qualitative and quantitative changes in the face processing network. *PLOS ONE*, *7*(6), e40371. doi: 10.1371/journal.pone.0040371

Sui, J., Adali, T., Yu, Q., Chen, J., & Calhoun, V. D. (2012). A review of multivariate methods for multimodal fusion of brain imaging data. *Journal of Neuroscience Methods*, *204*, 68-81. doi: 10.1016/j.jneumeth.2011.10.031

Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2013). ERP evidence on the interaction between information structure and emotional salience of words. *Cognitive, Affective, and Behavioral Neuroscience*, *13*, 297-310. doi: 10.3758/s13415-012-0146-2

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, *156*, 249-268.

RECEIVED 19.03.2013 | ACCEPTED 04.06.2013