NeuroImage 109 (2015) 273–282

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Affective responses to emotional words are boosted in communicative situations

Lana Rohr *, Rasha Abdel Rahman

Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

ARTICLE INFO

Article history:
Accepted 9 January 2015
Available online 14 January 2015

Keywords:
Emotional words
ERPs
Communication
Auditory emotion
Social situations

ABSTRACT

Emotional verbal messages are typically encountered in meaningful contexts, for instance, during face-to-face communication in social situations. Yet, they are often investigated by confronting single participants with isolated words on a computer screen, thus potentially lacking ecological validity. In the present study we recorded event-related brain potentials (ERPs) during emotional word processing in communicative situations provided by videos of a speaker, assuming that emotion effects should be augmented by the presence of a speaker addressing the listener. Indeed, compared to non-communicative situations or isolated word processing, emotion effects were more pronounced, started earlier and lasted longer in communicative situations. Furthermore, while the brain responded most strongly to negative words when presented in isolation, a positivity bias with more pronounced emotion effects for positive words was observed in communicative situations. These findings demonstrate that communicative situations – in which verbal emotions are typically encountered – strongly enhance emotion effects, underlining the importance of social and meaningful contexts in processing emotional and verbal messages.

Introduction

Emotional verbal messages, even in the form of single words, are typically encountered in wider meaningful contexts and rarely seen or heard in isolation. For instance, during reading, emotional words are often embedded in longer sentence or text passages. Crucially, in social communicative situations they are experienced in the presence of a speaker directly addressing the listener. Such meaningful contexts may have a strong influence on how we experience and evaluate single words and their emotional and semantic contents (e.g. Hagoort and van Berkum, 2007; Nieuwland and Van Berkum, 2006). Yet, and in contrast to natural language processing in real life situations, communicative aspects have not yet been taken into account in studies on emotional word processing. In the typical lab situation, single participants are placed in front of a computer screen, reading visually presented words of varying emotional contents. Arguably, such situations lack ecological validity, and the potential of verbal stimuli to induce emotion effects may be largely underestimated (e.g. Recio et al., 2011; Trautmann et al., 2009).

Indeed, emotional words (e.g. Kissler et al., 2007; Schacht and Sommer, 2009a; Schacht and Sommer, 2009b) tend to induce weaker affective responses than other emotionally arousing stimuli such as facial expressions (Schupp et al., 2004b), body postures (Aviezer et al., 2012), gestures (Flaisch et al., 2011), visual scenes or objects (e.g. Lang et al., 1993; Schupp et al., 2003). Two recent studies have suggested higher visual complexity of pictorial stimuli to contribute to those differences (Schlochtermeier et al., 2013; Tempel et al., 2013). However, using simple line-drawings and only positive and neutral stimuli that were furthermore matched for arousal, these studies seem not fully conclusive. Another possible explanation following from the discussion above is that isolated words lack personal relevance when they are not embedded in personally meaningful contexts such as communicative situations that frame the emotional meaning of the words. First evidence for the contribution of context relevance in emotional word processing derives from studies providing self-relevant verbal contexts using sentences or personal pronouns preceding the emotional target words (Fields and Kuperberg, 2012; Herbert et al., 2011a,b). Taking additionally the symbolic nature of verbal stimuli into account – in contrast to the more direct emotional appeal of facial expressions or arousing scenes – the personal relevance of emotional words such as "love" or "idiot" may be considerably enhanced when experienced during face-to-face communication, resulting in stronger and more immediate affective responses.

Within current two-dimensional theories of emotion processing that focus on valence and arousal (Bradley and Lang, 2000; Lang, 1995; Lang et al., 1993, 1998) communicative situations can be assumed to increase the subjective valence of the words and/or the arousal they induce, and may thus intensify the emotional experience induced by those words. Alternatively, in appraisal theories of emotion (Ellsworth

* Corresponding author at: Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany. Fax: +49 30 2093 4910.
E-mail address: lana.rohr@hu-berlin.de (L. Rohr).

http://dx.doi.org/10.1016/j.neuroimage.2015.01.031
1053-8119/

and Scherer, 2003; Scherer and Peper, 2001) communicative situations can be assumed to directly affect appraisal of the pleasantness and self-relevance of emotional words, including their potential consequences for the listener (Grandjean and Scherer, 2008; Grandjean et al., 2008). In particular, the self-relevance of emotional words may differ largely between face-to-face communication and encounters of context-free single words.

To date, there is to the best of our knowledge no direct empirical evidence on the processing of emotional words in social-communicative contexts. Thus, this study was designed to investigate the consequences of the presence of a speaker on emotional word processing with electrophysiological measures of brain activity. Our goal was to provide insights into how the brain responds to emotional words in more realistic communicative, and thus personally more relevant and ecologically more valid, situations. We contrasted affective responses to emotional words experienced during communicative situations with the processing of the identical words in non-communicative situations. To this end we presented videos of a speaker with direct eye gaze, conveying a neutral facial expression, uttering emotional and emotionally neutral words. The speaker's gaze turned towards the perceiver during verbal and non-verbal communication signals attention to the perceiver which can be seen as a basic ingredient of face-to-face communication (e.g. Kampe et al., 2003; Vertegaal et al., 2001), enhances attention for the seen face (e.g. Bockler et al., 2014) and may facilitate speech perception, especially when there is only speech but no accompanying gesture information present (Holler et al., 2014). Thus, as a control that also included the presentation of a face, we introduced a non-communicative condition in which videos of the same speaker's face with closed eyes and mouth were presented, signaling that the words were not uttered by the person seen in the video. Seeing a person's face while hearing other persons talk is a rather common situation in real life.

Please note that our focus here is on the social communicative effects of a speaker directly addressing the listener, rather than investigating mechanisms of audiovisual integration. Ample evidence has demonstrated that the congruency of multimodal stimuli may facilitate the perception and identification of emotional (e.g. Paulmann and Pell, 2011; Paulmann et al., 2009; see Klasen et al., 2012 for a review) and non-emotional speech signals (e.g. Schwartz et al., 2004; van Wassenhove et al., 2005) and is mandatorily processed (e.g. de Gelder and Vroomen, 2000) already during early perceptual processing stages (e.g. de Gelder et al., 1999; Gerdes et al., 2013; Pourtois et al., 2000, 2002; Stekelenburg and Vroomen, 2007) probably involving specialized structures (e.g. de Gelder and Van den Stock, 2011), while incongruent audiovisual input can even lead to perceptual illusions (cf. McGurk-effect; e.g. McGurk and MacDonald, 1976). Our aim here was not to add evidence on this issue, but instead to concentrate on comparing the effects of emotional and non-emotional words separately in communicative and non-communicative situations, rather than directly contrasting word processing in the presence vs. absence of converging information from the visual modality or the congruency between visual and auditory information. Of course, the processing and integration of combined auditory and visual information is an integral component of face-to-face communication that should affect our responses to verbal messages. However, other determinants like enhanced attention towards the speaker (e.g. Bockler et al., 2014) and enhanced personal/social relevance induced by being directly addressed should have strong – and yet to be determined – effects. We believe that here, the social-communicative aspects play a crucial role. For instance, in contrast to the social relevance manipulated here, there is no a priori reason to assume that audio-visual integration affects the processing of emotional and neutral words differentially. Thus, while audio-visual integration plays an undisputed role in face-to-face communication, the social and communicative aspects can be expected to specifically influence the processing of communicated emotional and personally relevant messages.

At last, because in contrast to the well-investigated effects of emotional word reading little is known about the electrophysiological correlates of these effects in the auditory modality, we additionally conducted a pre-experiment in which the identical words were presented in isolation in the visual and auditory modality.

We focused on a temporal and functional characterization of affective responses to socially communicated emotional words, exploiting the high temporal resolution of event-related brain potentials (ERPs). Two ERP components have been repeatedly reported to reflect emotional responses to different types of visual stimuli such as faces, scenes, objects or words. The first component is the early posterior negativity (EPN), a relative negative deflection over posterior brain regions, occurring around 200 to 300 ms (e.g. Flaisch et al., 2011; Herbert et al., 2008; Kissler et al., 2006; Recio et al., 2011; Schacht and Sommer, 2009a,b; Schupp et al., 2004a). The EPN has been reported primarily for visual stimuli and is taken to reflect early reflexive attention to and enhanced visual perception of affective stimuli (e.g. Junghöfer et al., 2001; Kissler et al., 2007; Schupp et al., 2003; for reviews see Citron, 2012; Kissler et al., 2006). The EPN does not seem to be strongly modulated by the (semantic) depth of word processing or the nature of the task (Kissler et al., 2006; Schacht and Sommer, 2009b). Furthermore, this component has been demonstrated to vary independent of the self-reference of emotional stimuli. Specifically, emotional visual words induced comparable EPN modulations when preceded by personal pronouns ("my") or definite articles without self-reference ("the"; e.g. Herbert et al., 2011b).

At later stages, the late positive potential (LPP), peaking at about 400 to 700 ms over centro-parietal regions, has been associated with the more elaborate processing and appraisal of the intrinsic relevance of emotional stimuli (e.g. Bayer et al., 2010; Cacioppo et al., 1993; Cuthbert et al., 2000; Schacht and Sommer, 2009a). This component is directly affected by task demands and the relevance of emotion for the task (e.g. Fischler and Bradley, 2006; Rellecke et al., 2011; Schacht and Sommer, 2009b). Furthermore, LPP amplitude has been shown to be enhanced for self-referential emotional stimuli (e.g. "my success" vs. "the success"; Herbert et al., 2011a,b).

If social communicative contexts, as hypothesized above, increase the personal relevance of emotional words and therefore the subjective valence and arousal levels, this intensified experience should be reflected in augmented effects of emotion at early points in time (reflecting fast reflexive processing of emotional stimuli) and in later more sustained evaluative processes (effects in the LPP component). While the LPP effects can be expected to be present irrespective of the presentation modality, the predictions for the expected early reflexive effects cannot directly be related to a specific component because thus far, early emotion effects were mostly found at posterior sites for visual materials (EPN). However, analogously to the visual effects, there might be fast reflexive responses to auditorily presented emotional words (possibly at fronto-central regions, see below), which should be enhanced in communicative situations.

Pre-experiment

In the pre-experiment the processing of context-free visual and auditory emotional words was investigated. The purpose was to evaluate the stimulus materials and to test whether the typical emotion effects of isolated visual words can be observed with the present materials. Furthermore, we aimed at comparing the effects for visual words to the processing of the same words in the auditory modality. As mentioned above, there is very little evidence on the electrophysiological correlates of auditory emotional words. One of the major reasons for this discrepancy between modalities may be the incremental nature of auditory signals, resulting in a smearing of ERP components and effects that reduces the signal to noise ratio for experimental manipulations.

Generally, processing of auditory stimuli elicits a characteristic ERP pattern of subsequent P1, N1 and P2 components over central and fronto-central regions (e.g. Martin et al., 2008). Usually, early components like the auditory P1 and N1 are attributed mainly to physical stimulus characteristics, while the P2 has also been linked to the processing of emotional cues, provided, for example, via prosodic information. The auditory P2 might be functionally similar to the visual EPN (Kotz and Paulmann, 2011). As processing auditory emotional stimuli enhances activity of the auditory cortex (e.g. Plichta et al., 2011), processing of emotional words could also be associated with enhancements of auditory evoked potentials.

In addition to the P2, later ERP components have been associated with auditory emotion processing, such as the P3 (e.g.Thierry and Roberts, 2007; Wambacq and Jerger, 2004) and the N400 component (review: Kotz and Paulmann, 2011). Up until today most evidence on auditory emotion processing originates from investigations of nonverbal stimuli as, for instance, the sound of a crying baby or a growling dog (e.g. Bradley and Lang, 2000; Czigler et al., 2007; Plichta et al., 2011; Thierry and Roberts, 2007) while few studies varied verbal emotion (e.g. Graß et al., 2014; Kotz and Paulmann, 2007; Paulmann and Kotz, 2008).

Method

Participants

Twenty-four native speakers of German (all women, 23 right-handed; mean age: 25 years, range: 18–34) gave informed consent to participate in the study. They received payment or course credit for participation. All participants reported normal or corrected-to-normal vision and normal hearing. Data of two additional participants were excluded due to artifacts in the EEG signal. The study was approved by the local ethics committee and conducted in accordance with the declaration of Helsinki.

Materials and procedure

We selected 240 German nouns and adjectives of neutral, negative and positive valence. Because German nouns are gender-marked, the female word forms were selected for all stimuli. Including female participants only, this was done to increase the potential relevance of the words for the listener in the communicative situations investigated in the main experiment. Based on normative values taken from the German dlexDB database (accessible via www.dlexdb.de; Heister, 2011) the words were matched for length (letters, syllables), word frequency and number of orthographic neighbors (cf. Table 1). We did not control initial phonemes and can therefore not exclude differences between emotion conditions regarding this parameter. However, the main conclusions of the study should not be affected, as the identical stimuli are presented in and compared between the different communication conditions.

Separate ratings of valence and arousal were conducted with a group of 12 participants who did not participate in the pre-experiment or main experiment. All visually presented words were rated on five-point SAM-scales (Bradley and Lang, 1994) for valence and arousal. Emotion conditions (cf. Table 1) differed significantly in valence, $F(1,11) = 151.85$; $p < .001$, $\eta^2 = .93$, and arousal ratings, $F(1,11) = 9.72$; $p = .006$, $\eta^2 = .47$. Planned contrasts confirmed that negative words were rated more negative, $F(1,11) = 122.75$; $p < .001$, $\eta^2 = .918$, and more arousing than neutral words, $F(1,11) = 16.76$; $p = .002$, $\eta^2 = .604$, and positive words were rated as more positive, $F(1,11) = 197.9$, $p < .001$, $\eta^2 = .947$, and more arousing than neutral words, $F(1,11) = 34.97$; $p < .001$, $\eta^2 = .761$.

Visual words were presented in white font (Arial 24 pt) at the center of a black screen. A trial started with a fixation cross presented for 400 ms (cf. Fig. 1). Then the word was presented and remained on the screen for 1 s, followed by a blank screen for 1 s. Auditory words were presented 400 ms after the onset of the fixation cross which remained on the screen for 1 s. The sound files of the spoken words presented in the auditory condition were identical to those used in the main experiment (see recording procedure below). Auditory word onsets were determined manually using GoldWave and correspond to the first visible deviation of the auditory signal from the silence level. In the pre-experiment we adjusted lead-in time auf the auditory files to a constant auditory word onset at 400 ms after fixation cross onset.

Participants were instructed to attend to the words and to memorize them. To draw participant's attention to the stimuli, control words were presented randomly after 7, 9, 11 or 13 trials, and participants were instructed to indicate via button press whether they had read or heard the word before. This procedure without trial-wise button press responses was used to keep the experimental setting as realistic as possible. Typically, we process verbal messages more implicitly, without directly classifying the words via button presses as emotional or arousing. The presentation modalities (visual and auditory) were blocked in counterbalanced order across participants. The sequence of emotion conditions within the blocks was fully randomized for each participant and block individually. Each word was presented two times in each modality, leading to 160 trials per emotion condition, with a duration of approximately 50 min.

Table 1
Descriptive statistics of the selected stimuli.

		Negative	Neutral	Positive	Main effect of emotion $df = 2{,}237$
Word properties	Letters	7.4	7.4	7.3	$F = .12$ $p = .89$
	Syllables	2.4	2.4	2.4	$F = .30$ $p = .74$
	Orth. neighbors	3.3	3.4	3.3	$F = .03$ $p = .98$
	Freq/1 mio	5.7	8.2	7.4	$F = 1.30$ $p = .27$
	Valence (rating)	1.9	3.1	3.9	$F = 151.85$ $p < .001$
	Arousal (rating)	2.9	2.1	2.6	$F = 9.72$ $p = .006$
Video properties	Word duration [ms]	658	648	661	$F = .35$ $p = .71$
	Word onset [ms]	543	532	550	$F = 1.01$ $p = .37$
	Mean sound intensity [dB]	75.0	75.2	75.6	$F = 2.06$ $p = .13$
	Valence (FaceReader)	−.26	−.27	−.26	$F = .71$ $p = .49$

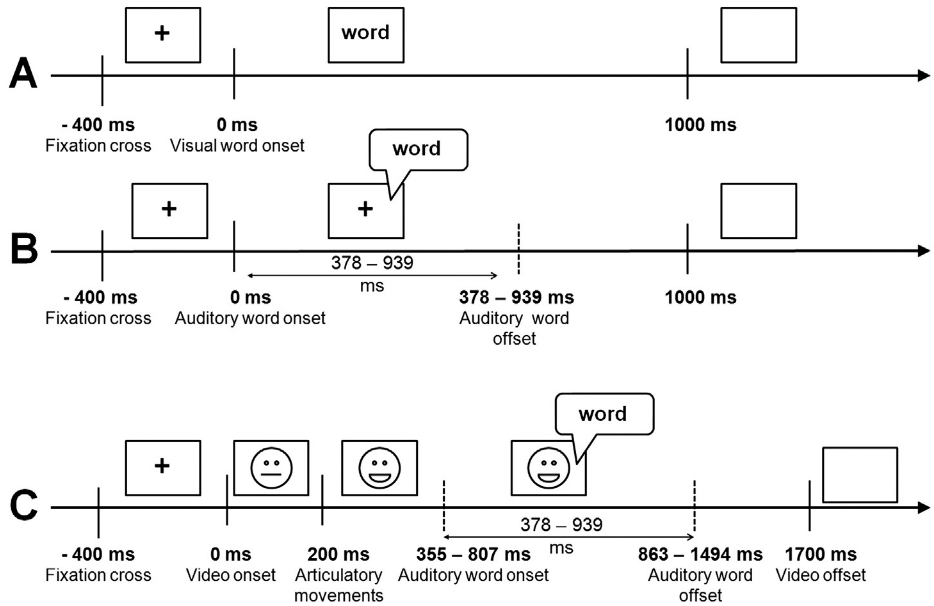

Fig. 1. Illustration of the visual (A) and auditory (B) condition in the pre-experiment and the communicative condition in the main experiment (C).

EEG-recording and analysis

The EEG was recorded from 62 sites according to the extended 10–20 system with Ag/AgCl-electrodes at a sampling rate of 500 Hz and online-referenced to the left mastoid. Impedances were kept below 5 kΩ. Horizontal and vertical electrooculograms were recorded with electrodes attached to the left and right canthi of the eyes and above and below the left eye. Offline the EEG was re-referenced to an average reference and low-pass-filtered (30 Hz). Eye-movement artifacts were corrected using a spatio-temporal dipole modeling procedure implemented in the BESA-software (Berg and Scherg, 1991). Remaining artifacts were rejected in a semiautomatic procedure (amplitudes and amplitude changes higher than 200 μV, voltage steps higher than 50 μV/ms). Artifact-free EEG was segmented time-locked to the auditory or visual word onset, respectively, to prevent smearing of averaged ERPs due the variation of onset times between different auditory words. Segments were corrected to a 100 ms baseline before auditory or visual word onset.

We performed repeated measures ANOVAs including all electrodes on successive 50 ms time windows from 0 to 800 ms with the factors electrode (62 levels) and emotion (positive, negative, and neutral), separately for visually and auditorily presented words (omnibus ANOVA). For this analysis, because an average reference was used, only effects in interaction with electrode are reported as significant effects. Furthermore, to more specifically characterize early and late emotion effects, we focused on regions and time windows of interest that are based on reports of temporal and topographical distributions of emotion effects in the literature (e.g. Schupp et al., 2003). Early posterior modulations in the EPN time window were investigated in two consecutive 50 ms intervals between 200 and 300 ms (electrode sites: PO7, PO8, PO9, PO10, O1, O2, O9, O10). For auditory words we additionally defined a fronto-central ROI (electrode sites: Fz, AFz, F3, F4, AF3, AF4) to capture effects in the early auditory components. Here, we analyzed six consecutive time-windows of 50 ms between 0 and 300 ms. Later centro-parietal effects in the LPP component were analyzed in the time-windows between 500–650 and 650–800 ms (electrode sites: POz, Pz, CPz, CP1, CP2). Because visual inspection of the data showed a central negativity between 300 and 400 ms, we additionally included this time window in the analyses on the central ROI. Huynh–Feldt corrections were applied throughout when the sphericity assumption was violated. If a significant emotion effect was found in the ROI-analyses, planned contrasts were used to check whether positive words, negative words or both differed from neutral words.

Results and discussion

Visual words

For visual words the omnibus ANOVA across all electrodes (cf. Table 2) revealed an emotion effect in the EPN between 250 and 300 ms. This was confirmed by the EPN ROI analyses, yielding early effects of emotion only in this time window, $F(2,46) = 4.71$; $p = .014$, $\eta^2 = .17$. Here, negative words induced a relative posterior negativity compared to neutral words (cf. Fig. 2), $F(1,23) = 5.79$; $p = .025$, $\eta^2 = .20$, whereas the comparison between positive and neutral words failed to reach significance, $F(1,23) = 0.46$; $p = .504$, $\eta^2 = .02$. A later modulation was found in the omnibus analysis between 400 and 550 ms. However, the topographical distribution with an posterior

Table 2
F-values and significance levels of the omnibus-ANOVAs for the pre-experiment.

Time window (ms)	Visual presentation	Auditory presentation
df	*122,3538*	*122,3538*
0–50		
50–100		
100–150		
150–200		
200–250		
250–300	2.205*	
300–350		
350–400		
400–450	3.280**	
450–500	2.884**	
500–550	2.771**	
550–600		
600–650		
650–700		
700–750		
750–800		

Because an average reference was used, all factors are reported in interaction with electrode.
Reported values are Huynh–Feldt corrected.
* $p < .05$.
** $p < .01$.

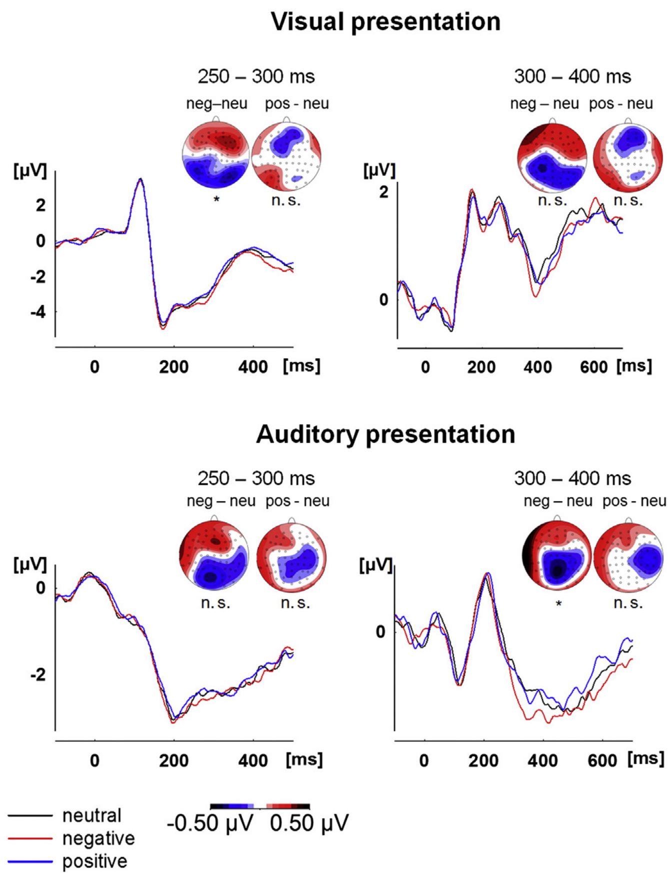

Fig. 2. Effects of emotion in visual and auditory word presentation in the pre-experiment pooled over the posterior (left) and the central sites (right) used in the ROI analyses. $^*p < .05$. n.s. not significant.

negativity did not resemble emotion-induced LPP modulations reported in the literature (see above). Accordingly, no effects were found in the LPP ROI at centro-parietal sites in any analyzed time window (300–400 ms and between 500 and 800 ms), $Fs \leq 1.06$; $ps \geq .351$, $\eta^2 \leq .044$. Instead, this effect may reflect a continuation of the EPN we found at 250 ms. To test this hypothesis and check whether we could have missed an earlier, short lived LPP-effect, we extended the EPN and LPP ROI analyses to two time-windows not yet included in the ROI analyses, namely 400–450 and 450–500 ms. In these analyses emotion effects were not significant at central sites, $Fs \leq 2.13$; $ps \geq .137$, $\eta^2 \leq .085$, but in the posterior ROI, $Fs \geq 5.33$; $ps = .008$, $\eta^2 \geq .188$, supporting the assumption, of a continuation of the EPN-effect.

Typically, EPN effects are found at around 200–300 ms, but there are also studies reporting later onsets (e.g. Palazova et al., 2011; Schacht and Sommer, 2009b). Therefore, for visually presented words we replicated known effects, demonstrating that our materials are suitable for the present purpose. We find emotion effects in the EPN component, including a negativity bias with strong effects primarily for negative but not for positive words (e.g. Holt et al., 2009; Kanske and Kotz, 2007). Emotionally induced LPP effects were not observed. This is in line with several reports that the LPP is influenced by task demands and more pronounced for tasks that require in-depth semantic processing (Hinojosa et al., 2010; Schacht and Sommer, 2009b) or attention towards the emotional content of stimuli (Rellecke et al., 2011), which was not the case here.

Auditory words

In the auditory modality neither the omnibus ANOVA (cf. Table 2) nor the EPN, $Fs \leq 1.93$; $ps \geq .179$, $\eta^2 \leq .077$, or LPP ROI analyses from 500 to 800 ms, $Fs(2,46) \leq 3.09$; $ps \geq .055$, $\eta^2 \leq .118$, revealed significant emotion effects. However, early EPN effects were not expected, in line with the interpretation that this component reflects enhanced visual processing of affective stimuli. Instead, emotion-induced modulations of the early auditory components are possible. However, the analyses in the fronto-central ROI analysis did not reveal significant emotion effects between 0 and 300 ms, $F(2,46) \leq 3.214$; $p \geq .054$, $\eta^2 \leq .123$. The absence of the LPP effect replicates the findings from visual words and is probably again task-related (see above). Even though no emotion effects were visible in the analysis over all electrodes, between 300 and 400 ms we observed a significant emotion effect at the central ROI in this time window, $F(2,46) = 7.45$; $p = .002$, $\eta^2 = .245$, for negative, $F(1,23) = 12.03$; $p = .002$, $\eta^2 = .343$, but not for positive words, $F(1,23) = 0.26$; $p = .613$, $\eta^2 = .011$. There are several possible explanations for this effect. First, it could be due to enhanced (and relatively late) sensory processing of emotional compared to neutral words. Alternatively, this effect could reflect semantic aspects of emotional word processing, as emotion-related N400 modulations for visual words have been reported before (Herbert et al., 2008; Kanske and Kotz, 2007; Wabnitz et al., 2012). Another possible origin may be the so-called self-positivity bias. Specifically, an N400-like central negativity for negative words that do not match the individual (usually positive) self-concept was reported by Watson and colleagues (Watson et al., 2007). This effect would also be in accordance with our finding of a pronounced negativity in particular for negative words. In any case, the central negativity demonstrates that the emotional content of auditorily presented words was processed by the listeners. Like in the visual presentation, a negativity bias was found for isolated auditory emotional words.

Main experiment

Method

Participants

Thirty native speakers of German (all women, right-handed, mean age: 25 years, range 18–37) received payment or course credit for participation. All participants gave informed consent to participate in the study and reported normal or corrected-to-normal vision and normal hearing. Data of two additional participants were excluded due to EEG artifacts. None of the participants took part in the rating of the materials or in the pre-experiment. The experiment was approved by the local ethics committee and conducted in accordance with the declaration of Helsinki.

Materials and design

We recorded 240 short videos containing headshots of a female professional speaker uttering single words, directly fixating the camera. To avoid confounding influences from contextual emotional sources, the speaker was instructed to keep prosody and facial expressions neutral. The videos were edited to start precisely 200 ms before the first articulatory movement was visible, which was manually determined offline. In the main experiment, the auditory word onsets occurred between 355 and 807 ms after video-file onset. The onset time of a given stimulus depended mainly on the initial phoneme of the word. Neither mean articulation duration nor mean auditory word onset (interval between video onset and onset of the auditory signal) or mean sound intensity differed significantly between the emotion conditions (cf. Table 1).

Differences in emotional facial expressions between emotional word conditions were tested using the FaceReader software (Version 4.0.8). Valence values were calculated for each frame of each video and averaged over all frames within each utterance. Mean valence values of the videos were submitted to a one-way ANOVA with the factor emotion (neutral, positive, negative). According to the software rating, the speaker's facial expression was slightly negative in all conditions (cf. Table 1) but it was not affected by emotion condition, $F(2,237) = .71$, $p = .493$, $\eta^2 = .006$.

In the communicative condition the words were presented along with their original video recordings. For the non-communicative control condition 20 additional video clips of the speaker with closed mouth and eyes were taken. Identical auditory files were presented in the communicative and non-communicative contexts by extracting and recombining the audio tracks of the videos. In the non-communicative condition the auditory words were presented with randomly selected videos of the non-articulating speaker with closed eyes and mouth. Thus, the visual input in the communicative and non-communicative situation was kept as similar as possible. We have also included a third presentation mode in which the auditory word stimuli were combined with a video of the empty studio to prevent habituation to the presence of a face.

All videos were presented at an approximate size of 10.5 × 12 cm and a viewing distance of approximately 90 cm. Each trial started with a white fixation cross on a black background. After 400 ms the audio file started and the fixation cross was replaced by a video. All videos were presented for 1700 ms (cf. Fig. 1). The task was the same as in the pre-experiment. The emotional and communicative conditions were presented in randomized order with the restriction that no more than 4 consecutive trials included the same emotion or communicative condition. Each word was presented twice in each condition, resulting in 160 trials per emotion in each communicative condition, and total experiment duration of approximately 90 min.

EEG-recording and analysis

EEG-recording and processing were identical to the procedures in the pre-experiment. The same omnibus and ROI analyses were performed as in the pre-experiment, separately for the communicative and non-communicative situation.

Results

Communicative situation

The ANOVAs including all electrodes (summarized in Table 3) revealed an early effect between 50 and 100 ms and continuous effects of emotion in each time window between 150 and 800 ms post-stimulus. In the early time windows we found an enhanced posterior positivity for emotional relative to neutral stimuli. Topographically, this effect resembles an EPN with a reversed polarity (cf. Fig. 3). EPN ROI analyses of the early emotion effects were significant in both time windows between 200 and 300 ms, $Fs(2,58) \geq 7.49$; $ps \leq .002$, $\eta^2 = .205$. Specifically, between 200 and 250 ms there was an effect of positive words only, $F(1,29) = 14.22$; $p = .001$, $\eta^2 = .329$, whereas both positive and negative words differed significantly from neutral words between 250 and 300 ms ($F(1,29) = 12.60$; $p = .001$, $\eta^2 = .303$ and $F(1,29) = 4.23$; $p = .049$, $\eta^2 = .127$, respectively). At the fronto-central ROI early emotion effects were significant from 200 to 300 ms, $Fs(2,58) \geq 3.19$; $ps \leq .048$, $\eta^2 \geq .099$, for positive but not for negative compared to neutral words (positive: $Fs(1,29) \geq 5.89$; $ps \leq .022$, $\eta^2 \geq .169$, negative: $Fs(1,29) \leq 0.91$; $ps \geq .347$, $\eta^2 \leq .031$).

Whereas robust and long-lasting emotion effects were found in the LPP time windows across all electrodes (cf. Table 3), the LPP ROI analyses did not reflect this pattern, $Fs(2,58) \leq 1.67$; $ps \geq .140$, $\eta^2 \leq .066$, suggesting emotion effects with a different topographical distribution that cannot be related to the LPP (see also Fig. 3). In fact, the topographical distribution suggests that the posterior positivity continued in later time-windows for positive words. Indeed, post-hoc tests extending the analyses on the EPN ROI on all the subsequent 50 ms time windows until 800 ms revealed significant effects in all time windows for positive words and between 650 and 800 ms for negative words.

As in the pre-study, there was a central negativity between 300 and 400 ms, which was significant in the central ROI, $F(2,46) = 3.82$; $p = .028$ $\eta^2 = .116$, for negative, $F(1,23) = 6.06$; $p = .020$ $\eta^2 = .173$, but not for positive words, $F(1,23) = 0.26$; $p = .614$ $\eta^2 = .009$.

Non-communicative situation

As can be seen in Table 3/Fig. 3, the effects of emotion observed in the communicative situation are massively reduced or entirely absent in the non-communicative situation. In the omnibus ANOVA main effects of emotion were found between 250 and 350 ms and from 550 to 600 ms. However, an EPN-like effect was neither confirmed by the EPN ROI analyses, $Fs(2,46) \leq 2.01$; $ps \geq .143$, $\eta^2 \leq .065$ nor were there significant emotion effects in the LPP ROI analysis starting at 500 ms, $Fs(2,46) \leq 1.67$; $ps \geq .167$, $\eta^2 \leq .054$, suggesting that classic emotion effects were not present. Here, the fronto-central ROI showed a main effect of emotion between 250 and 300 ms, $F(2,46) = 3.69$; $p = .031$, $\eta^2 = .113$, that reached significance neither for positive nor for negative compared to neutral words in the follow-up analyses, $Fs(1,29) \leq 2.81$; $ps \geq .104$, $\eta^2 \leq .088$. Instead, a negativity was found in the central ROI between 300 and 400 ms, $F(2,46) = 4.09$; $p = .024$ $\eta^2 = .124$, for negative, $F(1,23) = 9.24$; $p = .005$ $\eta^2 = .242$, and, despite a slightly different distribution, also for positive words, $F(1,23) = 5.22$; $p = .030$ $\eta^2 = .153$.

Table 3
F-values and significance levels of the OMNIBUS-ANOVAs for the main experiment.

Time window (ms)	Communicative situation	Non-communicative situation
df	*122,3538*	*122,3538*
0–50		
50–100	1.939*	
100–150		
150–200	2.251*	
200–250	2.822**	
250–300	2.961**	1.790*
300–350	3.079**	3.585***
350–400	3.415***	
400–450	3.292**	
450–500	3.611***	
500–550	3.214**	
550–600	2.206*	2.097*
600–650	2.899**	
650–700	2.889**	
700–750	2.207*	
750–800	2.324**	

Because an average reference was used, all factors are reported in interaction with electrode.
Reported values are Huynh–Feldt corrected.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Discussion

In the present study we investigated the impact of social-communicative contexts on affective brain responses to emotional words, assuming that communicative situations enhance the personal relevance, and therefore the susceptibility, for verbal emotional messages. The pre-experiment established the baseline effects of the emotional words used in the present study. In a standard situation in which isolated words are visually presented to single participants we observed an often reported modulation in the early posterior negativity that is taken to reflect reflexive attention to emotional stimuli and enhanced visual perception. Thus, the present materials induce classic and well-documented emotion effects. Different effects were observed when the same words were presented in the auditory modality. We have not found an early brain response that may be viewed as an auditory pendent to the EPN observed in the visual modality. For auditorily presented words emotion effects were generally weaker (as revealed by the analyses across all electrodes) and mainly characterized by a focal N400-like central negativity. Generally, processes related to the meaning of auditory words (such as their emotional quality) are difficult to investigate using ERPs because of their temporal dynamics. Auditory

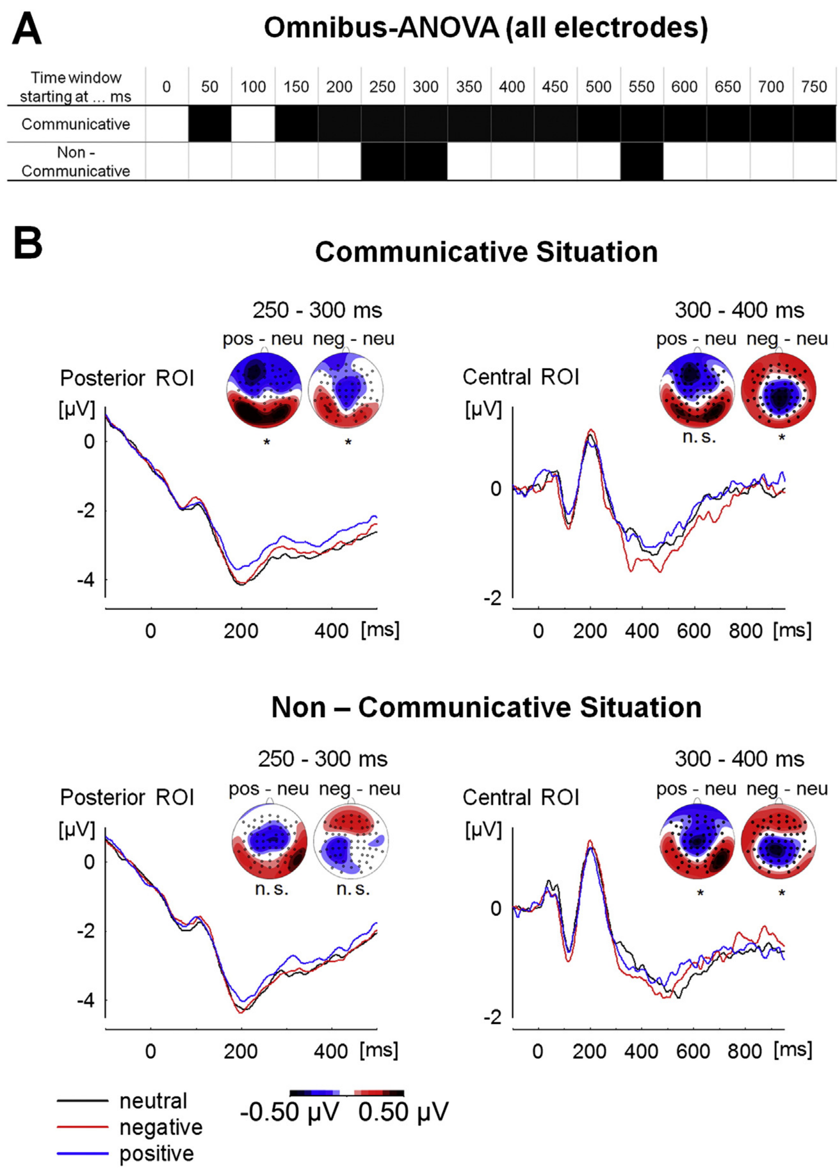

Fig. 3. Effects of word emotion on ERPs in the main experiment. (A) Illustration of the results of the omnibus-ANOVA for the communicative and the non-communicative situation. Time windows with a significant main effect of emotion are shaded black in the table. (B) Emotion effects in the analyzed regions and time-windows of interest. Presented ERPs are pooled over the posterior (left) and the central sites (right) used in the ROI analyses. *$p < .05$. n.s. not significant.

information unfolds over time and the moment of word identification is – in addition to the lexical word – also affected by contextual factors like how constraining the context is. Thus, word identification can neither be related (only) to word-length, nor can the variation be entirely prevented by time-locking to auditory stimulus onset, making it very difficult to calculate ERPs time-locked to "semantic access" (please see also the discussion on uniqueness point and word identification speed below) The resulting strong inter-item variation might lead to a smearing in ERP components. However, this should affect all conditions similarly and therefore should not cause differential effects between conditions. The N400-like modulation may be closely related to auditory perception that is reflected in topographically similar modulations with a central maximum, even though the effect is comparatively late, starting about 100 ms after the P2 component associated with auditory perception and emotion. In sum, the pre-experiment revealed classic effects of visual emotional word processing and weaker effects in the auditory modality that reflect semantic processing of emotional meaning and may be closely related to auditory perception. N400 like modulations related to auditory emotion processing have been reported before (e.g. Paulmann and Pell, 2010; Schirmer and Kotz, 2003).

The non-communicative condition of the main experiment included the presentation of the speaker's face with closed eyes and mouth. Here we observed an effect pattern comparable to the pattern found for auditory words presented in isolation, namely, N400 and P2 like ERP modulations at (fronto-)central sites that may reflect auditory/semantic processing of emotional words. The similarity of the effects to the presentation of isolated words confirms that we have successfully created a non-communicative situation despite the presence of a human face. When eyes are closed and the person is not speaking emotional words are not experienced as emotional messages from the seen person.

In marked contrast to the non-communicative situation and to the presentation of isolated visual or auditory words, the same words elicited very strong and robust emotion effects in the communicative situation. These effects started as early as about 50 to 100 ms, and were present continuously between 150 and 800 ms after word onset. The onset of continuous emotion effects from 150 ms onwards is remarkably fast, given that the words were presented auditorily. Auditory words are incremental in the sense that the information is delivered over time, in contrast to the direct and immediate presence of the entire information for visual words (e.g. Bradley and Lang, 2000). Thus, a cohort of several potential lexical candidates is initially co-activated, and this cohort narrows down over time until only one candidate is left. Indeed, emotion effects in such early time windows are often attributed to physical stimulus properties such as pitch and intensity, which are modulated by arousal, while meaning effects are considered to start in time windows of about 300 ms (e.g. Kotz and Paulmann, 2011). This is in line with the present observations of emotion effects in this time range when the auditory words were presented in isolation or in the non-communicative situation. Crucially, the early onset of auditory emotional word effects is restricted to the communicative situation. Thus, communicative situations may not only enhance the magnitude of emotion effects, they also seem to induce very early and long lasting affective brain responses to emotional words. As discussed above, we account for the fast access to emotional meaning with the enhanced impact and personal relevance of verbal emotional messages in communicative situations. The early onset of emotional meaning effects is plausible considering that lexical access can be achieved within 50 ms after crossing the uniqueness point (MacGregor et al., 2012). Furthermore, some evidence suggests that semantic processing can start even before a word is uniquely identified (Van Petten et al., 1999). At last, auditory stimulus identification is enhanced by increasing levels of predictability (van Wassenhove et al., 2005) and should therefore be pronounced for repeated stimuli — as was the case here. Thus, fast semantic processing before the uniqueness point is reached, the repetition of the relatively small stimulus set presented here, and processing advantages of audiovisual stimulation (see below) may have narrowed down the lexical cohort and facilitated word processing, paving the way for the very fast extraction of emotional word meanings during communicative situations.

As the non-communicative condition here was established by presenting a video of a non-talking person, we cannot easily determine to what extent the congruency of the audiovisual signal in the communicative situation contributes to the increased emotion effects. Crucially, we consider audiovisual integration one of the central ingredients of face-to-face communication that is missing in isolated word processing. Thus, while the combination of audiovisual integration and social-communicative processing is a natural feature of communicative situations, audiovisual integration as such cannot explain the augmented emotion effects or their time course. Congruent audiovisual words are known to be recognized more easily and the visual information provided by the face and mouth before the onset of the auditory signal contributes to the recognition of the uttered word (e.g. Schwartz et al., 2004; van Wassenhove et al., 2005). However, such processing advantages should facilitate the perception of emotional and neutral words in a similar way. Furthermore, the boosted emotion effects show not only an early onset, they also last longer, which is not easily explained with facilitated audiovisual processing of the words. Therefore, we conclude that the enhanced impact of emotional words in communicative situations is a consequence of the higher personal relevance and associated increased valence and arousal. Nevertheless, the potential interplay between audiovisual integration and top-down-modulated social-communicative factors may be an interesting subject for future studies.

To test this conclusion more directly, we conducted an additional valence and arousal rating (7-point Likert scale) in which participants ($N = 12$) rated the words in the communicative and non-communicative condition, presented in random order. To avoid repeated ratings of identical words (once on the communicative and once in the non-communicative condition), half of the words were presented in the communicative and the other half in the non-communicative condition. The assignment of words to conditions was randomized for each participant with the restriction that an equal number of neutral, positive and negative words were assigned to each communicative condition. ANOVAs with the factors emotion and communicative condition revealed valence and arousal effects for positive and negative compared to neutral words (all Fs ≥ 23.3, ps $\leq .001$, $\eta^2 \geq .680$), replicating the ratings for visually presented words (see above). Crucially, higher arousal values were found for words presented in the communicative compared to the non-communicative condition ($F = 6.4$, $p = .028$, $\eta^2 = .369$). Emotion and communicative condition did not interact in the valence rating ($F = 0.7$, $p = .49$, $\eta^2 = .060$), but a significant interaction of both factors was found in the arousal rating ($F = 4.7$, $p = .023$, $\eta^2 = .299$), confirming more pronounced emotion effects in the communicative ($F = 39.8$, $p \leq .001$, $\eta^2 = .784$; neutral $M = 2.21$, $SD = .91$; negative $M = 3.63$, $SD = 1.17$; positive $M = 2.75$, $SD = 1.05$) than in the non-communicative condition ($F = 22.7$, $p \leq .001$, $\eta^2 = .674$; neutral $M = 2.10$, $SD = .82$; negative $M = 3.30$, $SD = 1.18$; positive $M = 2.71$, $SD = 1.18$). The enhanced arousal values in the communicative condition confirm our conclusion that social-communicative situations enhance the personal relevance of emotional words, as reflected in the boosted emotion effects in ERPs. This finding is difficult to reconcile exclusively with mechanisms of audio-visual integration.

As described above, care was taken during construction of the materials to avoid differences in emotional prosody and facial expressions between the emotion conditions. While we cannot rule out prosodic differences between auditory emotional and neutral words entirely, such differences cannot explain our findings because emotional prosody and any other systematic differences between the emotion conditions should have similar effects in the non-communicative situation and for isolated auditory word processing because identical auditory word files were presented. Can differences in emotional facial expressions explain the current findings? Even though the speaker was instructed to keep her facial expression neutral, subtle differences in expressions when producing emotional compared to neutral words may be a confounding factor. In this case, part of the effects could be driven by processing an emotional facial expression, varying only in the communicative condition. However, this is unlikely because we have not found any differences in facial expressions between the emotional word conditions using FaceReader software. Moreover, emotional face processing has been associated with EPN and LPC modulations (e.g. Recio et al., 2011; Schupp et al., 2004b), which we did not find here. In contrast, we observed an early posterior positivity. Therefore, facial expressions can be ruled out as a source of the emotion effects observed in the communicative situation.

Another point is the validity of our experimental design. Natural language in real-life occurs usually in a rich contextual setting, including relationships, shared knowledge, larger semantic and situational contexts, different speaker identities etc. that are taken into account immediately (e.g. Van Berkum et al., 2008). Here, participants were confronted with an unfamiliar person, lacking these contexts. That we still find augmented emotion effects demonstrates, in our view, the considerable power of social-communicative situations to modulate our reception of verbal messages. Providing richer social situations and contexts may even result in stronger modulations.

Affective brain responses in communicative situations are topographically reflected simultaneously in an early posterior positivity and a negativity at fronto-central regions, starting (at the latest) at about 150 ms after word onset. Given the relevance of the integrated audiovisual information in the communicative relative to the non-communicative and isolated presentation conditions this pattern of brain activation most likely reflects the combined modulation of (posterior) visual and (fronto-central) auditory processing. The perception of synchronous visual and auditory input in the communicative situation can also be assumed to trigger activity in audiovisual integration structures, and correlates of audiovisual integration can be measured during sensory processing already before 200 ms (e.g. Giard and Peronnet, 1999; Pourtois et al., 2000). While the fronto-central negativity may reflect enhanced auditory processing in the P2, there are at least two possible accounts for the posterior positivity. First, very early visual areas may be activated by the emotional verbal input. Indeed, some evidence suggests that the processing of non-verbal emotional auditory stimuli can enhance early visual processing in the form of enhanced amplitudes in the P1 component (e.g. Brosch et al., 2009; Gerdes et al., 2013).

Similar to affective prosody, affective semantics may increase the salience of the emotional stimuli and elicit an analogous longer lasting posterior ERP effect in the P1 range.

Alternatively, concurrent attentive processing of the highly relevant auditory information in the emotional condition may directly influence activity in the visual cortex. This may reduce the extent to which the less informative visual information is processed attentively. According to this assumption, the early posterior positivity can be viewed as a "reversed EPN", induced by lower levels of attention for visual face processing. Indeed, selective attention to one modality while ignoring simultaneous information in another modality can modulate the activity in the to-be-ignored modality. An fMRI-study on sensory processing during bimodal selective attention using audiovisual stimuli (Johnson and Zatorre, 2005) indirectly supports the idea of a reversed EPN due to attention shifting to the auditory modality: the authors observed a deactivation of auditory cortex when the participant's attention was directed to visual stimuli, even though auditory stimuli were presented simultaneously. Analogous to visual emotion processing, auditory emotional stimuli have been shown to enhance auditory cortex activation (Plichta et al., 2011). Therefore, cross-modal suppression in a similar way as described above may also be present for auditory effects on visual processing, resulting in an enhanced posterior positivity for emotional words.

In the present study we have also found an unexpected but interesting reversal of a negativity bias during isolated word processing to a positivity bias in social-communicative situations. While in the isolated condition of the pre-study negative words tended to induce stronger emotion effects, the identical stimulus set presented in the presence of a speaker showed a positivity bias with stronger emotion effects for positive words. Thus, in addition to augmented and earlier emotion effects, we also found a marked qualitative difference between communicative and non-communicative situations: The negativity bias observed for isolated words turns into a positivity bias in the communicative situation, even though identical word sets were presented. Herbert and colleagues (e.g. Herbert et al., 2011b) discuss the occurrence of a positivity bias in studies employing self-descriptive stimuli like trait adjectives or explicit self-referential tasks. In such studies, a positivity bias could be mediated by the match between positive stimuli and the positive self-concept that healthy individuals typically show. As our communicative situation can be viewed as inducing a self-referential situation, such a mechanism could be the source of the positivity bias observed here. Thus, our findings suggest that meaningful contexts might be crucial factors determining the impact of positive and negative emotional words. This may lead to a better understanding of the mixed evidence on positivity and negativity biases in the literature (negativity bias: e.g. Holt et al., 2009; Kanske and Kotz, 2007; positivity bias: e.g. Bayer et al., 2012; Herbert et al., 2009; Kissler et al., 2009; Palazova et al., 2011).

Conclusions

To summarize, we have demonstrated that affective brain responses to emotional words are enhanced if the words are encountered in communicative situations: the effects are amplified, begin earlier and last longer than in different non-communicative control conditions. This is in line with recent evidence suggesting that verbal contexts and assumptions about the sender of verbal messages (e.g. human sender vs. computer) affect the processing of neutral and emotional words (e.g. Graß et al., 2014; Fields and Kuperberg, 2012; Herbert et al., 2011a,b; Schindler et al., 2014). Furthermore, communicative situations seem to be associated with a positivity bias, with enhanced affective responses for positive, relative to negative and neutral words. The present study is the first to demonstrate enhanced affective responses to emotional words that are embedded in socially meaningful contexts, emphasizing the importance of social-communicative factors in verbal emotion processing.

Acknowledgments

We thank Julia Baum, Marie Borowikow, Philipp Glage, Martin Maier and Johannes Rost for their help during data collection and Guido Kiecker for technical assistance.

Funding

This research was supported by a grant (AB 277/5) from the German Research Council to Rasha Abdel Rahman.

Conflict of interest

All the authors declare no conflict of interest.

References

Aviezer, H., Trope, Y., Todorov, A., 2012. Body cues, not facial expressions, discriminate between intense positive and negative emotions. Science 338 (6111), 1225–1229. http://dx.doi.org/10.1126/science.1224313.

Bayer, M., Sommer, W., Schacht, A., 2010. Reading emotional words within sentences: the impact of arousal and valence on event-related potentials. International Journal of Psychophysiology 78 (3), 299–307.

Bayer, M., Sommer, W., Schacht, A., 2012. P1 and beyond: functional separation of multiple emotion effects in word recognition. Psychophysiology 49 (7), 959–969. http://dx.doi.org/10.1111/j.1469-8986.2012.01381.x.

Berg, P., Scherg, M., 1991. Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG. Clin Phys Physiol Meas 12 Suppl. A, 49–54. http://dx.doi.org/10.1088/0143-0815/12/A/010.

Bockler, A., van der Wel, R.P., Welsh, T.N., 2014. Catching eyes: effects of social and non-social cues on attention capture. Psychol Sci 25 (3), 720–727. http://dx.doi.org/10.1177/0956797613516147.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry 25 (1), 49–59. http://dx.doi.org/10.1016/0005-7916(94)90063-9.

Bradley, M.M., Lang, P.J., 2000. Affective reactions to acoustic stimuli. Psychophysiology 37 (2), 204–215. http://dx.doi.org/10.1017/S0048577200990012.

Brosch, T., Grandjean, D., Sander, D., Scherer, K.R., 2009. Cross-modal emotional attention: emotional voices modulate early stages of visual processing. Journal of Cognitive Neuroscience 21 (9), 1670–1679. http://dx.doi.org/10.1162/jocn.2009.21110.

Cacioppo, J.T., Crites, S.L., Berntson, G.G., Coles, M.G., 1993. If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? Psychological Science 4 (2), 108–112. http://dx.doi.org/10.1111/j.1467-9280.1993.tb00470.x.

Citron, F.M., 2012. Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. Brain Lang 122 (3), 211–226. http://dx.doi.org/10.1016/j.bandl.2011.12.007.

Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J., 2000. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. Biol Psychol 52 (2), 95–111. http://dx.doi.org/10.1016/S0301-0511(99)00044-7.

Czigler, I., Cox, T.J., Gyimesi, K., Horváth, J., 2007. Event-related potential study to aversive auditory stimuli. Neuroscience letters 420 (3), 251–256. http://dx.doi.org/10.1016/j.neulet.2007.05.007.

de Gelder, B., Van den Stock, J., 2011. Real Faces, Real Emotions: Perceiving Facial Expressions in Naturalistic Contexts of Voices, Bodies and Scenes. Oxford University Press, USA, pp. 535–550.

de Gelder, B., Vroomen, J., 2000. The perception of emotions by ear and by eye. Cognition & Emotion 14 (3), 289–311. http://dx.doi.org/10.1080/026999300378824.

de Gelder, B., Bocker, K.B.E., Tuomainen, J., Hensen, M., Vroomen, J., 1999. The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. Neuroscience letters 260 (2), 133–136. http://dx.doi.org/10.1016/S0304-3940(98)00963-X.

Ellsworth, P.C., Scherer, K.R., 2003. Appraisal processes in emotion. Handbook of affective sciences 572, V595.

Fields, E.C., Kuperberg, G.R., 2012. It's All About You: an ERP study of emotion and self-relevance in discourse. Neuroimage 62 (1), 562–574. http://dx.doi.org/10.1016/j.neuroimage.2012.05.003.

Fischler, I., Bradley, M., 2006. Event-related potential studies of language and emotion: words, phrases, and task effects. Progress in brain research 156, 185–203. http://dx.doi.org/10.1016/S0079-6123(06)56009-1.

Flaisch, T., Hacker, F., Renner, B., Schupp, H.T., 2011. Emotion and the processing of symbolic gestures: an event-related brain potential study. Soc Cogn Affect Neurosci 6 (1), 109–118. http://dx.doi.org/10.1093/scan/nsq022.

Gerdes, A.B., Wieser, M.J., Bublatzky, F., Kusay, A., Plichta, M.M., Alpers, G.W., 2013. Emotional sounds modulate early neural processing of emotional pictures. Front Psychol 4, 741. http://dx.doi.org/10.3389/fpsyg.2013.00741.

Giard, M.H., Peronnet, F., 1999. Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. Journal of Cognitive Neuroscience 11 (5), 473–490. http://dx.doi.org/10.1162/089892999563544.

Grandjean, D., Scherer, K.R., 2008. Unpacking the cognitive architecture of emotion processes. Emotion 8 (3), 341–351.

Grandjean, D., Sander, D., Scherer, K.R., 2008. Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. Conscious. Cogn. 17 (2), 484–495. http://dx.doi.org/10.1016/j.concog.2008.03.019 (June).

Graß, A., Hammerschmidt, W., Bayer, M., Schacht, A., 2014. Emotion effects in spoken and written word processing: evidence from ERPs in cross-modal comparison. Poster Presented at the European Society for Cognitive and Affective Neuroscience Conference 2014, Dortmund, Germany (May).

Hagoort, P., van Berkum, J., 2007. Beyond the sentence given. Philos Trans R Soc Lond B Biol Sci 362 (1481), 801–811. http://dx.doi.org/10.1098/rstb.2007.2089.
Heister, J., 2011. dlex-DB eine lexikalische Datenbank für die psychologische und linguistische Forschung. Psychologische Rundschau 62 (1), 10–20. http://dx.doi.org/10.1026/0033-3042/a000029.
Herbert, C., Junghofer, M., Kissler, J., 2008. Event related potentials to emotional adjectives during reading. Psychophysiology 45 (3), 487–498. http://dx.doi.org/10.1111/j.1469-8986.2007.00638.x.
Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., Kissler, J., 2009. Amygdala activation during reading of emotional adjectives—an advantage for pleasant content. Social cognitive and affective neuroscience 4 (1), 35–49. http://dx.doi.org/10.1093/scan/nsn027.
Herbert, C., Herbert, B.M., Ethofer, T., Pauli, P., 2011a. His or mine? The time course of self-other discrimination in emotion processing. Soc Neurosci 6 (3), 277–288. http://dx.doi.org/10.1080/17470919.2010.523543.
Herbert, C., Pauli, P., Herbert, B.M., 2011b. Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. Soc Cogn Affect Neurosci 6 (5), 653–661. http://dx.doi.org/10.1093/scan/nsq082.
Hinojosa, J.A., Mendez-Bertolo, C., Pozo, M.A., 2010. Looking at emotional words is not the same as reading emotional words: behavioral and neural correlates. Psychophysiology 47 (4), 748–757. http://dx.doi.org/10.1111/j.1469-8986.2010.00982.x.
Holler, J., Schubotz, L., Kelly, S., Hagoort, P., Schuetze, M., Ozyurek, A., 2014. Social eye gaze modulates processing of speech and co-speech gesture. Cognition 133 (3), 692–697. http://dx.doi.org/10.1016/j.cognition.2014.08.008.
Holt, D.J., Lynn, S.K., Kuperberg, G.R., 2009. Neurophysiological correlates of comprehending emotional meaning in context. Journal of Cognitive Neuroscience 21 (11), 2245–2262. http://dx.doi.org/10.1162/jocn.2008.21151.
Johnson, J.A., Zatorre, R.J., 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. Cerebral Cortex 15 (10), 1609–1620. http://dx.doi.org/10.1093/cercor/bhi039.
Junghöfer, M., Bradley, M.M., Elbert, T.R., Lang, P.J., 2001. Fleeting images: a new look at early emotion discrimination. Psychophysiology 38 (2), 175–178.
Kampe, K.K., Frith, C.D., Frith, U., 2003. "Hey John": signals conveying communicative intention toward the self activate brain regions associated with mentalizing "mentalizing", regardless of modality. The Journal of neuroscience 23 (12), 5258–5263.
Kanske, P., Kotz, S.A., 2007. Concreteness in emotional words: ERP evidence from a hemifield study. Brain Res 1148, 138–148. http://dx.doi.org/10.1016/j.brainres.2007.02.044.
Kissler, J., Assadollahi, R., Herbert, C., 2006. Emotional and semantic networks in visual word processing: insights from ERP studies. Understanding Emotions 156, 147–183. http://dx.doi.org/10.1016/S0079-6123(06)56008-X.
Kissler, J., Herbert, C., Peyk, P., Junghofer, M., 2007. Buzzwords: early cortical responses to emotional words during reading. Psychol Sci 18 (6), 475–480. http://dx.doi.org/10.1111/j.1467-9280.2007.01924.x.
Kissler, J., Herbert, C., Winkler, I., Junghofer, M., 2009. Emotion and attention in visual word processing: an ERP study. Biol Psychol 80 (1), 75–83. http://dx.doi.org/10.1016/j.biopsycho.2008.03.004.
Klasen, M., Chen, Y.H., Mathiak, K., 2012. Multisensory emotions: perception, combination and underlying neural processes. Rev Neurosci 23 (4), 381–392. http://dx.doi.org/10.1515/revneuro-2012-0040.
Kotz, S.A., Paulmann, S., 2007. When emotional prosody and semantics dance cheek to cheek: ERP evidence. Brain Res 1151, 107–118. http://dx.doi.org/10.1016/j.brainres.2007.03.015.
Kotz, S.A., Paulmann, S., 2011. Emotion, language, and the brain. Language and Linguistics Compass 5 (3), 108–125. http://dx.doi.org/10.1111/j.1749-818X.2010.00267.x.
Lang, P.J., 1995. The emotion probe: studies of motivation and attention. American psychologist 50 (5), 372. http://dx.doi.org/10.1037/0003-066X.50.5.372.
Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O., 1993. Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 30 (3), 261–273. http://dx.doi.org/10.1111/j.1469-8986.1993.tb03352.x.
Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1998. Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. Biological psychiatry 44 (12), 1248–1263. http://dx.doi.org/10.1016/S0006-3223(98)00275-3.
MacGregor, L.J., Pulvermuller, F., van Casteren, M., Shtyrov, Y., 2012. Ultra-rapid access to words in the brain. Nat Commun 3, 711. http://dx.doi.org/10.1038/ncomms1715.
Martin, B.A., Tremblay, K.L., Korczak, P., 2008. Speech evoked potentials: from the laboratory to the clinic. Ear and hearing 29 (3), 285–313. http://dx.doi.org/10.1097/AUD.0b013e3181662c0e.
McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. Nature 264 (5588), 746–748. http://dx.doi.org/10.1038/264746a0.
Nieuwland, M.S., Van Berkum, J.J., 2006. When peanuts fall in love: N400 evidence for the power of discourse. J Cogn Neurosci 18 (7), 1098–1111. http://dx.doi.org/10.1162/jocn.2006.18.7.1098.
Palazova, M., Mantwill, K., Sommer, W., Schacht, A., 2011. Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. Neuropsychologia 49 (9), 2766–2775. http://dx.doi.org/10.1016/j.neuropsychologia.2011.06.005.
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 Lang 105 (1), 59–69. http://dx.doi.org/10.1016/j.bandl.2007.11.005.
Paulmann, S., Pell, M.D., 2010. Contextual influences of emotional speech prosody on face processing: how much is enough? Cognitive, Affective, & Behavioral Neuroscience 10 (2), 230–242. http://dx.doi.org/10.3758/CABN.10.2.230.
Paulmann, S., Pell, M.D., 2011. Is there an advantage for recognizing multi-modal emotional stimuli? Motivation and Emotion 35 (2), 192–201. http://dx.doi.org/10.1007/s11031-011-9206-0.
Paulmann, S., Jessen, S., Kotz, S.A., 2009. Investigating the multimodal nature of human communication. Journal of Psychophysiology 23 (2), 63–76. http://dx.doi.org/10.1027/0269-8803.23.2.63.
Plichta, M.M., Gerdes, A.B.M., Alpers, G.W., Harnisch, W., Brill, S., Wieser, M.J., Fallgatter, A.J., 2011. Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (fNIRS) study. Neuroimage 55 (3), 1200–1207. http://dx.doi.org/10.1016/j.neuroimage.2011.01.011.
Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., Crommelinck, M., 2000. The time-course of intermodal binding between seeing and hearing affective information. Neuroreport 11 (6), 1329–1333. http://dx.doi.org/10.1097/00001756-200004270-00036.
Pourtois, G., Debatisse, D., Despland, P.A., de Gelder, B., 2002. Facial expressions modulate the time course of long latency auditory brain potentials. Cognitive Brain Research 14 (1), 99–105. http://dx.doi.org/10.1016/S0926-6410(02)00064-2.
Recio, G., Sommer, W., Schacht, A., 2011. Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. Brain research 1376, 66–75. http://dx.doi.org/10.1016/j.brainres.2010.12.041.
Rellecke, J., Palazova, M., Sommer, W., Schacht, A., 2011. On the automaticity of emotion processing in words and faces: event-related brain potentials evidence from a superficial task. Brain Cogn 77 (1), 23–32. http://dx.doi.org/10.1016/j.bandc.2011.07.001.
Schacht, A., Sommer, W., 2009a. Emotions in word and face processing: early and late cortical responses. Brain Cogn 69 (3), 538–550. http://dx.doi.org/10.1016/j.bandc.2008.11.005.
Schacht, A., Sommer, W., 2009b. Time course and task dependence of emotion effects in word processing. Cognitive Affective & Behavioral Neuroscience 9 (1), 28–43. http://dx.doi.org/10.3758/Cabn.9.1.28.
Scherer, K.R., Peper, M., 2001. Psychological theories of emotion and neuropsychological research. In: Gamom, G. (Ed.), Handbook of Neuropsychology, 2nd ed., Vol. 5 Elsevier, pp. 17–48.
Schindler, S., Wegrzyn, M., Steppacher, I., Kissler, J.M., 2014. It's all in your head—how anticipating evaluation affects the processing of emotional trait adjectives. Language Sciences 5, 1292. http://dx.doi.org/10.3389/fpsyg.2014.01292.
Schirmer, A., Kotz, S.A., 2003. ERP evidence for a sex-specific Stroop effect in emotional speech. Journal of Cognitive Neuroscience 15 (8), 1135–1148. http://dx.doi.org/10.1162/089892903322598102.
Schlochtermeier, L.H., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., Jacobs, A.M., 2013. Emotional picture and word processing: an FMRI study on effects of stimulus complexity. PLoS One 8 (2), e55619. http://dx.doi.org/10.1371/journal.pone.0055619.
Schupp, H.T., Junghofer, M., Weike, A.I., Hamm, A.O., 2003. Emotional facilitation of sensory processing in the visual cortex. Psychol Sci 14 (1), 7–13. http://dx.doi.org/10.1111/1467-9280.01411.
Schupp, H.T., Junghofer, M., Weike, A.I., Hamm, A.O., 2004a. The selective processing of briefly presented affective pictures: an ERP analysis. Psychophysiology 41 (3), 441–449. http://dx.doi.org/10.1111/j.1469-8986.2004.00174.x.
Schupp, H.T., Öhman, A., Junghöfer, M., Weike, A.I., Stockburger, J., Hamm, A.O., 2004b. The facilitated processing of threatening faces: an ERP analysis. Emotion 4 (2), 189–200. http://dx.doi.org/10.1037/1528-3542.4.2.189.
Schwartz, J.-L., Berthommier, F., Savariaux, C., 2004. Seeing to hear better: evidence for early audio-visual interactions in speech identification. Cognition 93 (2), B69–B78. http://dx.doi.org/10.1016/j.cognition.2004.01.006.
Stekelenburg, J.J., Vroomen, J., 2007. Neural correlates of multisensory integration of ecologically valid audiovisual events. Journal of Cognitive Neuroscience 19 (12), 1964–1973. http://dx.doi.org/10.1162/jocn.2007.19.12.1964.
Tempel, K., Kuchinke, L., Urton, K., Schlochtermeier, L.H., Kappelhoff, H., Jacobs, A.M., 2013. Effects of positive pictograms and words: an emotional word superiority effect? Journal of Neurolinguistics 26 (6), 637–648. http://dx.doi.org/10.1016/j.jneuroling.2013.05.002.
Thierry, G., Roberts, M.V., 2007. Event-related potential study of attention capture by affective sounds. Neuroreport 18 (3), 245–248. http://dx.doi.org/10.1097/WNR.0b013e328011dc95.
Trautmann, S.A., Fehr, T., Herrmann, M., 2009. Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain research 1284, 100–115. http://dx.doi.org/10.1016/j.brainres.2009.05.075.
Van Berkum, J.J., van den Brink, D., Tesink, C.M., Kos, M., Hagoort, P., 2008. The neural integration of speaker and message. J Cogn Neurosci 20 (4), 580–591. http://dx.doi.org/10.1162/jocn.2008.20054.
Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J Exp Psychol Learn Mem Cogn 25 (2), 394–417. http://dx.doi.org/10.1037/0278-7393.25.2.394.
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. http://dx.doi.org/10.1073/pnas.0408949102.
Vertegaal, R., Slagter, R., Van der Veer, G., Nijholt, A., 2001. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. Paper Presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
Wabnitz, P., Martens, U., Neuner, F., 2012. Cortical reactions to verbal abuse: event-related brain potentials reflecting the processing of socially threatening words. Neuroreport 23 (13), 774–779. http://dx.doi.org/10.1097/WNR.0b013e328356f7a6.
Wambacq, I.J., Jerger, J.F., 2004. Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. Cognitive Brain Research 20 (3), 427–437. http://dx.doi.org/10.1016/j.cogbrainres.2004.03.015.
Watson, L.A., Dritschel, B., Obonsawin, M.C., Jentzsch, I., 2007. Seeing yourself in a positive light: brain correlates of the self-positivity bias. Brain Res 1152, 106–110. http://dx.doi.org/10.1016/j.brainres.2007.03.049.