Feel between the Lines: Implied Emotion in Sentence Comprehension

Vicky Tzuyin Lai1,2, Roel M. Willems1,3, and Peter Hagoort1,3

Abstract

This study investigated the brain regions for the comprehension of implied emotion in sentences. Participants read negative sentences without negative words, for example, “The boy fell asleep and never woke up again,” and their neutral counterparts “The boy stood up and grabbed his bag.” This kind of negative sentence allows us to examine implied emotion derived at the sentence level, without associative emotion coming from word retrieval. We found that implied emotion in sentences, relative to neutral sentences, led to activation in some emotion-related areas, including the medial prefrontal cortex, the amygdala, and the insula, as well as certain language-related areas, including the inferior frontal gyrus, which has been implicated in combinatorial processing. These results suggest that the emotional network involved in implied emotion is intricately related to the network for combinatorial processing in language, supporting the view that sentence meaning is more than simply concatenating the meanings of its lexical building blocks.

INTRODUCTION

We read, and oftentimes we feel emotional when we read. How does the brain gather emotion from language during reading? There are several ways that emotion can be induced by language. One straightforward way is through word associations, including words associated with emotions (e.g., sad, happy) and words associated with emotionally significant events (e.g., assassinate). Past research has shown that these words lead to activations in a range of brain regions, including amygdala, parahippocampal and lingual gyrus, cingulate cortex, and visual areas (Moseley, Carota, Hauk, Mohr, & Pulvermuller, 2012; Herbert et al., 2009; Isenberg et al., 1999). Another way language induces emotion is less straightforward. In our daily life, we understand not only emotion explicitly stated in words but also emotion implied by a series of neutral words. For example, in the sentence “The boy fell asleep and never woke up again,” the unpleasant emotion is not explicitly stated, but the sentence triggers negative feelings—that is, the emotion is implied. This study is interested in how the brain supports the computation of implied emotion in language. Specifically, we ask two questions: First, would implied emotion engage brain regions involved in explicit, associative emotion? Second, would implied emotion influence areas supportive of language comprehension?

In what follows, we first review the regions commonly reported for emotion and then for language. Most emotion studies used pictures, and based on the pictorial modality, the regions commonly reported for emotion include the amygdala, the medial prefrontal cortex (mPFC), the anterior cingulate gyrus, and the insula (Luan Phan, Wager, Taylor, & Liberzon, 2004). Different functional roles have been proposed for these regions. Amygdala responds to emotionally arousing information, including both negative emotions such as fear and threat (Davis & Whalen, 2001) and positive emotions such as happiness (Hamann & Mao, 2002). The mPFC has been argued to play a general role for many kinds of emotions across various stimulus formats (Lane, Reiman, Ahern, Schwartz, & Davidson, 1997) and is related to the cognitive generation and regulation of affect (Teasdale et al., 1999), as well as introspective mental activities (Gusnard, Akbudak, Shulman, & Raichle, 2001). Moreover, this region and its subregions have been discussed in the research contexts of social cognition and theory of mind. In particular, the anterior region of the rostral medial frontal cortex has been implicated in the evaluation of one’s own emotion, the perception and judgment about other people, and the ability to represent another person’s psychological perspective (Amodio & Frith, 2006). The last function is also known as “mentalizing”—that is, the ability to predict someone else’s mental state based on their own belief and the emotional impact of the belief (Frith & Frith, 2006). Closely related to the mPFC/medial frontal cortex is the ACC. It has been proposed that the dorsal ACC and mPFC are involved in cognitively generated emotional responses, whereas the ventral portions of ACC and mPFC play more of a regulatory role for emotional responses (Etkin, Egner, & Kalisch, 2011). The insula is another region that is commonly observed in emotion studies.
The insula has been implicated for internally generated emotion, but not so much for externally induced emotion via pictures (Reiman et al., 1997). One proposal suggested that the insula is a somatic marker that maps our bodily states to the associated emotional experiences (Damasio, Everitt, & Bishop, 1996; but see Dunn, Dalgleish, & Lawrence, 2006). Another proposal suggested that the insula plays an important role in social emotion, such as empathy and compassion (Lamm & Singer, 2010), and serves as an interface that integrates several functional systems involved in affect, sensory-motor processing, and cognition (Chang, Yarkoni, Khaw, & Sanfey, 2013).

Words differ from pictures in that words represent a type of emotional event indirectly, whereas pictures depict the specifics, representing a token of the emotional event of interest directly. Several studies have examined emotional words, but to date there is no clear consensus as to which regions are definitely involved. Amygdala involvement has been reported in some (Strange, Henson, Friston, & Dolan, 2000; Isenberg et al., 1999), but not others (Kuchinke et al., 2005; Cato et al., 2004). Within those that found amygdala involvement, they also found that amygdala modulated the visual processing of emotionally valenced words in the occipital cortex (e.g., Tabert et al., 2001). In those that did not show amygdalar responses, some observed activation in the posterior cingulate cortex instead (Kuchinke et al., 2005; Maddock, Garrett, & Buonocore, 2003), whereas others observed activation in the left inferior frontal cortex instead (Nakic, Smith, Busis, Vythilingam, & Blair, 2006). These mixed findings may be driven by task: The more evaluative the task is, the more likely it is to find amygdala activation.

At the sentence level, only a couple of studies have examined emotion and explicit emotion. In a lesion study, Adolphs, Russell, and Tranel (1999) examined the functional role of the amygdala in a rare patient with bilateral amygdala damage. The patient was presented with sentences such as “Sally waved her hands in the air and yelled for help, as the boat was sinking.” The patient was impaired in judging how emotionally arousing these sentences were but was intact in judging whether the sentences were pleasant or unpleasant. Moll, de Oliveira-Souza, Bramati, and Grafman (2002) examined moral statements. Relevant to this study, within their nonmoral sentences, enhanced activities were observed for the unpleasant relative to the neural statements in the left amygdala, orbital frontal cortex, lingual gyri, and fusiform areas. Holt et al. (2011) examined the comprehension of emotional descriptions in schizophrenia and healthy controls. Negative relative to neutral sentences in healthy participants showed activations in the right posterior cingulate gyrus, right precuneus, and left posterior cingulate gyrus. Finally, Willems, Clevis, and Hagoort (2011) found that emotional sentences compared to nonemotional sentences showed more activation in the left insula, the left temporal pole, and the anterior part of the left inferior frontal region.

Holt et al. (2011), Moll et al. (2002), and Adolphs et al. (1999) used emotional sentences that contain emotional words (e.g., yelled, sinking), providing word associations with explicit emotion. In strong contrast and crucially, this study used emotional sentences without emotionally valenced words, aiming at investigating the neurobiological underpinnings of the computation of implied emotion. We suggest that implied emotion in sentences require “combinatorial processing.” Combinatorial processing in language refers to the process in which words are combined into multiword representations to give rise to sentential meaning beyond the meaning of single words (Hagoort, 2013; Humphries, Binder, Medler, & Liebenthal, 2007; Hagoort, 2005). Combinatorial processing is assumed to happen at multiple levels, such as sound, syntax, and meaning (Jackendoff, 2002). To test emotion generated by combinatorial processing and combinatorial processing alone, we designed negative sentences that contain no negative words (e.g., “The boy fell asleep and never woke up again”) and their neutral counterparts (“The boy stood up and grabbed his bag”). In both the negative and neutral sentences, none of the individual lexical items are associated with negative feelings, but the negative sentence as a whole triggers negative feelings, whereas the neutral sentence does not.

Most language studies examining combinatorial processing of this kind reported enhanced activations in the regions of the left inferior frontal gyrus (LIFG), the anterior and posterior temporal regions, and the angular gyrus (Brennan & Pylkäinen, 2012; Rodd, Johnsruude, & Davis, 2012; Bemis & Pylkäinen, 2011; Graves, Binder, Desai, Conant, & Seidenberg, 2010; Snijders et al., 2009; Moll et al., 2002; Hagoort et al., 2005). Activations in the LIFG have been reliably found for various types of stimuli, ranging from simple combinations such as two-word compounds (e.g., meaningful “flower girl” vs. reversed “girl flower”; Graves et al., 2010) to complex combinations in multiword utterances (e.g., sentences with ambiguous words “The shell was fired towards the tank” vs. sentences with unambiguous words “Her secrets were written in her diary”; Rodd et al., 2005). Several hypotheses have been put forth for the functional role of LIFG during language processing: The LIFG may contribute to the parsing and reparsing of syntactic information (Novick, Trueswell, & Thompson-Schill, 2005), the reinterpretation of semantic information (Rodd et al., 2012), and the integration of information that is not stored in long-term memory (Hagoort, 2005, 2013; Willems, Ozurek, & Hagoort, 2007). The LIFG can be anatomically subdivided into pars orbitalis, pars opercularis, and pars triangularis, each of which has been implicated in the subprocesses in language processing (Hagoort, 2005). Specifically, the dorsal parts of the LIFG (pars triangularis and pars opercularis) are implicated in syntactic processing, and the ventral parts of the LIFG (pars triangularis and pars orbitalis) are implicated in semantic processing (Hagoort & Indefrey, 2007; Van Berkum et al., 2006; Hagoort, 2005).
compute meaning beyond the single word level. We processing is related to combinatorial processing as both can also be computed via inferential processing. Inferential emotion or, more generally speaking, implied meaning (Emmorey, Hickok, & Pylkkänen, 2012). Moreover, implied activations in LIFG because of its role in combinatorial processing. However, most of these studies examined emotion associated with words, not emotion derived from word combinations in sentences. It is not clear if the emotion-by-simulation theories can account for implied emotion at the sentence level, which cannot come from word retrieval, but must stem from combinatorial processing in sentences. In addition, some studies (e.g., Moseley et al., 2012; Kuchinke et al., 2005) also found activations for emotional words in classic language regions such as inferior frontal gyrus (IFG), typically implicated for combinatorial language processing.

This study examines the comprehension of implied emotion in sentences. To this end, we measured hemodynamic changes while participants read negative sentences without negative words (e.g., The boy fell asleep and never woke up again), negative words (e.g., coma), and negative pictures (e.g., a picture of a man being bullied), along with their neutral counterparts. We added the conditions of words and pictures because, although our primary interest is sentences, we are also interested in the commonalities and differences between words, pictures, and sentences. We predict that implied emotion, as revealed by subtracting the neutral sentences from the negative sentences, will likely lead to activation in the mPFC, the amygdala, and the insula, because implied emotion tested here can be viewed as an internally produced, high-level cognitive representation of emotion. In addition, we may also observe enhanced activations in LIFG because of its role in combinatorial processing (Hagoort, 2005, 2013), as well as other language-relevant regions in the temporal cortex, the ventromedial pFC, and/or the angular gyrus (Friederici, 2012; Poeppel, Emmorey, Hickok, & Pylkkänen, 2012). Moreover, implied emotion or, more generally speaking, implied meaning can also be computed via inferential processing. Inferential processing is related to combinatorial processing as both compute meaning beyond the single word level. We view combinatorial processing as a necessity for language, whereas inferential processing is available when called upon by communicative needs, which is often the case during language communication. Several studies have shown that inferring the speaker meaning involves the theory of mind network including the TPJ and the mPFC (Bašnáková, Weber, Petersson, van Berkum, & Hagoort, 2014; Hagoort 2013; van Ackeren, Casasanto, Bekkering, Hagoort, & Rueschemeyer, 2012; Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006).

**METHODS**

**Participants**

Twenty-four native Dutch speakers participated in the study for payment. Eight were excluded for the following reasons: Two did not fit the language criterion of being a native speaker of Dutch, two had noncorrected vision during the experiment, and four had too much movement during scanning (>3.5 mm). The remaining 16 participants (14 women, mean age = 20.5 years, range = 18–28 years) were right-handed, as assessed by the Dutch version of the Edinburgh handedness inventory (Oldfield, 1971). All had normal or corrected-to-normal vision, and none had a history of neurological, psychiatric, or language disorder problems. None was particularly depressed or happy, as assessed by the Positive Affect Negative Affect System Trait questionnaire (Watson, 1988; positive affect = 3.59, SD = 0.34, range = 3.1–4.0, on a scale from 1 to 5; negative affect = 1.87, SD = 0.44, range = 1.3–2.7). The study fell under the legal ethical approval procedure in the Netherlands. Before the experiment, each participant signed an informed consent and was told that they could stop their participation any time.

**Materials**

The materials consisted of 140 pairs of negative/neutral sentences (Table 1), 140 pairs of negative/neutral words (Table 1), and 140 pairs of negative/neutral scenes. We constructed the sentences from scratch and selected words based on the Dutch Affective Word norm database (Moors et al., 2013) and the word frequency database SubTLex-NL (Keuleers, Brysbaert, & New, 2010). The semantic content of the words and sentences are matched, but not with the pictures. We selected pictures from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008) and from the Internet because of the lack of animate beings in the International Affective Picture System neutral images. Several norming tests were carried out to verify the manipulated content and the quality of materials detailed below.

For verifying the emotional manipulations of all materials, 41 additional participants who did not take part in the scanning session participated in the computerized pretests for sentences (13 participants), words (13 participants), and pictures (15 participants). Participants rated each stimulus on their valence, arousal, and concreteness dimensions by clicking on a 9-point Self-Assessment Manikin scale, adapted from Bradley and Lang (1994), a rating scheme typically used in emotion research. The emotional scale ranged from a frowning manikin at the negative end (=1) to a neutral-faced manikin in the middle (=5) and a smiling manikin on the right (=9). Likewise, the arousal
The scale ranged from a nonarousing manikin on one end (=1) to a more arousing manikin in the middle (=5) and a very arousing manikin on the right (=9). The concreteness scale ranged from concrete (=1) to abstract (=9).

The rating results are summarized in Table 2. Overall, negative items are indeed more negatively valenced than neutral items for sentences, words, and pictures, with similar negative-neutral difference between sentences, words, and pictures. Paired t tests indicated that, in sentences, the negative sentences were more negative than the neutral ones ($t(12) = 19.796, p < .0001$), more arousing than the neutral ones ($t(12) = 5.022, p < .0001$), and as concrete as the neutral ones ($t(12) = 1.172, p = .264$).

For words, the negative words were more negative than the neutral ones ($t(12) = 12.041, p < .0001$), more arousing than the neutral ones ($t(12) = 7.321, p < .0001$), and as concrete as the neutral ones ($t(12) = 1.664, p = .122$). In pictures, the negative scenes were more negative than the neutral ones ($t(14) = 14.749, p < .0001$), more arousing than the neutral ones ($t(14) = 8.917, p < .0001$), and as concrete as the neutral ones ($t(14) = 1.050, p = .311$).

Crucially, for sentences, we made sure that (1) the words in the negative and neutral sentences were matched in terms of the emotional characteristics and nonemotional features at the single word level, so that all emotional meaning came from combinations of words and not from

### Table 2. Pretest Ratings of Valence, Arousal, and Concreteness for Negative and Neutral Sentences, Words, and Pictures on a 9-Point Scale

<table>
<thead>
<tr>
<th>Eg. Condition</th>
<th>Sentences: Whole Sentence, Mean (SD)</th>
<th>Sentences: Content Words, Mean (SD)</th>
<th>Words, Mean (SD)</th>
<th>Pictures, Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Negative</td>
<td>Ze kwam uit het water en zag dat er een zwart beest vastgezogen zat aan haar been.</td>
<td>2.36 (0.5)</td>
<td>4.40 (0.32)</td>
<td>2.38 (0.5)</td>
</tr>
<tr>
<td></td>
<td>“She came out of the water and saw a black animal attached on her leg.”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Negative</td>
<td>De man zag een rood spoor op de vloer en zag zijn vrouw in de gang liggen.</td>
<td>5.22 (0.2)</td>
<td>4.47 (0.36)</td>
<td>5.48 (0.6)</td>
</tr>
<tr>
<td></td>
<td>“The man saw a red trace on the floor and saw his wife lying in the hallway.”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Negative</td>
<td>Hij heeft nog maar één Been en nu moet deze ook verwijderd worden.</td>
<td>5.59 (1.1)</td>
<td>4.16 (0.36)</td>
<td>5.62 (1.1)</td>
</tr>
<tr>
<td></td>
<td>“He has only one leg and now this one must also be removed.”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Negative</td>
<td>Er kropen bonderden beestjes over haar benen die hun tanden in haar vlees zetten.</td>
<td>6.41 (1.1)</td>
<td>4.18 (0.45)</td>
<td>2.71 (1.4)</td>
</tr>
<tr>
<td></td>
<td>“Hundreds of small animals were crawling on her legs and were putting their teeth into her flesh.”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Negative</td>
<td>Er reden honderden auto’s over de weg waarvan de achterste lampen op het wegdek schenen.</td>
<td>6.20 (1.3)</td>
<td>–</td>
<td>6.72 (1.3)</td>
</tr>
<tr>
<td></td>
<td>“Hundreds of cars were driving on the road of which the rear lights were shining on the road”</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Examples of Negative/Neutral Sentences/Words in Dutch with English Translations Provided in Quotes

<table>
<thead>
<tr>
<th>Eg. Condition</th>
<th>Sentences</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Negative</td>
<td>Ze kwam uit het water en zag dat er een zwart beest vastgezogen zat aan haar been.</td>
<td>bloedzuiger</td>
</tr>
<tr>
<td></td>
<td>“She came out of the water and saw a black animal attached on her leg.”</td>
<td>“leech”</td>
</tr>
<tr>
<td>2 Negative</td>
<td>De man zag een rood spoor op de vloer en zag zijn vrouw in de gang liggen.</td>
<td>bloedsporen</td>
</tr>
<tr>
<td></td>
<td>“The man saw a red trace on the floor and saw his wife lying in the hallway.”</td>
<td>“blood traces”</td>
</tr>
<tr>
<td>3 Negative</td>
<td>Hij heeft nog maar één been en nu moet deze ook verwijderd worden.</td>
<td>amputatie</td>
</tr>
<tr>
<td></td>
<td>“He has only one leg and now this one must also be removed.”</td>
<td>“amputation”</td>
</tr>
<tr>
<td>4 Negative</td>
<td>Er kropen bonderden beestjes over haar benen die hun tanden in haar vlees zetten.</td>
<td>insectenbeet</td>
</tr>
<tr>
<td></td>
<td>“Hundreds of small animals were crawling on her legs and were putting their teeth into her flesh.”</td>
<td>“insect bite”</td>
</tr>
<tr>
<td>5 Negative</td>
<td>Er reden honderden auto’s over de weg waarvan de achterste lampen op het wegdek schenen.</td>
<td>achterlichten</td>
</tr>
<tr>
<td></td>
<td>“Hundreds of cars were driving on the road of which the rear lights were shining on the road”</td>
<td>“tail lights”</td>
</tr>
</tbody>
</table>
words themselves; (2) the negative and neutral sentences
were matched in terms of their nonemotional implied
meanings, so that the inferential processing was matched
between conditions; and (3) the negative and neutral
sentences differed in terms of their emotional implied
meanings.

Regarding (1), we matched the mean valence, arousal,
and word frequencies of the content words of the sen-
tences between conditions. We ensured this by first finding
the valence and arousal values of all of the content words in
the sentences based on the Dutch Affective Word norm
database (Moors et al., 2013), averaging the values of the
content words of a given sentence and matching the aver-
ages between the conditions. In this norm, the arousal ranges from 1 to 7 (1 = very negative, 7 = very positive;
1 = very calm and 7 = very arousing). The averaged valence ratings were 4.4 (SD = 0.5, range = 3.5–4.9) for the
content words in the negative sentences and 4.5 (SD =
0.4, range = 3.4–5.8) in the neutral sentences (p = .998).
The averaged arousal ratings were 4.2 (SD = 0.4, range =
3.2–4.8) for the content words in the negative sentences
and 4.2 (SD = 0.5, range = 3.1–5.1) for the content words
in the neutral sentences (p = .926). We consulted Keuleers
et al. (2010) for word frequency values. The averaged log
frequencies of the content words were 2.3 (SD = 0.4) for
the negative sentences and 2.2 (SD = 0.4) for the neutral
sentences (p = .389).

Regarding (2), the nonemotional implied meaning and
animacy were assessed by two native Dutch-speaking
research assistants. One of the assistants first wrote down
for each sentence what the nonemotional implied mean-
ing was. For example, for a sentence such as “The window
was open and the paper on the desk fell to the floor,” the
research assistant wrote down “There was wind,” as the
nonemotional implied meaning. The other then rated
whether the generated nonemotional implied meaning
was emotional (=1) or nonemotional (=0). The averaged
ratings were 0.02 (SD = 0.15) for the negative and 0.08
(SD = 0.3) for the neutral sentences, with the difference
being nonsignificant (p = .09).3 For animacy, both assis-
tants judged whether there was anything animate in a
given sentence (1 = yes, 0 = no). The averaged scores
for animacy were 1.0 (SD = 0.2) for the negative and 0.9
(SD = 0.2) for the neutral sentences (p = .741). The
assistants also discussed and made sure that the syntactic
structures are matched. The sentence length was mea-
ured both in terms of the number of letters and words.
The averaged numbers of letters were 78.4 (SD = 15.2) for
the negative and 80.6 (SD = 17.4) for the neutral sentences
(p = .424), and the averaged numbers of words were 14.9
(SD = 3.1) for the negative and 14.7 (SD = 3.0) for the
neutral sentences.

Regarding (3), two research assistants generated one
sentence that they considered to represent the implied
meaning for each sentence. Then, 15 participants rated
whether those generated sentences were negative on a
1–7 scale (1 = negative, 4 = neutral, 7 = positive).

On average, implied meanings generated from the nega-
tive sentences were rated 1.80 (SD = 0.31) and implied
meanings generated from the neutral sentences were
rated 4.53 (SD = 0.59). This difference was statistically
significant (p < .0001), confirming our manipulation.

Additionally, 20 nonsense sentences (e.g., The patch
will gladly reveal against a watery involved magic),
20 nonwords, and 20 pictures with a black dot somewhere
in the picture were created as the catch trials.

Procedure

Participants were debriefed and signed an informed con-
sent form before entering the scanner room and lying
down in the scanner. Each participant did three runs
(Figure 1), with the anatomical scan and breaks in be-
 tween runs. In Run 1, 70 negative sentences, 70 neutral
sentences, and 20 nonsense sentences were presented in
randomized order via the Presentation software (Neuro-
behavioral Systems, Inc., www.nbs.com). In each trial, a
fixation sign “+” was presented for 500 msec to signal
the upcoming sentence. The presentation duration of a
given sentence was length dependent (the number of
words × 500 msec) if the number of words in a sentence
was ≤8. If the number of words was >8, the sentence
was presented for 4000 msec. Between trials, a black
screen was presented with the intertrial interval varying
from 2.5 sec to 5.5 sec (mean = 4 sec within a participant)
to effectively jitter trial onset with respect to volume
acquisition (Dale, 1999).

In Run 2, 70 negative words, 70 neutral words, and
20 nonwords were presented in random order. The trial
structure was the same as Run 1 except for the presenta-
tion duration of the stimulus word. If the word length
was >8 letters, the word was presented for (the number
of letters × 50) msec. If the length was >8 letters, the
word was presented for 1500 msec. In Run 3, 70 negative
pictures, 70 neutral pictures, 10 negative pictures with
a black dot, and 10 neutral pictures with a black dot were
presented in random order. The trial structure was the
same as Runs 1 and 2, except for the presentation dura-
tion of the stimulus picture (2.5 sec).

Participants were instructed to read/view the sentences/
words/pictures attentively and only press a button when
the sentence did not make sense, when the word was
not a real word, and when a black dot appeared some-
where in the picture. Participants had four practice trials
before the run, with materials not used in the actual
experiment. After the scanning, a short exit interview
was conducted to verify if the experiment went well, for
example, double-checking if the participants could see
the sentences well.

fMRI Data Acquisition

EPIs covering the whole brain were acquired with an eight-
channel head coil on a Siemens MR system with a magnetic

3
field strength of 1.5 T (repetition time = 2340 msec; echo time = 35 msec; flip angle = 90°). Each volume consisted of 32 transversal slices. The voxel size was 3.3 mm × 3.3 mm × 3.5 mm. Additionally, a high-resolution T1-weighted anatomical scan of the head was taken (repetition time = 2250 msec; echo time = 2.95 msec; 1 mm isocubic voxel size).

fMRI Data Analysis
Data were analyzed with Statistical Parametric Mapping 8 (www.fil.ion.ucl.ac.uk/spm/software/spm8/). Preprocessing involved realignment through rigid body registration to correct for head motion and correction for differences in slice timing acquisition to the onset of the first slice. Subsequently, for each participant the anatomical scan was normalized to Montreal Neurological Institute (MNI) space, and normalization parameters were applied to all functional (EPI) images. Functional data were interpolated to 2 mm × 2 mm × 2 mm voxel size and spatially smoothed with an isotropic Gaussian kernel of 8 mm FWHM. First-level analysis involved estimation of beta weights in a multiple regression analysis with regressors describing the expected hemodynamic responses (Friston et al., 1995) for each of the conditions (Sentence_negative, sentence_neutral, word_negative, word_neutral, picture_negative, picture_neutral). Stimuli were modeled as their actual duration, and regressors were convolved with a canonical two gamma hemodynamic response function (e.g., Friston, Holmes, Poline, Price, & Frith, 1996). The six motion parameters obtained from the motion correction algorithm (three translations and three rotations) were included as regressors of no interest. Second-level group analysis involved testing a mixed model with participants as a random factor (“random effects analysis”; Friston, Holmes, Price, Buchel, & Worsley, 1999). Group statistical maps were corrected for multiple comparisons by combining an activation level threshold of $p < .001$ at the individual participant level with a cluster extent threshold computed using the theory of Gaussian random fields to arrive at a statistical threshold with a $p < .05$ significance level, corrected for multiple comparisons (Poline, Worsley, Evans, & Friston, 1997; Friston et al., 1996).

ROI Analysis
To understand implied emotion and how it may be related to combinatorial processing and/or motor simulation, we assessed responses from the IFG including its three subdivisions, pars opercularis, pars triangularis, and pars orbitalis, and motor regions including motor cortex and premotor cortex (Eickhoff et al., 2005). We created these ROIs by using the automated anatomical labeling template based on a landmark-guided parcellation of the MNI template (Tzourio-Mazoyer et al., 2002).

We additionally carried out ROI analyses for typical emotion regions, including amygdala, mPFC, and insula. The ROIs for amygdala and mPFC were anatomically defined. As for the ROI for the anterior insula, because the insula comprises a very large part of the cortex and different subparts of the insula are known to be differently involved in cognitive processing, this ROI was created on the basis of an extensive meta-analysis of neuroimaging studies investigating emotional processing (Kober et al., 2008). The
mean coordinates of the maxima described in the meta-
analysis were taken for the left and right insula separately
(Kober et al., 2008; Table 3). The left insula coordinates
\([x, y, z]\) are \([-28.6, -22, -40.24, -6, -34.12, -10, -40, 10, -20]\), and the right insula coordinates are \([42.24, -8, 40.4, -14, -44.16, -2]\). The ROIs were spheres with the
mean coordinates for each hemisphere as center (L: \([-36.13, -15, R: 42.15, -8]\)), with a 10-mm radius.

**RESULTS**

**Behavioral Results**

The accuracies for the catch trials were 99.1% \((SD = 2.0\%)\) for sentences in Run 1, 100% \((SD = 0)\) for words in Run 2, and 99.1% \((SD = 3.8\%)\) for pictures in Run 3.

**Whole-brain Analysis Results**

**Sentences**

Comparing negative versus neutral sentences revealed
increased activation in a large cluster stretching from the
left insular cortex to the hippocampus including the amyg-
dala (Figure 2; Table 3). A similar but less extended cluster of activation was also observed in the right hemisphere. Bilateral activation of the middle temporal gyrus, the

| Table 3. Whole-brain (Corrected): Negative versus Neutral Sentences (Significance Level \(p < .05\), Corrected) |
|----------------------------------|---|---|---|---|
| Sentences | \(x\) | \(y\) | \(z\) | \(T\) |
| **Negative > Neutral** | | | | |
| R Insula | 26 | 10 | -10 | 5.49 |
| | 28 | 12 | -22 | 3.2 |
| LR medial superior frontal gyrus | -2 | 38 | 42 | 7.12 |
| | -2 | 54 | 30 | 7.22 |
| | -14 | 46 | 24 | 6.51 |
| LR amygdala | -30 | 2 | -22 | 6.32 |
| | 24 | -12 | -8 | 5.51 |
| L insula | -30 | 18 | -4 | 3.6 |
| LR middle temporal gyrus | -54 | -30 | -8 | 4.91 |
| | 56 | -36 | 0 | 3.85 |
| R inferior frontal gyrus | 54 | 26 | -2 | 5.66 |
| L precuneus | -8 | -60 | 30 | 5.26 |
| **Neutral > Negative** | | | | |
| R superior frontal sulcus | 28 | 2 | 62 | 4.64 |

**Figure 2.** Activations for the negative > neutral contrast in sentences (yellow), words (red), and pictures (blue) and their overlap (green). All activations are corrected for multiple comparisons as described in the Methods section \((p < .05)\).
medial superior frontal gyrus (mPFC), the left precuneus, and the right inferior frontal gyrus (RIFG) were observed. The opposite contrast (neutral > negative) only revealed one cluster of activation in the right superior frontal sulcus.

Words
Comparing negative versus neutral words revealed one cluster of activation in the left medial superior frontal gyrus (mPFC; Figure 2; Table 4). The opposite contrast (neutral > negative) did not reveal any activation.

Pictures
Comparing negative versus neutral pictures led to widespread bilateral activation in the limbic system, including the amygdalae, stretching into the temporal poles (Figure 2; Table 5). Another cluster of activation was found in the bilateral fusiform gyrus, stretching into the inferior and posterior middle temporal cortices. Additionally, clusters of activation were observed in both LIFG and RIFG, with the cluster on the right being much larger than that on the left. The opposite contrast (neutral > negative) showed a set of brain regions some of which have been implicated as part of the default mode network, including the bilateral superior parietal cortex, bilateral posterior insula, posterior cingulated cortex, and bilateral anterior middle frontal sulci.

Conjunction Analysis Results
We next performed a conjunction analysis, testing for regions involved in negative > neutral in all three modalities (i.e., Sentence negative > Sentence neutral ∩ Picture negative > Picture neutral ∩ Word negative > Word neutral). We tested for the so-called conjunction null (Nichols, Brett, Andersson, Wager, & Poline, 2005), in which each statistical map is thresholded at \( p < .05 \) corrected, using the correction procedure described above. This revealed one area of activation in the medial superior frontal gyrus (mPFC; MNI peak voxel \(-6 49 32\); Figure 3).

Table 4. Whole-brain (Corrected): Negative versus Neutral Words (Significance Level \( p < .05 \), Corrected)

<table>
<thead>
<tr>
<th>Words</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
<th>( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative &gt; Neutral</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. medial superior frontal gyrus</td>
<td>(-8)</td>
<td>(50)</td>
<td>(32)</td>
<td>(4.47)</td>
</tr>
<tr>
<td><strong>Neutral &gt; Negative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5. Whole-brain (Corrected): Negative versus Neutral Pictures (Significance Level \( p < .05 \), Corrected)

<table>
<thead>
<tr>
<th>Pictures</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
<th>( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative &gt; Neutral</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR amygdalae, stretching into the temporal poles</td>
<td>(-22)</td>
<td>(-6)</td>
<td>(-16)</td>
<td>(6)</td>
</tr>
<tr>
<td></td>
<td>(-35)</td>
<td>(0)</td>
<td>(-28)</td>
<td>(3.8)</td>
</tr>
<tr>
<td></td>
<td>(20)</td>
<td>(-4)</td>
<td>(-16)</td>
<td>(7.14)</td>
</tr>
<tr>
<td></td>
<td>(44)</td>
<td>(16)</td>
<td>(-32)</td>
<td>(5.87)</td>
</tr>
<tr>
<td>LR inferior frontal gyrus</td>
<td>(52)</td>
<td>(28)</td>
<td>(0)</td>
<td>(4.88)</td>
</tr>
<tr>
<td></td>
<td>(50)</td>
<td>(14)</td>
<td>(28)</td>
<td>(4.35)</td>
</tr>
<tr>
<td></td>
<td>(-42)</td>
<td>(28)</td>
<td>(8)</td>
<td>(3.92)</td>
</tr>
<tr>
<td>LR fusiform gyrus</td>
<td>(-44)</td>
<td>(-48)</td>
<td>(-16)</td>
<td>(6.58)</td>
</tr>
<tr>
<td></td>
<td>(44)</td>
<td>(-44)</td>
<td>(-14)</td>
<td>(7.11)</td>
</tr>
<tr>
<td>LR inferior temporal sulcus</td>
<td>(-42)</td>
<td>(-74)</td>
<td>(-4)</td>
<td>(4.64)</td>
</tr>
<tr>
<td></td>
<td>(40)</td>
<td>(-72)</td>
<td>(-10)</td>
<td>(5.64)</td>
</tr>
<tr>
<td>LR posterior middle temporal gyrus</td>
<td>(46)</td>
<td>(-68)</td>
<td>(8)</td>
<td>(5.1)</td>
</tr>
<tr>
<td></td>
<td>(-44)</td>
<td>(-70)</td>
<td>(8)</td>
<td>(4.59)</td>
</tr>
<tr>
<td>L. medial superior frontal gyrus</td>
<td>(-4)</td>
<td>(56)</td>
<td>(32)</td>
<td>(6.3)</td>
</tr>
<tr>
<td><strong>Neutral &gt; Negative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR posterior insula</td>
<td>(-43)</td>
<td>(-21)</td>
<td>(12)</td>
<td>(4.5)</td>
</tr>
<tr>
<td></td>
<td>(50)</td>
<td>(-56)</td>
<td>(46)</td>
<td>(8.3)</td>
</tr>
<tr>
<td>LR superior parietal sulcus</td>
<td>(-48)</td>
<td>(-56)</td>
<td>(46)</td>
<td>(6.55)</td>
</tr>
<tr>
<td></td>
<td>(50)</td>
<td>(-54)</td>
<td>(46)</td>
<td>(9.21)</td>
</tr>
<tr>
<td>Posterior cingulated cortex</td>
<td>(-4)</td>
<td>(-77)</td>
<td>(46)</td>
<td>(5.35)</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(-34)</td>
<td>(34)</td>
<td>(7.41)</td>
</tr>
<tr>
<td>LR middle frontal sulcus</td>
<td>(30)</td>
<td>(56)</td>
<td>(6)</td>
<td>(6.57)</td>
</tr>
<tr>
<td></td>
<td>(-30)</td>
<td>(50)</td>
<td>(8)</td>
<td>(5.69)</td>
</tr>
</tbody>
</table>

ROI Analysis Results
Inferior Frontal Cortex
We examined inferior frontal cortex and its subdivisions of pars opercularis, pars triangularis, and pars orbitalis, bilaterally, for sentences and for words. The results are summarized in Figure 4. Mean activation level indicated by beta weights was extracted and entered into a repeated-measures ANOVA of 2 Conditions (negative, neutral) × 2 Hemispheres (left, right) × 3 Locations (pars opercularis, pars triangularis, pars orbitalis). In sentences, there was a Condition × Location interaction \([F(2, 30) = 5.001, p = .026]\) with no hemisphere interaction \((F < 1)\). Negative sentences showed more activation than neutral sentences in pars orbitalis \([F(1, 15) = 16.413, p = .001]\) and pars triangularis \([F(1, 15) = 9.445, p = .008]\), but not in pars...
opercularis $[F(1, 15) = 1.87, p = .192]$. For words, there was no significant effect.

**Motor Areas**

We examined motor areas including motor cortex (BA 4) and premotor cortex (BA 6). Repeated-measures ANOVAs of 2 Condition × 2 Hemisphere revealed no Condition × Hemisphere interaction or any main effect of condition, neither in motor ($F < 1$) nor in premotor areas ($F < 1$), for both sentences and words.

**Emotion Network**

We carried out focused analyses for sentences and words in the following ROIs: amygdala, mPFC, and insula. The results are summarized in Figure 5. In amygdala, for sentences, a repeated-measures ANOVA of 2 Condition × 2 Hemisphere revealed a main effect of condition $[F(1, 15) = 16.12, p < .001]$. The amygdala was more activated for negative than neutral sentences in both hemispheres $[\text{left: } F(1, 15) = 18.45, p < .001; \text{right: } F(1, 15) = 8.37, p < .01]$. For words, there was no significant effect.

In mPFC (Figure 5, middle) for sentences, there was a Condition × Hemisphere interaction $[F(1, 15) = 17.73, p = .001]$. Negative sentences led to more activations than neutral sentences in both hemispheres $[\text{left: } F(1, 15) = 67.55, p < .0001; \text{right: } F(1, 15) = 33.31, p < .0001]$, and such difference was greater in the right than in the left hemisphere. For words, there was also a Condition × Hemisphere interaction $[F(1, 15) = 10.42, p = .006]$. The negative > neutral difference in the right mPFC was larger than that in the left. Within each hemisphere, there was, however, no main effect of condition $[\text{left: } F(1, 15) = 1.55, p = .23; \text{right: } F < 1]$.

For the insula (Figure 5, bottom), we obtained a Condition × Hemisphere interaction $[F(1, 15) = 6.45, p = .02]$. Pairwise comparisons indicate that the activation level was higher for negative than for neutral sentences, but only in the left hemisphere $[F(1, 15) = 14.71, p = .002]$. For words, no significant effect was found.
DISCUSSION

This study is one of the first studies that investigated the brain regions for the comprehension of implied emotion in sentences. Participants read negative sentences without negative words, for example, “The boy fell asleep and never woke up again,” and their neutral counterparts “The boy stood up and grabbed his bag.” This kind of negative sentence allows us to examine implied emotion derived at the sentence level, without associative emotion coming from word retrieval. We found that implied emotion led to activations in emotion-related areas including the amygdala, insula, and mPFC. In addition, we found that implied emotion also led to increased activation in language-related areas including the IFG. No activation in the motor areas was observed.

Regions for Implied Emotion in Sentences

The most pronounced activation in sentences is within the mPFC, which is also the activation overlap across the three stimulus formats examined in the current study. There are at least three functional interpretations. First, the mPFC has been implicated in general emotion processing (Lane et al., 1997). The finding that all three stimulus formats have led to activations in the mPFC is consistent with this proposal. Adding to this literature, we showed that the mPFC is involved not only in associative emotion in words and pictures but also in implied emotion in sentences. Within the mPFC, the activation for the implied emotion in our negative sentences was more dorsal than ventral. This is in line with past studies reporting larger signal changes in the dorsal than the ventral mPFC when participants interpreted the same stimuli (faces) as being negatively valenced as opposed to being positively valenced (Kim, Somerville, Johnstone, Alexander, & Whalen, 2003).

Second, the mPFC has been implicated in the cognitive aspect of emotion processing. Implied emotion in sentences requires combinations of words and therefore is likely to be more cognitively costly than retrieving emotion explicitly associated with words, as reflected by the greater mPFC activation in sentences than in words. With regard to the subdivision of the mPFC for mentalizing, it has been shown that the ventral mPFC is more involved in the evaluative judgments about self, whereas the dorsal mPFC is more involved in mentalizing about others (Denny, Kober, Wager, & Ochsner, 2012). This fits our observation of the dorsal mPFC activation in sentences, as our sentences contain no first person pronouns and are always about the sufferings of others. Future studies can manipulate the first/third person pronouns in sentences with implied emotion to further verify this interpretation.

Third, a number of language studies, the mPFC has been implicated in inferential processes such as causal inferencing (Ferstl, 2010; Kuperberg et al., 2006; Ferstl & von Cramon, 2002), indirect speech (van Ackeren et al., 2012), and conversational implicatures (Bašnáková et al., 2014). Thus, the observed activation in the mPFC could be driven by general inferential processing. This interpretation would suggest that more inferential processing was needed for the condition with implied meaning than the one without, rendering the output of the inferential
processing (emotion, in this case) not crucial as a function of the mPFC. Although this is possible, it is unlikely, as the mPFC is found in the vast majority of emotion studies. In addition, in the current study, the amount of nonemotional inferential processing was matched between conditions (cf. Materials). A more likely scenario is that the implied meaning was jointly output via inferential processing in mPFC in addition to the core language combinatorial processing in areas such as LIFG. After all, successful communication requires both coded meaning and speaker meaning.

A second prominent activation is found in the insula for implied emotion in sentences, but not in words or pictures. The insula is important for social emotion such as compassion and empathy (Lamm & Singer, 2010). The insula activation we observed may reflect our readers’ compassionate/empathizing feelings about the scenarios described in the sentences. This latter interpretation is consistent with Willems et al. (2011), who found activation in the left anterior insula for fearful relative to neutral captions without scenes. They suggested that their participants “feel along” with the implied meaning of the fearful captions and that the insula is involved when explicit attention is drawn to emotion (Singer, Critchley, & Preuschoff, 2009).

A third prominent activation for implied emotion in sentences is found in the amygdala. This is not surprising given that our negative sentences were rated more arousing than the neutral ones and that past research has shown that amygdala responds to the arousing dimension of the stimuli (cf. Adolphs et al., 1999). An interesting next question is whether implied emotion could be more arousing than sentences that convey emotion explicitly. This is an empirical question, but we suspect that this is probably not necessarily the case, at least not based on comparing the present data with the one study that found amygdalar responses for sentences with explicit emotion (Moll et al., 2002; Holt et al., 2011).

Regions for Associative Emotion in Words and Pictures

For words, the only area more activated for the negative relative to neutral words is the mPFC, which has been shown before in a number of studies on emotional words (Herbert et al., 2009; Cato et al., 2004; Tabert et al., 2001). Relating to the mentalizing function in the mPFC, our results indicate that reading negative words requires more cognitive processing of emotion than neutral words. However, no activation in amygdala is observed, although the valence/arousal difference between the negative and neutral words has been matched to be the same as that in the sentences (Table 2). The lack of amygdala activation may be due to the nonevaluative nature of our passive reading task, which would be consistent with the literature that so far only studies that employed an evaluative task on words found amygdalar responses (cf. Introduction). However, this cannot account for the amygdala activation that we observed for the emotional sentences, because they were read passively as well. We contend that emotional language can lead to activation in amygdala, but only when there is sufficient context including both the speaker context (task/processing goal) and the sentential context that might make the emotional information salient and noticeable.

In pictures, negative pictures led to more activation in the bilateral amygdala and mPFC relative to neutral words, findings that are typically observed in emotion research. There was also enhanced activation in IFG and middle temporal gyrus (MTG). It is not uncommon to observe activation in IFG in imaging studies of emotional scenes (Vytal & Hamann, 2010). The IFG and MTG are also commonly implicated in language studies. It is possible that our participants made up stories for the scenarios depicted in the pictures internally, leading to some verbal processing. The latter explanation is supported by the ROI analyses, which showed that both pars triangularis and pars orbitalis are involved, which are usually regarded as nodes in the language processing network (Hagoort, 2005, 2013). It should be noted that the latter interpretation was constructed post hoc, because we had no a priori expectation that participants would engage in more verbalization in the negative as compared to the neutral condition.

Regions for Combinatorial Processing in Sentences

Implied emotion in sentences enhanced activation in the RIFG and bilateral MTG in the whole-brain analysis and bilateral IFG in the ROI analysis. These areas are known to be critical for language processing, with an especially important contribution of left and right IFG to combinatorial operations in language. Through these operations, sentence level meaning is constructed from the meanings of the individual lexical items. Crucially, in our sentence condition, the emotional content was not due to the meaning of the individual words but resulted fully from the combinatorial operations that created the overall sentence meaning, with, in our case, an emotional quality.

The coactivation of these language-relevant areas suggests that the emotion network is intricately connected with the processing in the left and right IFG. As concerns right IFG, previous studies have found that increased combinatorial processing leads to increased activation in the right IFG, especially for stimuli consisting of brief discourse (Ferstl, 2010; Hagoort, Baggio, & Willems, 2009; Menenti, Petersson, Scheerlinga, & Hagoort, 2009). Within the subdivisions of the left IFG, enhanced activations in pars orbitalis and pars triangularis were found, but not in pars opercularis. Thus, implied emotion in visual sentences recruit the pars orbitalis and pars triangularis, most likely for the semantic and the syntactic processing that they support, respectively (Hagoort,


resolution in spoken sentences. *Cerebral Cortex*, 22, 1761–1775.
This article has been cited by:

1. Rutvik H. Desai, Megan Reilly, Wessel van Dam. 2018. The multifaceted abstract brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373:1752, 20170122. [Crossref]


3. Sibylla Wolter, Carolin Dudschig, Barbara Kaup. 2016. Reading sentences describing high- or low-pitched auditory events: only pianists show evidence for a horizontal space–pitch association. *Psychological Research* 92. [Crossref]

