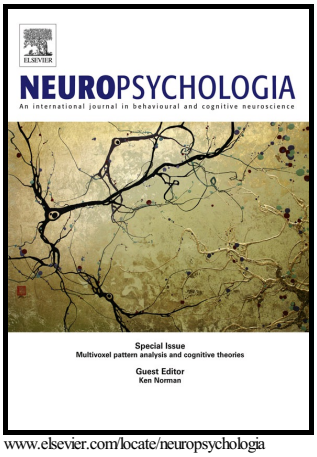
# Coherent emotional perception from body expressions and the voice

Pei-wen Yeh, Elena Geangu, Vincent Reid




www.elsevier.com/locate/neuropsychologia

PII: S0028-3932(16)30282-2
DOI: http://dx.doi.org/10.1016/j.neuropsychologia.2016.07.038
Reference: NSY6091

To appear in: *Neuropsychologia*

Received date: 2 December 2015
Revised date: 7 July 2016
Accepted date: 29 July 2016

Cite this article as: Pei-wen Yeh, Elena Geangu and Vincent Reid, Coheren emotional perception from body expressions and the voice, *Neuropsychologia* http://dx.doi.org/10.1016/j.neuropsychologia.2016.07.038

# Coherent emotional perception from body expressions and the voice

Pei-wen Yeh[1], Elena Geangu[1] and Vincent, Reid[1]

1 Department of Psychology, Lancaster University, UK

Correspondence author:

Peiwen Yeh

Department of Psychology

Lancaster University

Fylde College, Bailrigg

Lancaster, LA1 4YF, UK

Email: p.yeh@lancaster.ac.uk

**Abstract**

Perceiving emotion from multiple modalities enhances the perceptual sensitivity of an individual. This allows more accurate judgments of others' emotional states, which is crucial to appropriate social interactions. It is known that body expressions effectively convey emotional messages, although fewer studies have examined how this information is combined with the auditory cues. The present study used event-related potentials (ERP) to investigate the interaction between emotional body expressions and vocalizations. We also examined emotional congruency between auditory and visual information to determine how preceding visual context influences later auditory processing. Consistent with prior findings, a reduced N1 amplitude was observed in the audiovisual condition compared to an auditory-only condition. While this component was not sensitive to the modality congruency, the P2 was sensitive to the emotionally incompatible audiovisual pairs. Further, the direction of these congruency effects was different in terms of facilitation or suppression based on the preceding contexts. Overall, the results indicate a functionally dissociated mechanism underlying two stages of emotional processing whereby N1 is involved in cross-modal processing, whereas P2 is related to assessing a unifying perceptual content. These data also indicate that emotion integration can be affected by the specific emotion that is presented.



**Keywords:** EEG/ERP, audiovisual processing, body expressions, emotion, congruency, cross-modal prediction

## 1. Introduction

In our daily life, the perception of others' emotions gives us a good insight into their dispositions and allows us to anticipate suitable responses during complex dynamic social interactions. Emotions are typically expressed through different sensory modalities (e.g., faces and bodies, or vocalization). The combination of multiple emotional cues can be particularly useful in making a more accurate and rapid detection and discrimination of emotional content (de Gelder & Vroomen, 2000; Massaro & Egan, 1996; Van den Stock, Righart, & de Gelder, 2007). This is advantageous in life when information from one modality is unclear (Collignon et al., 2008). For instance, the affective prosody in someone's voice can help us disambiguate the emotional expression of their body posture when this is partially occluded in a crowded room. In order to understand how we process emotions, it is essential to elucidate how emotional information from multiple modalities can be unified into a coherent percept. It is for this reason that this study will investigate how auditory and visual information from voices and bodies are jointly processed.

Body postures are often essential visual cues that convey reliable emotional content (for a review, see, e.g., de Gelder, 2006). One such circumstance is when attending to distal events, prior to the ability to see an emotional expression displayed on a face. Thus, body expressions provide important complementary emotional information in our daily life (de Gelder & Beatrice, 2009). Electrophysiological (EEG/ERP) data has provided evidence that the processing of emotional information from the body occurs at an early stage of visual processing at approximately 100 ms (Stekelenburg & de Gelder, 2004; van Heijnsbergen, Meeren, Grezes, & de Gelder, 2007). How our bodily expressions interact with other social cues, such as those from the voice, illustrates a further challenging issue. To date, only a few studies on

emotion perception have focused on the body and the voice (Jessen & Kotz, 2011; Jessen, Obleser, & Kotz, 2012; Van den Stock et al., 2007). Recently, Jessen and colleagues (Jessen & Kotz, 2011; Jessen et al., 2012) used ERPs to examine neural mechanisms underlying the interaction of emotional perceptions from body expressions and affective interjections. This investigation reported a decrease in N1 amplitude in the bimodal condition compared to an auditory-only condition. The auditory N1 is usually reported at around 100 ms after the sound onset and it has shown sensitivity to sensory information such as intensity or frequency (e.g., Naatanen & Picton, 1987; Naatanen et al., 1988). Other multisensory studies (Besle, Fort, Delpuech, & Giard, 2004; Stekelenburg & Vroomen, 2007) also observed the reduction in N1 amplitude to multisensory modalities compared to the sum of the unimodal modalities. If information from each modality was processed independently, the bimodal response is supposed to equal to the sum of unisensory response (AV = A+V). However, if the bimodal response differs from the sum of the unimodal responses in a sub-additive (AV < A+V) or supra-additive manner (AV > A+V), then this points towards interactions occurring between the two modalities (Giard & Peronnet, 1999). As such, the interaction of the auditory and visual information is likely to take place during an early stage of sensory processing. When interpreting how the visual stimuli modulate the auditory processing, van Wassenhove, Grant, and Poeppel (2005) proposed that the preceding visual stimulus acts as a predictor for the forthcoming information. There might be a deactivation mechanism that minimizes the processing of redundant information for multiple modalities, with the consequence that the auditory cortices decrease responses to the relevant information.

Nevertheless, it should be noted that information from multiple modalities is not always presented simultaneously to an observer. The information from one modality

may well precede other modalities within the perceptual system, either in a way of suppression or facilitation (Ho, Schroger, & Kotz, 2014; Takagi, Hiramatsu, Tabei, & Tanaka, 2015). Thus, the preceding one might be a prediction or a constraint to subsequent perceptual processing. However, Jessen's findings (Jessen & Kotz, 2011; Jessen et al., 2012) of the comparison between unimodal and bimodal information does not explore how the preceding visual context influences the processing of different emotions. Irrespective of this, studies on facial expression with voices have exploited the presentation of emotionally conflicted visual and auditory stimuli to reveal the contextual influence on emotional integration. Kokinous, Kotz, Tavano, and Schroger (2014) provided evidence that N1 amplitudes were suppressed in both congruent and incongruent auditory-visual conditions with neutral sounds compared to neutral sound-only conditions. The N1 was only reduced in the congruent pairs with angry sounds when compared to the other two conditions. This emotion-specific suppression in N1 was interpreted in terms of the preceding angry visual stimulus being a stronger predictor compared to the neutral stimulus despite the presentation of incongruent information. In that case, the saliency of emotional contexts compared to non-emotional contexts is preferentially processed during early audiovisual integration.

Another component, the P2 (P200), was also reported in response to congruency and incongruency of audiovisual information (Kokinous et al., 2014). The P2 showing a positive deflection at 200-ms post-stimulus is modulated by the emotional quality of a stimulus (Paulmann, Jessen, & Kotz, 2009). The component is also associated with attention to the competition between multisensory incompatible information (Knowland, Mercure, Karmiloff-Smith, Dick, & Thomas, 2014; Stekelenburg & Vroomen, 2007). More precisely, P2 is correlated to assessing a unifying perceptual

content, dependent upon preceding contexts (van Wassenhove et al., 2005). Ho et al. (2014) have shown that a suppression of the P2 amplitude occurs for a neutral sound presented with an angry face compared to a neutral face. This could be interpreted as an effect of incongruency. However, the P2 amplitude increased when an angry sound was paired with a neutral face than when both sound and face were angry. The P2 implied the modulation of the previous emotional expression on the following neural responses. As such, it has been considered that the P2 is likely to be functionally separated processes to the N1 component during multisensory integration of the emotional percept. While the N1 is associated with visual anticipation for the following auditory processing, the P2 is considered to be content-dependent processing (Stekelenburg & Vroomen, 2007; van Wassenhove et al., 2005).

In addition, there seems to be different cognitive processes from one emotion to another. A reduced N1 latency (Jessen & Kotz, 2011) in response to anger was observed when contrasted with a fearful stimulus either in auditory or in audiovisual conditions. Although the authors did not have a conclusive explanation for this effect, several brain imaging studies provided evidence for common and specific neural circuits during the perception of anger and fear derived from body expressions. For instance, the amygdala and temporal cortices were activated when participants recognized both angry and fearful behaviors compared to neutral (non-emotional) ones (Grezes, Pichon, & de Gelder, 2007; Pichon, de Gelder, & Grezes, 2008, 2009). More specifically, the perception of angry bodies particularly triggered activation within a wider array of the anterior temporal lobes whereas the perception of fearful bodies elicited responses in the right temporoparietal junction (TPJ) (Pichon et al., 2009). Based on these results, it is likely that there are particular neural routes for the

perception of angry and fearful body expressions, respectively, which might modulate the integration of emotion perception information differently.

Moreover, moving stimuli and static stimuli may be processed differently. Generally speaking, dynamic stimuli compared to static stimuli contain explicit movements, which arguably provide more information associated with emotion recognition. Behavioral findings indicate that accuracy rates of emotion recognition for dynamic body expressions are generally higher than for static expressions (Atkinson, Dittrich, Gemmell, & Young, 2004). Supported by fMRI data, responses to emotions were more pronounced when a body was presented with movement than when a still body was shown. For instance, the expression of fear elicited more activation of the TPJ when displayed in a dynamic compared to a static way (Grezes et al., 2007); and the regions of the premotor cortex were more engaged for the dynamic angry body (Pichon et al., 2008). These more pronounced activation areas for dynamic stimuli are linked to the understanding of actions during action observation; therefore, biological motion is likely to be contributing to emotion understanding (Gallese, Keysers, & Rizzolatti, 2004; Iacoboni, 2005).

This is not to say, however, that static body postures are not a reliable source of information for emotion recognition. With static postures of expressions displayed from three angles to different types of emotions, Coulson (2004) revealed that anger and happiness were accurately recognized for large numbers of postures whereas only a small number of postures were perceived for fear and surprise. Atkinson et al. (2004) also found that the classification accuracy for expressions of anger and fear was improved, but for sadness was impeded when increasing exaggeration presentation of moving body expressions. These results are in line with the natural differences in velocity between different emotional body expressions, with sadness featuring less

movement or at times being even motionless, whereas anger is typically associated with a higher velocity movement (Roether, Omlor, Christensen, & Giese, 2009; Volkova, Mohler, Dodds, Tesch, & Bulthoff, 2014). Taken together, each type of emotion is likely to be optimized specifically, whether in a dynamic or static way, in order to be recognized successfully.

The aim of the current study was to investigate the mechanisms underlying the interaction of emotion perceptions presented in body expressions and affective sounds. We examined ERPs in order to compare both the N1 and P2 to emotions (anger vs. fear) and visual stimulus types (dynamic vs. static body expressions) in three conditions: auditory-only, visual-only and audiovisual. We also included emotionally congruent/incongruent body-voice pairs to explore the influence of the preceding visual context to the bimodal interaction. Since emotion processing is thought to be an automatic response (Mauss, Bunge, & Gross, 2007; Mauss, Cook, & Gross, 2007), we conducted the study without directing attention to the emotional characteristics of the stimuli. Based on previous work (Jessen & Kotz, 2011), the N1 is expected to be reduced in amplitude and increased in speed in the audiovisual when compared with the auditory conditions. This differentiation will particularly be observed with the presence of dynamic visual information. It is predicted that the N1 for the emotions of anger and fear will be different either in terms of latency and/or amplitude, and it will also be modulated by the emotional content within the audiovisual information. The P2 is hypothesized to reflect attention on incompatible information and process content of the binding perception; therefore, it is predicted that this will be influenced by emotional audiovisual congruency and visual type (body expression with/without movement)

## 2. Methods

*2.1. Participants*

Twenty-two students from Lancaster University (5 males) with a mean age of 21.5 years old (SD. = 4.0 years) participated in this study. Three participants were excluded from the analysis because of fatigue and one further participant was excluded due to poor signal-to-noise ratio compared to other datasets. All participants had normal vision and hearing, and none reported any neurological or psychiatric disorders. Participants provided written informed consent and were paid (£10) for their participation. The study was approved by Lancaster University Ethics Committee.

*2.2. Stimuli*

All visual stimuli were obtained from the research group of Beatrice de Gelder. To compare the motion effect, there were two types of visual stimuli: a video depicting an actor expressing bodily emotions of anger or fear either with movements (i.e., dynamic condition) or with static postures only (i.e., static condition). The static visual stimuli were based on the results of the Bodily Expressive Action Stimulus Test (de Gelder & Van den Stock, 2011) whereas the dynamic stimuli were extracted from those used by Kret, Pichon, Grezes, and de Gelder (2011). The body expressions for anger included shaking a clenched fist and raising the arm, while fear expressions involved bending the body backwards and defensive movements of the hands. The face area was blurred in all conditions involving the visual modality. The characters were all male dressed in black and performed the body movements against a gray background. The luminance of each video clip was analyzed by taking into account each pixel within a frame (33 frames/clip, 480 × 854-pixel/frame). Each pixel was measured on a gray-scale using MATLAB, with values ranging from 0 to 255. The values of all pixels within a frame were the averaged to obtain a luminance score for

that frame. This allowed us to explore any potential variations in luminance that may appear with time due to the velocity and frequency of motion. Following the procedure described by Jessen and Kotz (2011), we found out that the average luminance of the individual frames in the dynamic stimuli ranges from 64 to 68, with differences of no more than 1 between two consecutive frames. The luminance of the static stimuli was slightly lower than that of the dynamic ones, and varied between 30 to 44.

The auditory stimuli were audio recordings of interjections spoken with a fearful or angry prosody. The sounds were produced by male speakers as included in the Montreal Affective Voices database (Belin, Fillion-Bilodeau, & Gosselin, 2008). All the voices were edited to last 700ms. The mean pitch (anger = 240.47 Hz (SD. = 60.72); fear = 298.45 Hz (SD. = 38.02)) and the mean intensity (anger = 71.66 db (SD. = 9.60); fear = 73.19 db (SD. = 8.88)) were not statistically different between the two emotional sounds.

In the study, the auditory stimuli with or without the visual stimuli were presented in the following conditions: visual-only (V), auditory-only (A), emotionally congruent audio-visual (CAV), and emotionally incongruent audio-visual conditions (IAV). In the V condition, a video clip displayed either a dynamic (dV) or static human (sV) body expressing emotions in the absence of sound. In the A condition, only a sound was played against a black background. The CAV and IAV conditions played affective sounds with either emotionally congruent dynamic (dCAV) or static (sCAV) body expression, or emotionally incongruent ones (dIAV and sIAV, respectively).

In order to account for the emotional properties of the stimuli, we asked two new groups of participants to judge the emotions and rate the intensity of the visual-only

(N = 20) and the audiovisual stimuli (N = 20), respectively. For rating the intensity of the stimuli, we used a 5 point Likert scale ranging from 1 (= very weak) to 5 (= very strong). The **Table 1** shows the mean accuracy and the mean intensity in identifying the emotions, with standard deviation in brackets (D = dynamic body; S= static body)

---------- *Insert Table 1. about here* ----------

*2.3. Procedure*

Participants sat comfortably in a dimly lit/darkened room, and were asked to make their response by pressing a button. Each stimulus was presented using the Psychtoolbox 3.0 in Matlab 2012a. The visual stimuli were presented on a monitor at a distance 90-100 cm from the participants, and the auditory stimuli were binaurally played via two speakers at a sound pressure of 70 dB for all participants. Each trial started with a 800-ms white fixation on a black screen, followed by the presentation of a video clip (CAV, IAV and V condition) or a black background (A condition) for 1300 ms. The auditory stimuli were shown 600ms after the onset of the visual stimulus and ended synchronously with the video clips. In V, CAV and IAV conditions, participants were required to indicate what the person in the video was wearing (e.g., "Did the person wear a jumper/belt?") by pressing the left or the right button. A question mark was also presented in the A condition, and participants also pressed the space bar as a response without any judgement. The question mark disappeared once the participants had made their response. Each block included 64 trials. In order to avoid learning the regularities of question marks presentation, in each block we randomly showed them after a trial in less than 60% of the cases (ranging from 20 to 33), by using a custom Matlab script. The presentation of a question mark after a trial was presented less than 5 consecutive times. The testing started after a practice session consisting of 10 trials, and the participants were able to take a self-defined

break between blocks if required. The study consisted of 8 blocks, a total of 512 trials. In each of the 4 blocks, either the dynamic or static body expression (V) was presented (8 times/block) together with other factors of *condition* (A, CAV, IAV conditions) and *emotion* (anger and fear). The study lasted approximately 50 minutes, including breaks.

*2.4. EEG recording and analysis*

The data were recorded by EGI NetStation system (Geodesic Sensor Nets, Inc., Eugene, OR) with a 128-channel electrode net. The EEG signal was sampled at 500 Hz and the impedances were kept to 50 Hz or less during recording. All electrodes were on-line referenced to vertex (Cz). For computing the ERPs, the data was filtered with a 0.3-30 Hz bandpass filter and segmented off-line from 100 ms before to 700ms after sound onset. Baseline correction was applied to 100 ms prior to each segment before artifact rejections. Trials were rejected with EGI software once the eye movement exceeded +/- 140 uV, and eye blinks exceeded +/- 100 uV. Any channels that exceeded over +/- 200 uV for an electrode were marked as bad. If more than 12 electrodes within a trial were marked as bad, the trial was automatically discarded. The remaining trials were re-referenced into an average reference before averaged waveforms for each participant with each condition. The analysis was focused on the two ERP components, N1 and P2, which have been indexed in audiovisual emotion perception literature. Based on previous studies (e.g. Jessen & Kotz, 2011), and visual inspection of present data, two different analyses were conducted: the first involved the latency to the peak amplitude between 90-180 ms (N1) and 160-330 ms (P2) after sound onset, and the second involved the mean peak amplitude for the time window centered on the latency of each conditions (+/- 30 ms).

As the distribution between frontal-central and central-parietal sites showed a

reversed polarity of the potentials, the statistical analysis were therefore performed individually, taking the average of these electrode clusters for frontal (6, 11, 19, 4, 12, 5), central (Ref, 7, 106, 80, 31, 55) and central-parietal (62, 61, 78, 79, 54) regions of interest (ROI) (**Figure 1**). A 2 (*visual type:* dynamic, static body expression) x 4 (*conditions*: audio-only, visual-only, emotionally congruent audiovisual, and emotionally incongruent audiovisual) x 2 (*emotion*: anger, fear) x 3 (*ROI*: frontal-central, central, central-parietal sites) repeated-measures ANOVA was conducted on the two time windows. Post-hoc analyses (least significant difference) were run where any significant ($p$-value < 0.05) interaction effects were reported.

*- Insert Figure1.here (single column) -*

## 3. Results

The topography and the grand average of the N1 and the P2 at sequential time from 100 to 350 ms for each condition are presented separately for the dynamic (**Figure 2**) and static (**Figure 3**) visual stimuli. In the following sections, we only reported the key findings, particularly the comparison of condition for visual types (dynamic and static) and for emotional content (anger and fear) as we were interested in modality and congruency effects. A full list of all statistical comparisons can be found in the **Table 2**

------ *Insert Figure 2. here (double column)*-----

------ *Insert Figure 3. here (double column)*-----

----- *Insert Table 2. here* ------

### *3.1. ERP latency*

#### *3.1.1. N1*

Only the main effect of *emotion* ($F(1,17) = 62.65$, $p < .0001$, $\eta^2 = .787$) reached

significance. A significant interaction between *emotion*, *condition* and *site* (F(6, 102) = 2.74, $p$ =.017, $\eta^2$ = .139) (**Table 3**) was also found. *Post hoc* analysis of the interaction indicated that the N1 response to the angry stimuli peaked earlier than to the fearful stimuli, and the difference was most enhanced in both A and CAV conditions at central and central-parietal sites (all $p < .0001$).

In addition to emotion effects, we also considered the comparison of the conditions. However, no significant effects were found when the three-way interactions (*emotion, condition and site*) were unpacked by the other two factors. The *condition* only showed a significant two-way interaction with *emotion* (F(3, 51) = 2.67, $p$ =.029, $\eta^2$ = .151). Further analysis showed a shorter N1 latency was found for the angry stimulus in the CAV than in IAV condition ($p$ = .022), whereas the latency was only reduced in the IAV compared to the A condition ($p$ = .031) for the fearful stimulus.

------ *Insert Table 3. about here* -----

*3.1.2. P2*

Only the main effect of *emotion* was significant (F(1,17) = 9.47, $p$ = .007, $\eta^2$ = .358), revealing a rapid latency to the P2 peak for the angry compared to the fearful stimuli. The *emotion* also showed significantly interactions with *type* and *condition* (F(3,51) = 4.12, $p$ = .011, $\eta^2$ = .195) (**Table 4**). Further analysis showed the different latencies between emotion were pronounced in the sounds-only condition (dA and sA: all $p < .0001$), and sounds with dynamic visual information (dCAV: $p$ = .031; dIAV: $p < .0001$). However, the emotion effects were reduced when sounds were presented with static body expressions ( sCAV: $p$ = .042; sIAV: $p$ = .024).

With regard to the condition effects, we only found the difference when the static

body expressions were presented. Shorter latencies to angry sounds were observed in both sA and sCAV compared to sIAV conditions ($p$ = .01; $p$ < .0001, respectively). The peak was shorter for sCAV than for sIAV conditions when sounds were fearful ($p$ = .049).

------ *Insert Table 4. about here* -----

*3.2. ERP amplitude*

**Figure 4** shows the mean peak amplitude of the N1 (top) and of the P2 (bottom) components across emotions, visual type, conditions.

---- *Insert Figure 4. here (single column)* -----

*3.2.1. N1*

A significant main effect of *condition* was found ($F(3,51) = 15.98$, $p < .0001$, $\eta^2 = .485$), with reduced N1 amplitudes in both CAV and IAV conditions compared to the A condition ($p$ = .015 and $p$ = .018, respectively). Of interest is the marginally significant four-way interactions between *condition*, *emotion*, *visual types* and *sites* ($F(6,102) = 1.87$, $p$ = .093, $\eta^2$ = .099). When separated by *visual types*, *emotion*, and *sites*, smaller N1 amplitudes were observed for angry dCAV and dIAV conditions compared to the dA condition at frontal ($p$ = .005; $p$ = .001, respectively) and central sites ($p$ = .014; $p$ = .041, respectively). Conversely, no significant differences were found between conditions with static body expressions (sCAV vs. sA, $p$ = .375; sIAV vs. sA, $p$ = .282). In response to the fearful sounds, a reduced N1 amplitude for dCAV was found when contrasted with dA at central regions ($p$ = .036). However, the reduced N1 was less significant for sCAV and sIAV compared to sA conditions at frontal sites ($p$ = .018; $p$ = 0.088, respectively).

*3.2.2. P2*

We observed a significant main effect of *condition* ($F(3,51) = 38.42$, $p < .0001$, $\eta^2 = .693$). The *post hoc* analysis indicated that smaller P2 amplitudes were observed for IAV in comparison to A conditions ($p = .017$), but the reduction become less significant in CAV compared to A condition ($p = .071$). In addition, *type* ($F(1,17) = 11.55$, $p = .003$, $\eta^2 = .405$) as well as *emotion* showed significant main effects ($F(1,17) = 4.79$, $p = .043$, $\eta^2 = .220$). Planned comparison revealed a reduced P2 for dynamic compared to static visual stimuli, as well as for angry than for fearful expressions. Significant interactions between *type, condition*, *emotion* and *site* were also found ($F(6,102) = 2.45$, $p = .030$, $\eta^2 = .13$). Further analysis was separated by *visual types*, *emotion*, and *sites.* Generally, the differences between conditions were more pronounced with the presentation of dynamic contrasted with static body expressions. With presentation of the dynamic angry body, smaller P2 amplitudes were found in both dA and dCAV conditions compared to the dIAV condition at the frontal regions ($p = .011$; $p = .001$, respectively). In addition, smaller responses to dCAV compared to dA conditions nearly achieved significance at central sites ($p = .085$) but became robust at central-parietal sites ($p =. 013$). However, no significant differences were found in response to angry stimuli when the body expressions were static. In contrast, reduced P2 amplitudes were found for the fearful IAV condition compared to dA at frontal and central sites ($p = .033$, $p = .005$, respectively), and for the dIAV compared to dCAV conditions at frontal sites ($p = .004$). When static body expressions were presented, only larger P2 amplitudes were observed for the sCAV compared to sA condition at frontal regions ($p = .003$).

## 4. Discussion

In the current study, we used ERPs to measure the integration of emotion perception from body expressions and affective interjections. Both the emotion and

the presence of dynamic visual information significantly modulated both the N1 and P2 components. However, the modality in which the emotional information was presented significantly affected the N1, whereas the effect of the congruency between visual and auditory information was only observed within the P2. These findings indicate that processing the interaction between visual information, related to body posture, and auditory information, specific to prosody, during emotion perception may occur at different stages, as reflected by the response of the N1 and P2 components. The influences of modality, visual type, emotion and audiovisual congruency within these two components will be discussed in more detail below.

### *4.1. Modality Effects*

In agreement with the studies of Jessen and her colleagues (e.g. Jessen & Kotz, 2011; Jessen et al., 2012) on emotional integration from body postures and prosody, we found both reduced N1 latencies and amplitudes for the emotionally congruent and incongruent audiovisual compared to voice-only conditions. This observation is also consistent with other previous investigations of audio-visual integration outside the emotional domain (e.g. Stekelenburg & Vroomen, 2007), suggesting that the interaction of body posture and voice information occurs at a very early stage of perception. In addition, these modality effects in both amplitude and latency are likely to be activated by unspecific emotional information as the N1 was suppressed in both angry and fearful contexts.

### *4.2. Comparison between Emotions*

The reduced N1 latency for anger was robustly found when compared with fearful stimuli in the auditory-only and audio-visual conditions. Since the N1 is interpreted as a sensory component, it shows that faster processing for anger than for fear at a very early stage, rather than a later stage of processing (Paulmann et al.,

2009). With regard to the emotion component, both anger and fear are associated with high arousal and negative valence, yet they convey quite different social signals. In comparison to fear, anger often displays cues about the expressers' intentions to act, so it is an interactive message that requires observers to modify their behaviour in tune with the approaching interaction (Pichon et al., 2009). Neuroimaging studies also have demonstrated that the perceptions of the two emotions are different. For instance, the premotor area and temporal lobe, activate more when one perceives an angry rather than a fearful body (Pichon et al., 2009). The authors proposed that the function of the premotor area is to readjust our defensive behaviour in response to one monitoring a forthcoming threat, and the temporal area evaluates the emotional contexts by drawing from past experience. Consequently, this additional activation is crucial for one to be sensitive to the detection of anger, improving their social relationships.

However, the current differences between emotions could be the fact that the fearful stimuli we used do not evoke threat in the observer as efficiently as the angry stimuli. It has been indicated that if the expresser's signals of emotions are directed to or successfully shared with the observer, these might become threats to the observer which require an adjustment in their behaviour (Adams & Kleck, 2003). Therefore, whether the emotional signals are clearly related to the observers is likely to influence the observer's emotion perception.

*4.3. Congruency Effect*

The differentiation between response to congruent and incongruent audiovisual conditions was not reflected by the N1, which is consistent with the findings of Stekelenburg and Vroomen (2007), but is in contrast to two other prior studies (Ho et al., 2014; Kokinous et al., 2014). Several reasons might influence the results. Firstly,

Ho et al. (2014) and Kokinous et al. (2014) observed facial expressions whereas we presented body expressions as visual information. It is possible that perceiving emotions from bodily expressions is less sensitive than facial expressions at an early stage of processing; that is, faces compared to bodies appear to be better predictors of the auditory emotional information. Secondly, different combinations of emotions may modulate the congruency effects differently. Previous studies examined the audiovisual congruency effect by mismatching angry and neutral information, which is different from the present study, which paired the expressions of anger and fear. Both anger and fear are negative emotions conveying a message of threat, so it might be difficult to perceive the difference when the two emotions are displayed through separate modalities simultaneously. Differences could also arise due to distinctive methodology in analysis and instruction, and so further studies are required to demonstrate this assumption.

Although a congruency effect at the level of N1 was absent in our study, the same component reflected predominance of the information from bimodality than from unimodality. Conversely, only a significant congruency effect was observed for the P2 amplitude at frontal-central regions, which was specific to moving body expressions. These results suggest that the processing for the modal interaction emerges at an early sensory stage, but for conjunctions of emotional contents occurs at a later stage (Kokinous et al., 2014). The discrepancy within the two investigated components is also in line with the assumption that the AV effect on the N1 is modulated by visual anticipation but is independent of audiovisual coherence, whereas the P2 is driven by AV coherence and more dependent on specific contents (Ganesh, Berthommier, Vilain, Sato, & Schwartz, 2014).

More precisely, the direction for the congruency effects, either the suppression or

facilitation within P2, might depend on the preceding emotional contexts (Ho et al., 2014). The current data has shown that the P2 amplitude reduced when the angry body presentation preceded the fearful sounds compared with when the fearful sounds were paired with fearful body expressions. Conversely, the P2 amplitude increased when the fearful body preceded the angry sounds compared with when both the sounds and body expressions presented anger. The preceding angry body expression may be considered to convey a strong signal and lead to a greater expectation by the participant. This is strongly in conflict with the following fearful sounds, leading to reassessing the stimulus with consequent processing costs and attention (Crowley & Colrain, 2004). However, the fearful image seems not to carry this message as strongly as the anger stimuli; therefore, the P2 was not suppressed to the incongruent combination with the angry sounds.

An alternative perspective for the reversed congruency effects may be related to the different dominance within separate modalities for the two emotions. The modality dominance might be different for each type of emotion when presenting audiovisual information (Takagi et al., 2015) as voice dominance was shown for fear, whereas anger was most linked to a visual modality. Considering the auditory-only condition as a baseline, we observed that the amplitude of the P2 was reduced whenever the angry body expression was displayed before the voices, whereas no significant effects appeared within the P2 when a fearful body was presented. In that case, the image of an angry body might serve as a very strong predictor, modulating the brain responses irrespective of the information provided subsequently by the emotional voices. However, this modality dominance may not be due to differences in intensity. While the auditory stimuli were rated as being less intense than the visual ones, this difference was consistent across the two emotions. Despite this, we still

observed different directions of congruency effects, either facilitate or suppress response to congruent compared to incongruent pairs. As such, we consider the natural characteristic of emotion stimuli or other factors to be more associated the findings than an intensity explanation, which is not parsimonious with the data.

It has also previously been suggested that the P2 could represent a general stimulus classification process (Garcia-Larrea, Lukaszewicz, & Mauguiere, 1992), and that the mismatched audiovisual pairs might yield new percepts. A noticeable example is the McGurk effect (McGurk & MacDonald, 1976), which comprises a speech sound (/ba/) overlaid with a face articulating another sound (/ga/) resulting in a fused percept (/da/), whereas a reverse combination of the auditory (/ga/) and visual (/ba/)is perceived as /bga/. In that case, the combined information is likely to be perceived differently when the emotional information was reversed from the two modalities. Based on this assumption, in our study, the perception for the four types of combinations from two modalities during the processing of two emotions might be different, with consequent results found within the P2 in terms of latency or amplitude.

*4.4. The Modulation of Motion*

The current study showed that visual types of emotional body expressions are relevant for multisensory emotion processing. In particular, both modality and congruency effects were observed within N1 and P2, respectively when presenting dynamic materials; however, the results were not entirely extended to the static stimuli, especially for the angry stimulus. In support of the assumption that the kinetic cues from visual information fasten the process for the following auditory information (Stekelenburg & Vroomen, 2007). Some neuroimaging studies have provided evidence that specific brain areas are activated when one perceives a dynamic

represented body compared to a static one (e.g. Pichon et al., 2008). The additional engagement of brain areas, such as the premotor cortex, are noted for the perception of biological motion, but have also been observed for the processing of understanding emotion (e.g. Iacoboni, 2005). Consequently, our results feed into a literature that indicate that viewing a dynamic angry body activates sensory regions as well as motor areas, which helps one to understand the emotion that is being portrayed.

On the other hand, the benefit of dynamic cues appears to partially apply to the recognition of fear. Although a larger activation of the premotor area has also been reported for a fearful body in dynamic compared to still states (Grezes et al., 2007), our data nonetheless indicated N1 suppression for the fearful audiovisual condition in static conditions. The more easily recognized fearful stimuli might be related to the angles within the postures for the current static body stimuli, whereas this may not be the case for anger (Coulson, 2004). In addition, a fearful body is often well recognized with fewer high velocity movements than angry expressions (Roether et al., 2009). On this basis, we have assumed that a still body presentation is sufficient for the discrimination of fear.

*4.5. Limitations*

There are some potential limitations related to the present study. First, attention might not be balanced across the two modalities. We tried to divert participants' attention from the emotional information by asking them to make judgments about the non-emotional visual properties of the stimuli. However, this may not have fully removed attention from visual information, and attention away from auditory processes. Also, the auditory condition cannot be displayed in a dynamic and static way. We consequently presented twice the A compared to the other CAV, IAV and V conditions. To ensure the effects of the auditory-only were not attenuated, we

examined the response in the blocks with each visual type and found no differences. Another limitation might be due to the fact that the emotional intensity of the fearful stimuli was higher for the dynamic than for the static presentations. However, the modality effect can be observed for the static fearful body expression but not for the angry static expression with the same intensity, which suggests that the emotional intensity and the biological motion per se are not the main contributors to the observed effects. Other factors might contribute more specifically to the perceptual integration of fear. Moreover, males and females are known to differ in processing emotional prosody (e.g., Schirmer & Kotz, 2003; Schirmer, Kotz, & Friederici, 2002) and this might be an important aspect to consider when investigating emotional information processing. As such, the lack of balance with respect to gender in the present study is a restriction for generalizing the findings to broader populations. However, the present study is more focused on understanding whether different types of emotions and body exhibitions influence the emotion perception from body expression and sounds. Given the number of variables in the current study, which currently features four factors (visual types, condition, emotions, and sites), the addition of another factor would dramatically increase the complexity of the study and the associated interpretations of results. In this case, we reasoned it is better to address gender issues across body expression and voice in future studies.



### *4.6. Conclusion*

The present study reiterates the findings of Jessen and Kotz (2011) indicating a clear suppression of the N1 amplitude and latency for the emotionally congruent and incongruent audiovisual conditions than for auditory-only condition. Moreover, we have clearly shown that the availability of dynamic information about body expressions aids emotion processing, particularly at later stages as indexed by the P2.

The N1 and the P2 were separately influenced by the presence of multimodal emotional information and their congruency, leading us to conclude that these components index different emotion processing functions. The current evidence supports the previous assumption that the N1 is affected by multisensory signals in a manner that is independent of congruency information, whereas the P2 is sensitive to the coherence of the integration of emotional content.

**References**

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychol Sci, 14*(6), 644-647.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746.

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behav Res Methods, 40*(2), 531-539.

Besle, J., Fort, A., Delpuech, C., & Giard, M. H. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur J Neurosci, 20*(8), 2225-2234. doi: 10.1111/j.1460-9568.2004.03670.x

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Res, 1242*, 126-135. doi: 10.1016/j.brainres.2008.04.023

Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior, 28*(2), 117-139. doi: Doi 10.1023/B:Jonb.0000023655.25550.Be

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology, 115*(4), 732-744. doi: DOI 10.1016/j.clinph.2003.11.021

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat Rev Neurosci, 7*(3), 242-249. doi: 10.1038/nrn1872

de Gelder, B., & Van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST). Construction and Validation of a Stimulus Basis for Measuring Perception of Whole Body Expression of Emotions. *Front Psychol, 2*, 181. doi: 10.3389/fpsyg.2011.00181

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye.

*Cognition & Emotion, 14*(3), 289-311.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn Sci, 8*(9), 396-403. doi: 10.1016/j.tics.2004.07.002

Ganesh, A. C., Berthommier, F., Vilain, C., Sato, M., & Schwartz, J. L. (2014). A possible neurophysiological correlate of audiovisual binding and unbinding in speech perception. *Front Psychol, 5*, 1340. doi: 10.3389/fpsyg.2014.01340

Garcia-Larrea, L., Lukaszewicz, A. C., & Mauguiere, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia, 30*(8), 723-741.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci, 11*(5), 473-490.

Grezes, J., Pichon, S., & de Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage, 35*(2), 959-967. doi: DOI 10.1016/j.neuroimage.2006.11.030

Ho, H. T., Schroger, E., & Kotz, S. A. (2014). Selective Attention Modulates Early Human Evoked Potentials during Emotional Face-Voice Processing. *J Cogn Neurosci*, 1-21. doi: 10.1162/jocn_a_00734

Iacoboni, M. (2005). Neural mechanisms of imitation. *Curr Opin Neurobiol, 15*(6), 632-637. doi: 10.1016/j.conb.2005.10.010

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage, 58*(2), 665-674. doi: DOI 10.1016/j.neuroimage.2011.06.035

Jessen, S., Obleser, J., & Kotz, S. A. (2012). How Bodies and Voices Interact in Early Emotion Perception. *Plos One, 7*(4). doi: ARTN e36070

DOI 10.1371/journal.pone.0036070

Knowland, V. C., Mercure, E., Karmiloff-Smith, A., Dick, F., & Thomas, M. S. (2014). Audio-visual speech perception: a developmental ERP investigation. *Dev Sci, 17*(1), 110-124. doi: 10.1111/desc.12098

Kokinous, J., Kotz, S. A., Tavano, A., & Schroger, E. (2014). The role of emotion in dynamic audiovisual integration of faces and voices. *Soc Cogn Affect Neurosci*. doi: 10.1093/scan/nsu105

Kret, M. E., Pichon, S., Grezes, J., & de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage, 54*(2), 1755-1762. doi: 10.1016/j.neuroimage.2010.08.012

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychon Bull Rev, 3*(2), 215-221. doi: 10.3758/BF03212421

Mauss, I. B., Bunge, S. A., & Gross, J. J. (2007). Automatic Emotion Regulation.

*Personality and Social Psychology Review,, 8*, 220-247.

Mauss, I. B., Cook, C. L., & Gross, J. J. (2007). Automatic emotion regulation during anger provocation. *Journal of Experimental Social Psychology, 43*(5), 698-711. doi: DOI 10.1016/j.jesp.2006.07.003

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Naatanen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound - a Review and an Analysis of the Component Structure. *Psychophysiology, 24*(4), 375-425. doi: DOI 10.1111/j.1469-8986.1987.tb00311.x

Naatanen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., & Sokolov, E. N. (1988). Frequency and Location Specificity of the Human Vertex N1-Wave. *Electroencephalography and Clinical Neurophysiology, 69*(6), 523-531. doi: Doi 10.1016/0013-4694(88)90164-2

Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the Multimodal Nature of Human Communication Insights from ERPs. *Journal of Psychophysiology, 23*(2), 63-76. doi: Doi 10.1027/0269-8803.23.2.63

Pichon, S., de Gelder, B., & Grezes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social Neuroscience, 3*(3-4), 199-212. doi: Doi 10.1080/17470910701394368

Pichon, S., de Gelder, B., & Grezes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *Neuroimage, 47*(4), 1873-1883. doi: 10.1016/j.neuroimage.2009.03.084

Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *J Vis, 9*(6), 15 11-32. doi: 10.1167/9.6.15

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *J Cogn Neurosci, 15*(8), 1135-1148. doi: 10.1162/089892903322598102

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Res Cogn Brain Res, 14*(2), 228-233.

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport, 15*(5), 777-780. doi: DOI 10.1097/01.wnr.0000119730.93564.e8

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci, 19*(12), 1964-1973. doi: 10.1162/jocn.2007.19.12.1964

Takagi, S., Hiramatsu, S., Tabei, K., & Tanaka, A. (2015). Multisensory perception of the six basic emotions is modulated by attentional instruction and unattended modality. *Front Integr Neurosci, 9*, 1. doi: 10.3389/fnint.2015.00001

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7*(3), 487-494. doi: 10.1037/1528-3542.7.3.487

van Heijnsbergen, C. C. R. J., Meeren, H. K. M., Grezes, J., & de Gelder, B. (2007). Rapid detection of fear in body expressions, an ERP study. *Brain Research, 1186*, 233-241. doi: DOI 10.1016/j.brainres.2007.09.093

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci U S A, 102*(4), 1181-1186. doi: 10.1073/pnas.0408949102

Volkova, E. P., Mohler, B. J., Dodds, T. J., Tesch, J., & Bulthoff, H. H. (2014). Emotion categorization of body expressions in narrative scenarios. *Front Psychol, 5*, 623. doi: 10.3389/fpsyg.2014.00623

**Figure Legends**

Figure 1. Averages were calculated based on electrode ROIs for frontal (6, 11, 19, 4, 12, 5), central (Ref, 7, 106, 80, 31, 55) and central-parietal (62, 61, 78, 79, 54) channels.

Figure 2. The ERPs displaying (A) the topography distributions for angry (4 left) and fearful (4 right) information in dA, dCAV, dIAV and dV conditions from 100 to 350 ms after onset auditory stimulus when the dynamic body expressions were presented. (B) The grand average for each condition at central electrode sites.

Figure 3. The ERPs displaying (A) the topography distributions for angry (4 left) and fearful (4 right) information in sA, sCAV, sIAV and sV conditions from 100 to 350 ms after onset auditory stimulus when the static body expressions were presented. (B) The grand average for each condition at central electrode sites.

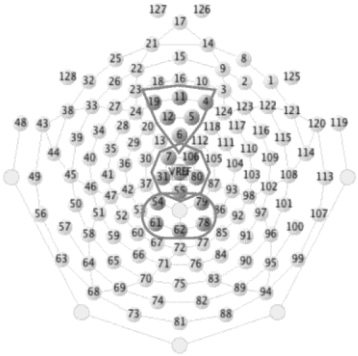
Figure 4. The N1 and P2 mean peak amplitudes for each factor (condition, visual types and emotions), which is indicative of effects in the region.

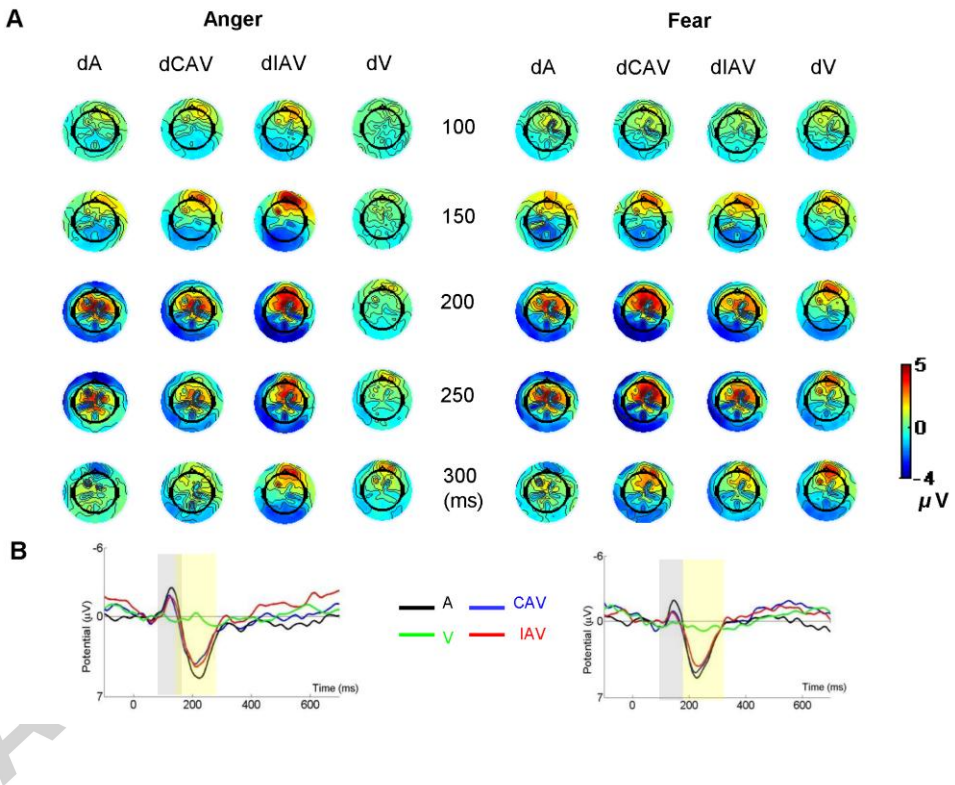


**Highlights:**

- **N1 and P2 reflect distinct processes for the emotionally perceptual integration.**
- **The emotional congruency effect was modulated by the preceding visual context.**
- **Emotion integration can be affected by the specific emotion that is**

**presented.**




A
Anger
Fear
dA
dCAV
dIAV
dV
100
150
200
250
300
(ms)
5
0
-4
μV
B
Potential (uV)
Time (ms)
A
CAV
V
IAV

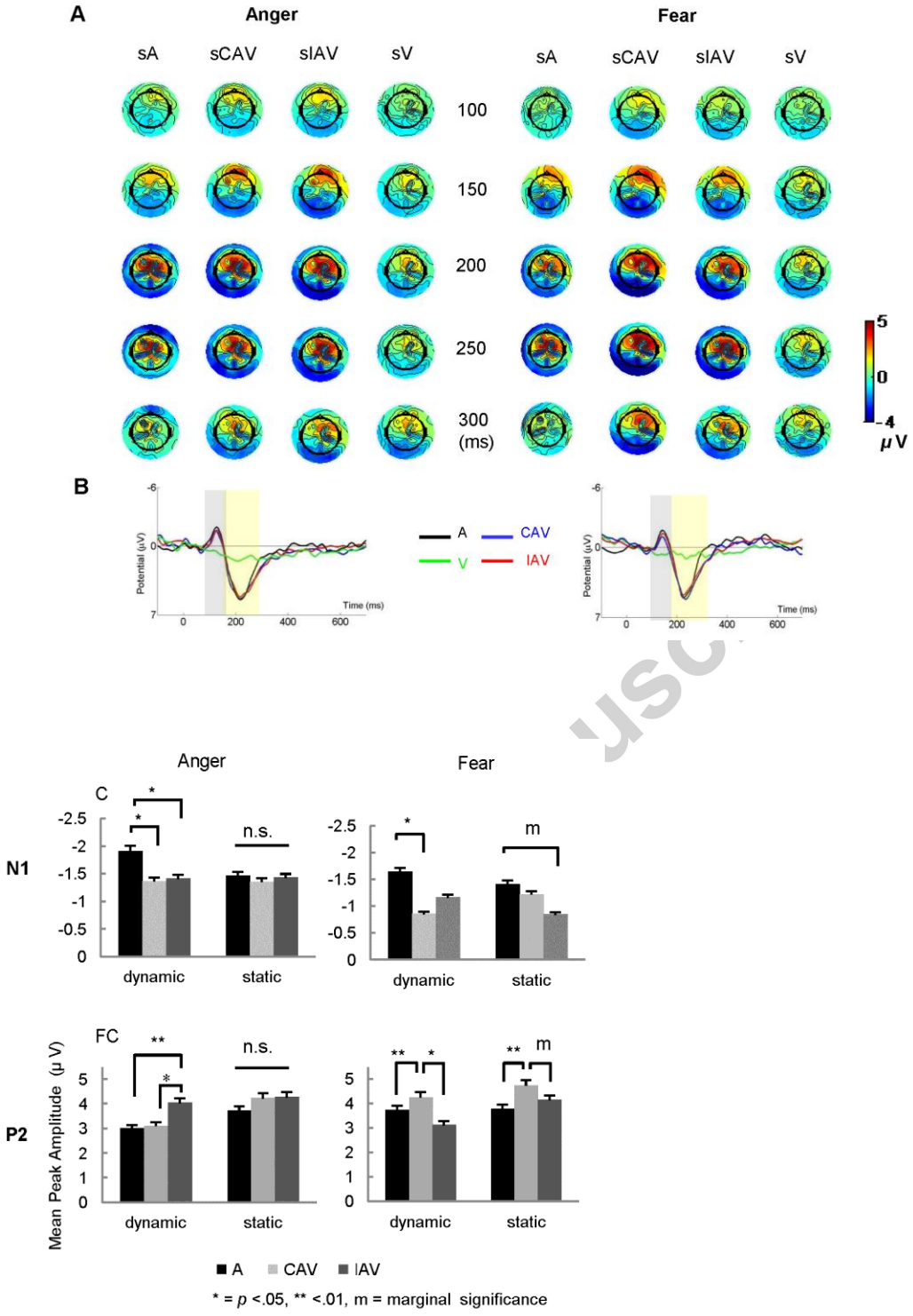

A
Anger
Fear
sA
sCAV
sIAV
sV
100
150
200
250
300
(ms)
5
0
-4
μV
B
Potential (μV)
Time (ms)
A
CAV
V
IAV
C
FC
N1
P2
Anger
Fear
dynamic
static
n.s.
m
Mean Peak Amplitude (μV)
A
CAV
IAV
* = p <.05, ** <.01, m = marginal significance

**Table 1.**



**Table 1**. Results of rating for the stimuli presented in the EEG study. Mean accuracies (%) and intensity (1 to 5 scale) for emotions of angry and fear in V (visual-only condition) and A (auditory-only condition), with standard deviant in parentheses.

| | V | | | | A | |
|---|---|---|---|---|---|---|
| | Anger | | Fear | | Anger | Fear |
| | D | S | D | S | | |
| Accuracy | 95.24% (0.15) | 100% (0) | 100% (0) | 97.62% (0.11) | 100 % (0) | 92.5% (0.13) |
| Intensity | 3.60 (0.90) | 3.35 (0.80) | 4.50 (0.51) | 3.35 (0.49) | 2.88 (0.79) | 2.52 (0.73) |

D = dynamic visual stimulus; S = static visual stimulus

**Table 2.**



**Table 2**. A summary of statistical analysis * $p < .05$; ** < .01; *** < .001

| | | | N1 Amp | | P2 Amp | | N1 latency | | P2 latency | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | df | F | *p* | F | *p* | F | *p* | F | *p* |
| Visual type | | 1,17 | 1.61 | | 11.55 | ** | 0.62 | | 0.11 | |
| Condition | | 3,51 | 15.98 | *** | 38.42 | *** | 0.83 | | 0.66 | |
| | CAV v.s A | | | * | | .071 | | | | |
| | CAV v.s V | | | *** | | *** | | | | |
| | CAV v.s IAV | | | | | | | | | |
| | IAV v.s A | | | * | | * | | | | |
| | IAV v.s V | | | *** | | *** | | * | | |
| | A v.s V | | | *** | | *** | | | | |
| Emotion | | 1,17 | 1.43 | | 4.79 | * | 62.65 | *** | 9.47 | ** |
| Site | | 2,34 | 50.80 | *** | 56.75 | *** | 14.21 | *** | 12.43 | *** |
| | | | | | | | | | | |
| Type*condition | | 3,51 | 1.71 | | 0.81 | | 0.88 | | 3.67 | * |
| Type* emotion | | 1,17 | 3.18 | .092 | 5.45 | * | 0.60 | | 5.77 | * |
| Condition *emotion | | 3,51 | 0.37 | | 0.26 | | 3.26 | * | 3.41 | * |
| Type*condition * emotion | | 3,51 | 0.86 | | 1.45 | | 0.44 | | 4.12 | * |
| Type *site | | 2,34 | 3,28 | .050 | 6.28 | ** | 0.10 | | 1.41 | |
| Condition *site | | 6,102 | 9.74 | *** | 9.09 | *** | 2.19 | * | 13.94 | *** |
| Type*condition * site | | 6,102 | 0.95 | | 1.03 | | 1.01 | | 0.92 | |
| Emotion *site | | 2,34 | 5.09 | * | 15.19 | *** | 33.49 | *** | 0.57 | |
| Type*emotion *site | | 2,34 | 0.56 | | 0.23 | | 1.11 | | 1.31 | |
| Condition *emotion*site | | 6,102 | 5.35 | *** | 10.36 | *** | 2.74 | * | 0.45 | |
| Type*condition *emotion*site | | 6,102 | 1.87 | .093 | 2.45 | * | 1.33 | | 1.24 | |

**Table 3.**



**Table 3**. The mean in milliseconds of N1 peak latency for each condition at frontal (F), central (C) and central-parietal (CP) sites (SD in parentheses)

| | | N1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | anger | | | fear | | |
| | | F | C | CP | F | C | CP |
| Dynamic visual type | A | 120.4 (22.54) | 126.0 (12.34) | 128.0 (14.26) | 130.2 (27.72) | 147.3 (15.44) | 151.7 (10.81) |
| | CAV | 113.2 (19.11) | 122.1 (12.19) | 126.2 (12.66) | 116.30 (24.00) | 128.2 (20.24) | 155.6 (11.38) |
| | IAV | 113.2 (22.00) | 124.0 (15.46) | 144.7 (15.69) | 125.1 (22.83) | 138.7 (20.21) | 148.0 (16.57) |
| | V | 137.7 (24.39) | 125.1 (23.88) | 127.5 (26.63) | 120.8 (25.52) | 142.0 (24.49) | 144.7 (20.34) |
| Static visual type | A | 117.4 (25.87) | 127.6 (14.38) | 123.0 (13.21) | 124.0 (30.49) | 140.60 (13.88) | 153.6 (14.70) |
| | CAV | 116.2 (21.74) | 123.7 (14.37) | 129.7 (13.21) | 124.0 (30.49) | 140.6 (13.88) | 153.6 (14.69) |
| | IAV | 115.4 (21.84) | 123.2 (15.83) | 134.4 (14.33) | 124.1 (27.81) | 135.4 (13.78) | 145.5 (14.85) |
| | V | 125.2 (24.47) | 128.0 (21.37) | 133.1 (24.17) | 122.0 (24.87) | 130.54 (27.72) | 145.2 (25.19) |

**Table 4.**



**Table 4**. The mean in milliseconds of P2 peak latency for each condition at frontal (F), central (C) and central-parietal (CP) sites (SD in parentheses)

| | | P2 | | | | | |
|---|---|---|---|---|---|---|---|
| | | anger | | | fear | | |
| | | F | C | CP | F | C | CP |
| Dynamic visual type | A | 203.7 (19.64) | 222.8 (16.41) | 245.0 (24.45) | 224.6 (23.83) | 232.8 (15.77) | 263.3 (26.58) |
| | CAV | 212.0 (30.69) | 222.8 (28.01) | 243.9 (35.37) | 231.4 (21.26) | 233.8 (18.53) | 252.5 (24.27) |
| | IAV | 214.4 (16.52) | 213.0 (18.11) | 240.11 (26.19) | 228.3 (29.65) | 239.5 (23.92) | 256.4 (19.91) |
| | V | 248.3 (36.22) | 248.0 (42.06) | 234.87 (36.10) | 255.3 (26.85) | 237.8 (35.56) | 229.4 (41.65) |
| Static visual type | A | 212.1 (23.28) | 225.2 (20.04) | 253.1 (27.17) | 216.8 (23.52) | 232.6 (12.01) | 257.9 (27.97) |
| | CAV | 218.9 (21.72) | 222.3 (25.18) | 238.3 (31.00) | 232.2 (26.93) | 239.0 (26.26) | 245.78 (35.71) |
| | IAV | 231.5 (25.59) | 241.3 (17.24) | 253.3 (30.77) | 223.3 (23.87) | 230.1 (26.84) | 243.5 (31.65) |
| | V | 242.9 (38.07) | 241.0 (34.81) | 221.60 (29.38) | 227.3 (38.59) | 242.4 (36.40) | 227.6 (34.42) |