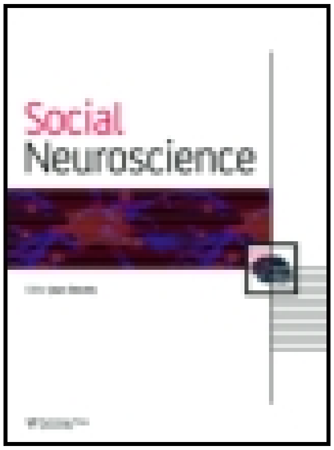

Social Neuroscience
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/psns20

Implicit agreeing/disagreeing intention while reading self-relevant sentences: A human fMRI study

Suh-Yeon Dong[a], Bo-Kyeong Kim[a] & Soo-Young Lee[a]
[a] Department of Electrical Engineering, KAIST, Daejeon, Republic of Korea
Published online: 10 Jul 2015.

SOCIAL NEUROSCIENCE, 2015
http://dx.doi.org/10.1080/17470919.2015.1059362

Implicit agreeing/disagreeing intention while reading self-relevant sentences: A human fMRI study

Suh-Yeon Dong, Bo-Kyeong Kim, and Soo-Young Lee

Department of Electrical Engineering, KAIST, Daejeon, Republic of Korea

The true intentions of humans are sometimes difficult to ascertain exclusively from explicit expressions, such as speech, gestures, or facial expressions. In this experiment, functional magnetic resonance imaging (fMRI) was used to investigate implicit intentions that were generated while a subject was reading self-relevant sentences. Short sentences, which were presented visually, consisted of self-relevant statements and a substantive verb, which indicated sentence polarity as either affirmative or negative. Each sentence was divided into the contents and the sentence ending, and the subjects were asked to respond with either agreement or disagreement after the complete sentence was presented. The overall group analysis suggested that the intention of the sentence response was found even before the reading of the complete sentences. Increased neural activation was found in the left medial prefrontal cortex (MPFC) during feelings of agreement compared to feelings of disagreement during self-relevant decision-making. In addition, according to the sentence ending, the decision of a response activated the fronto-polar cortex (FPC) in the switching condition. These findings indicated that the implicit intentions of responses to the given statements were internally generated before an explicit response occurred, and, hence, intentions can be used to predict a subject's future answer.

Keywords: Implicit intention; Functional magnetic resonance imaging; Self-relevance.

Good communication requires the ability to understand the intention of the interlocutor (Garrod & Pickering, 2009). For humans, it is natural to try to find out what is on other people's minds by observing their way of talking or their appearance. To help machines communicate with humans, human–computer interfaces have been developed. Since the 1970s, research on human–computer interactions has aimed at communicating with machines rather than at operating them (Card, Moran, & Newell, 1986). Thus, the key objective has been the development of machines that better understand human intentions. In humans, explicit expression is a straightforward way to represent one's thoughts or mindset. During the conversation, speech acts themselves play a crucial role in representing one's intentions (Malle, 2002; Smith, 1996), and the accompanying facial expressions (Ekman, Friesen, & Ellsworth, 1972; Kaliouby & Robinson, 2005) or gestures (Just & Marcel, 2009) can also be used.

Although we carefully read our partner's behavior in order to determine his mind state, it is critical to note that the reading of superficial behavior is open to both misinterpretation and deception (Dunbar, 1998). Moreover, sometimes we do not feel the need to express intentions, and sometimes it is beneficial to conceal them. In other words, intention understanding cannot always be explicit in nature. Thus, the importance of understanding implicit intention has recently been emphasized. The literature on understanding

Correspondence should be addressed to: Soo-Young Lee, Department of Electrical Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea. E-mail: sylee@kaist.ac.kr

No potential conflict of interest was reported by the authors.

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology [2013-035100], and the Asian Office of Aerospace R&D (AOARD), U.S. Air Force Research Laboratory.

implicit intentions has grown enormously, but the objectives have been limited to motion-based understanding, such as judging intention from motion patterns (Barrett, Todd, Miller, & Blythe, 2005) and action with human-like speed (Morewedge, Preston, & Wegner, 2007). Concerning motionless implicit intention, most advances since the early 1900s have examined deception, which is also called lie detection. However, emerging lie detection studies using event-related potentials (Farwell & Donchin, 1991; Rosenfeld, 2002) and functional magnetic resonance imaging (fMRI) (Davatzikos et al., 2005; Langleben et al., 2002) have been performed based on simulations of deception. These simulated lies were not self-motivated. Thus, the deception test is still controversial due to a number of limitations, including the validity of the measurements and the psychological paradigms (Wolpe, Foster, & Langleben, 2005). Our previous studies have verified that our experimental design can separately determine brain responses that are related to the pre-decision stage and the explicit decision stage by observing functional brain activation and analyzing multichannel electroencephalography (Dong & Lee, 2012, 2013a, 2013b; Dong, Kim, & Lee, 2013). The current report will provide a better understanding of implicit intentions and possibly support these previous findings. We aimed at understanding motionless implicit intention, which can apply to situations that are more general. To the best of our knowledge, no such studies have been performed to date.

The present study aimed to investigate the neural substrates of implicit intentions during the comprehension of sentences that contain self-relevant statements. Specifically, intentions of interest involve either an agreement or a disagreement with a given statement. Because there is no basic understanding of implicit intention, we therefore used brain measurements with fMRI, which has a good spatial resolution. If the sentence contents are related to personal experiences or opinions, then a feeling of agreement indicates self-relevance. We will start by reviewing previous research on self-relevance and sentence comprehension with a focus on agreement versus disagreement. We will then describe the experimental design of this study. The evidence obtained from fMRI that was performed during the intention task will then be presented.

Self-relevance

A tendency toward selective attention to self-relevant information, even in an inattentive situation, has been widely reported in human perception studies. The most famous example of this phenomenon is known as the "cocktail party effect" (Cherry, 1953), which enables an individual to focus auditory attention onto a particular stimulus while filtering out other stimuli. More recently, many researchers have suggested that evidence from event-related potentials and fMRI indicates that self-relevant information has preferential access to our perceptual systems. Electrophysiological evidence has shown that a subject's own name evokes larger N250s as well as P300s compared to other stimuli (Zhao, Wu, Zimmer, & Fu, 2011). It has also been reported that an autobiographical word or phrase evokes P300s (Gray, Ambady, Lowenthal, & Deldin, 2004). Recently, Fields and Kuperberg (2012) have shown that self-conditional sentences elicit the early components (P1, N1, and P2). Several fMRI studies have found that the medial prefrontal cortex (MPFC) is associated with self-relevance (Amodio and Frith, 2006; D'Argembeau et al., 2007; Johnson et al., 2002; Kelley et al., 2002; Ochsner et al., 2005; Oddo et al., 2010; Summerfield, Hassabis, & Maguire, 2009). Three of the studies of self-relevance included trait descriptiveness for self and others. Their judgment task for self activated the dorsal and ventral MPFC and the anterior cingulate cortex, compared to the judgment task for others. Two of the studies of self-relevance used autobiographical memory. Recollecting personal past experiences was associated with increased activity in the MPFC. One study compared yes-no decisions for a variety of statements requiring self-knowledge and semantic knowledge. Their results revealed anterior MPFC and posterior cingulate activation for all participants. These findings suggest that the MPFC might play an important role in the general processing of self-relevant information.

Sentence comprehension

Our goal was to investigate the effects of self-relevant contents on brain activity during sentence comprehension. Either the intention of agreement or disagreement can be identified by explicit responses; yes-no decisions were made by pressing a button. Johnson's research (Johnson et al., 2002) also involved yes-no decisions in response to self-relevant statements. However, their goal was not to discriminate activation according to the subjects' responses but rather to compare the activations to those to the control statements. In the current study, the subjects' responses were very important because their response was an indicator of self-relevance to the given statements. In the Korean language, the answer is affected by sentence polarity, which indicates affirmation or negation. Thus, sentence polarity was an important issue, as was the subject's response. This characteristic will be introduced in the following section. The two different intentions (agreement and disagreement) can be described by four different conditions, which includes "Yes" or "No" answers toward affirmative or negative sentences. However, there has been no research focusing on the reader's intention according to sentence polarity. Instead, as a default role for affirmative effects, a number of studies have examined the effects of negation in the brain while subjects are passively listening (Tettamanti et al., 2008) or reading a phrase, sentence, or paragraph (Carpenter, Just, Keller, Eddy, & Thulborn, 1999; Christensen, 2009); this has been mainly investigated in native speakers, with one study using bilingual subjects (Hasegawa, Carpenter, & Just, 2002).

Language characteristics

Intention to a sentence can be interpreted in different ways depending on the language. Korean is a subject-object-verb (SOV) language, in which the subject, object, and verb of a sentence always appear in the SOV order. Korean, Japanese, Hindi, Latin, and other languages also follow this order. If English followed SOV, "I book read" would be correct, which is in contrast to the correct English sentence, "I read a book." In other words, in SOV languages, sentence polarity is determined at the end of the sentence. Thus, before the polarity is decided, one's intention is directed toward the contents. After seeing the end of the sentence, the response will be determined by comparing the intention and sentence polarity. In Table 1, an example of sentence structure differences between languages is provided. For both polarities, before the verb is shown (before "does exist" and before "does not exist"), both contents are identical; hence, the reader cannot determine the polarity at this point. This is why the implicit intention is generated only toward the contents. If a reader has had a prior experience of stealing, he would agree with the given sentence part when he sees the contents, reflecting affirmative intention. If the following verb is "does exist," his response is still affirmative; otherwise, it turns negative. While an English speaker answers "Yes" when he agrees, regardless of the polarity, a Korean speaker answers "Yes" for an affirmative sentence and "No" for a negative sentence. Simply speaking, "Yes, I didn't" or "No, I did" can be correct answers in Korean.

The current study

We selected a substantive verb ("to be" or "not to be") for the stimulus sentences. By using the language characteristics, we can divide a sentence into the verb part and the remaining contents. The verb part has only the verb in an affirmative statement and the verb and negative adverb in the negative statement. We assumed that there might be two separate cognitive processes occurring while the subject is reading each part of the sentence. We analyzed the changes in neural activity for each intention state in each sentence part separately. We focused on the changes in the activation patterns according to the implicit intention state and the sentence part. In other words, we examined the details about how implicit intention was represented in the fMRI image and how the brain activation pattern changed in each intention state while the subject was reading contents and reading sentence endings.

METHODS

Subjects

Data were obtained for 19 subjects (aged 20–30 years; mean, 23.1 years; seven women). All of the subjects were undergraduate or graduate students at KAIST. They were recruited from the KAIST student community, and they voluntarily participated in the experiment. The subjects had no history of psychiatric disorders, significant physical illness, head injury, neurological disorders, and/or alcohol or drug depen-

TABLE 1
Sentence structure differences between languages

	Affirmative statement	*Negative statement*
Standard Korean	나는/물건을 훔친 적이/있다	나는/물건을 훔친 적이/없다
Translated into English	I/experience of stealing/does exist	I/experience of stealing/does not exist
Standard English	I have an experience of stealing	I have no experience of stealing

dence. Before starting the experiment, the study and fMRI experiment were explained to the participants. Then, written informed consents were obtained from all of the subjects. The study was submitted to the regular review in the KAIST institutional review board and approved. After the experiment, the subjects were asked to report any side effects they experienced due to the experiment.

Stimuli

Seventy-four stimulus sentences that were divided into two types (affirmative or negative) were given to each subject. They were all written in Korean. The common subject of all sentences ("I") was omitted, and unnecessary adverbs or adjectives were removed. Moreover, some components that implied negation, such as "any," "at all," or "even once," were removed from the contents block. Thus, the subjects could not deduce the type of sentences beforehand. Table 2 shows examples of the sentences that were used but translated into English. In the actual experiment, all of the sentences were presented in Korean.

Procedure

First, in the scanner, fMRI-compatible goggles (NordicNeuroLab Visual System, NordicNeuroLab AS, Bergen, Norway) were adjusted for each subject. Because they all had different visual abilities and different spacing between the eyes, fine-tuning was done so as not to result in unwanted load due to visual issues. The type of sentence was classified according to the sentence ending, but the order of presentation was random in order to prevent the subjects from becoming aware of which sentence type was coming next. Figure 1 shows the experimental paradigm. The subjects were initially asked to stare at a fixation cross for 4 s. A sentence contents block was presented for 4 s, and then a sentence-ending block was presented for 2 s. After a sentence was shown, an asterisk was presented for 2 s. Subjects were asked to press a two-button device if he/she had ever been in the situation described or if they agreed with the statements presented by pressing the left button or the right one. All of the answers were to be given while the asterisk was shown. The interstimulus interval was 2 s. The task length for all 74 sentences was 17 min. The subjects were asked to stay still and not move too much in the scanner.

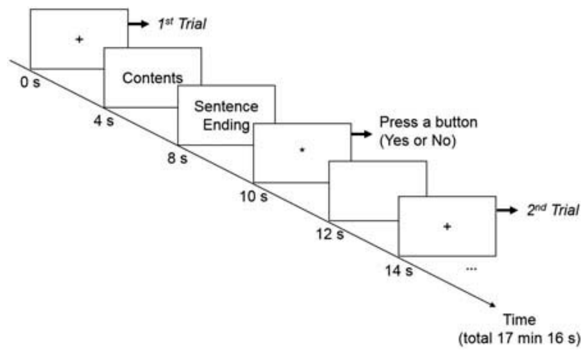

Figure 1. Experimental paradigm.

TABLE 2
Examples of sentences

Affirmative statement	*Negative statement*
Done anything dangerous for the thrill of it	Never been in trouble because of my sex behavior
Had the same dream over and over	Never worried over money and business
Heard that I sleepwalk	Never been afraid of my face becoming red
Kept from stealing something	Never been in trouble with the law
Lost sleep over worry	Never been in love with anyone
Worried about religion	Never had a peculiar experience
Had difficulty urinating	Never felt like swearing
Heard so well it bothered me	Never felt as if things were not real
Worried about my health	Never had a fainting spell

Notes: In the actual experiments, the sentences were given in Korean.

Image acquisition

The experiments were all performed at the KAIST fMRI center. Functional images were acquired with a 3 T magnetic resonance scanner (MAGNETOM® Verio, Siemens AG, Munich, Germany). Functional images were acquired with a gradient-echo echo-planar imaging sequence [36 slices; thickness, 4 mm; no gap between slices; field of view, 220 × 220 mm; matrix, 64 × 64; echo time (TE), 28 ms; repetition time (TR), 2.0 s; flip angle, 90°; voxel size, 3.4 mm × 3.4 mm × 4 mm]. A T2-weighted structural image was acquired with a three-dimensional magnetization-prepared rapid acquisition gradient-echo sequence on which the functional data was overlaid (TR, 1,800 ms; TE, 2.52 ms; matrix size, 256 × 256 × 128). Visual stimuli were displayed on a head-mounted display (NordicNeuroLab AS) that was compatible with the fMRI system. The display system had a visual angle of 30° × 23° and a resolution of 800 × 600.

Data analysis

The data were analyzed with a statistical parametric mapping software package (SPM8) (Wellcome Department of Imaging Neuroscience, University College London) running MATLAB (The MathWorks, Inc., Natick, MA, USA). The imaging data were realigned to correct for movement and normalized to the standard space defined by the Montreal Neurological Institute template. The normalized images were then spatially smoothed with a 6-mm Gaussian kernel. The experiment had two event conditions (agreement and disagreement). Together with SPM, xjView toolbox (http://www.alivelearn.net/xjview) was also used to visualize the results and identify the regions of activation. The experimental conditions were then contrasted to investigate the functional contributions of each region by using standard *t*-test analyses.

RESULTS

We focused on whether the subjects agreed or disagreed with the statements while viewing the contents block. Before the sentence end that indicated the sentence type as either positive or negative was presented, a decision of agreement or disagreement with the statements was determined in the contents block. The "Yes" or "No" answer was only dependent upon the ending block. Each sentence was classified according to the subject's answer. If the subject answered "Yes" for an affirmative sentence, it implied that the subject agreed with the contents. If the subject answered "No" for a negative sentence, it also implied agreement with the contents. Similarly, if the subject answered "Yes" for a negative sentence, it meant that the subject did not agree with the contents, and answering "No" to a positive sentence indicated disagreement. For example, let us assume that the subject sees the sentence, "I have never been to Paris." If he/she actually has been there before, the subject's answer will be "No" because of the style of responding in Korean (see *Language characteristics* in the Introduction). In this case, the explicit answer is negative toward the sentence, but his/her intention toward the contents is positive. It is the same as saying "Yes" to the sentence of "I have been to Paris." "Having been to Paris" is a common truth for the subject, and the answer can be either "Yes" or "No" depending on the sentence type. Thus, in order to determine the contrast between agreement and disagreement in the contents block, a yes response to an affirmative sentence and a no response to a negative sentence can be grouped together in the agreement condition, while a yes response to a negative sentence and a no response to an affirmative sentence are grouped in a disagreement condition.

We examined the brain regions that showed significant increases in signal intensity relative to the agreement or disagreement intention in the three kinds of contrasts: contents versus fixation for each intention, agreement versus disagreement in the contents block, and agreement versus disagreement in the ending block.

Comparison of blood-oxygen-level-dependent (BOLD) responses in different experimental blocks

To examine the effects of implicit intention in the current experiment, statistical inferences of the images were conducted as a group. For each intention, the images that were taken during the contents block were thresholded for inferences compared to those taken during the fixation block. Several clusters of cerebral areas survived with a threshold of $p < .05$ that was corrected for family-wise error (FWE) in multiple comparisons with a minimum cluster size of 10 voxels. Table 3 shows the results of the group activations that were associated with the contents versus the fixation in both the agreement and disagreement conditions. Table 3 provides the peak voxel information in each cluster, which are given as Montreal Neurological Institute coordinates and as anatomical labels according to Automated Anatomic Labeling. In Table 3, the clusters are listed in the order of peak intensity for each condition. The inferences were thresholded at an FWE-corrected p-value $< .05$, and they were excluded if the anatomical label was undefined or from the cerebellum.

Even if the intention condition was different, the explicit action was exactly the same during the contents block. Therefore, it was obvious that there were common activation regions for the contents versus the fixation contrasts for both intention conditions. Figure 2 shows the common regions, which were around the calcarine in the occipital lobe and, especially, the lingual gyrus. The lingual gyrus is known for visual processing, especially visual information that is related to letters. Other studies have implied that the lingual gyrus is involved in word processing by showing reading-specific activation in this region

TABLE 3
Regions of activation during the contents versus fixation blocks for each intention condition

Peak anatomical label	*Number of voxels*	*Peak intensity (T)*	*Peak MNI coordinate* x	y	z
Contents > Fixation in Agreement condition					
L Extranuclear	22	9.14	−12	−2	12
R Inferior parietal lobule	13	9.07	42	−52	48
L Middle frontal gyrus	18	8.84	−30	−8	64
L Lingual gyrus	34	8.83	−10	−90	−14
L Precuneus	26	8.54	−4	−78	52
L Inferior parietal lobule	21	8.36	−40	−52	48
Contents > Fixation in Disagreement condition					
R Superior frontal gyrus	33	10.1	14	16	60
R Superior frontal gyrus	11	9.74	16	16	68
L Extranuclear	27	9.51	−14	8	10
L Middle frontal gyrus	17	8.39	−28	4	56

Notes: The contrasts were thresholded at a family-wise error-corrected p-value < .05, which corresponded to t-statistics of 7.495 and 7.547 in each condition. The critical cluster size was 10 voxels. L, left; R, right; MNI, Montreal Neurological Institute.

(Buchweitz, Mason, Tomitch, & Just, 2009). Figure 3 shows the activation contrasts between the contents and fixation in the agreement condition. As shown in the figure, significant group differences were found in portions of the left extranuclear region, the left and right inferior parietal lobule, the left middle frontal gyrus, the left lingual gyrus, and the left precuneus. When the contents block showed the text, we observed increased signals on both sides of the inferior parietal lobule, which is known for its association with sentence comprehension studies (den Ouden, Fix, Parrish, & Thompson, 2009; Yokoyama et al., 2007). Intriguingly, we did not observe any significant activations in these brain regions during the disagreement condition. Two of the contrasts yielded a small significant cluster in the left middle frontal gyrus. Figure 4 shows the significant signal differences that yielded a single cluster that was found in the right superior frontal gyrus only in the disagreement condition.

Comparison of BOLD responses in different conditions

The relative differences between the two intentions were examined by observing the activation that was associated with agreement compared to that associated with disagreement. For the contrasts, we report several clusters that survived with a threshold of uncorrected p-values < .005 with a minimum cluster size of 20 voxels. No clusters survived with an FWE correction for this contrast. Figure 5 and Table 4 show the results of the group activations that were associated with agreement versus those that were associated with disagreement. We found that several regions were more active during the agreement condition, and these included the left superior frontal gyrus, both sides of the anterior cingulate, the right cingulate gyrus, the right paracentral lobule, the right supplementary motor area, and the left postcentral gyrus.

As previous self-relevant studies have shown, the activated regions mainly included the MPFC or, specifically, the left superior frontal gyrus and both sides of the anterior cingulate. One of the prominently activated regions, the peak of which was located in the left superior frontal gyrus, closely corresponded to the findings of Goldberg, Harel, and Malach (2006)'s

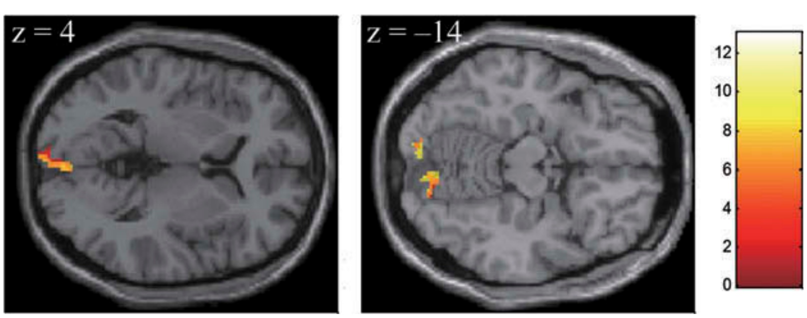

Figure 2. Regions of activation in both the agreement and disagreement conditions. Red indicates activation in the agreement condition and yellow indicates activation in the disagreement condition. The two images have overlapping activations for both conditions, which is indicated as the orange-colored region. The commonly activated region is a small part of the visual cortex in the left and interhemispheric occipital lobe (z = 4) and in both sides of the lingual gyrus (z = −14).

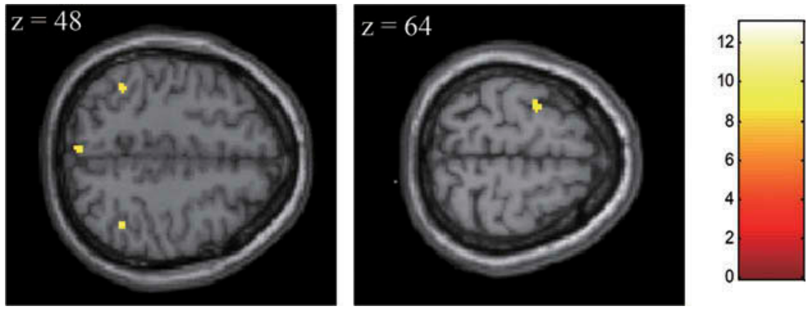

Figure 3. The functional magnetic resonance imaging (fMRI) results of the contrast (contents—fixation) in the agreement condition reveal increased activity in the inferior parietal lobule on both sides, the left precuneus (z = 48) and the left middle frontal gyrus (z = 64).

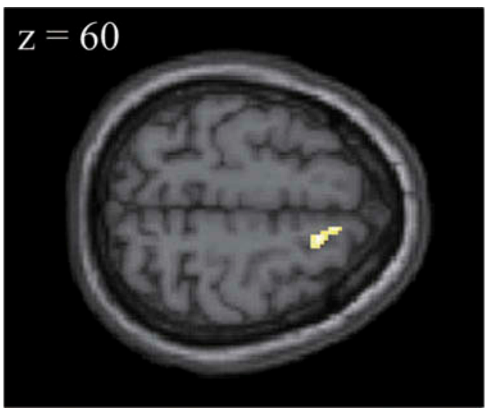

Figure 4. The fMRI results of the contrast (contents—fixation) in the disagreement condition reveal increased activity in the indicated brain area (white color, z = 60). The indicated cluster comprised 33 voxels in the right superior frontal gyrus.

self-awareness study. In his study, this region was found to mediate aspects of self-related information rather than aspects of emotional processing. In addition, relatively less significant activations were found in the right paracentral regions and in the right supplementary motor area.

In contrast, no higher signals, except in the left fusiform gyrus, were found for disagreement. The left fusiform gyrus is known to be related to written word recognition (Dehaene & Cohen, 2011; Pammer et al., 2004), but one study has shown that the left fusiform gyrus became preferentially engaged in response to unfamiliar stimuli (Szpunar, Chan, & McDermott, 2009), which can be called, as in the current study, disagreement.

In other words, the activations that were elicited by self-relevant stimuli were focused across regions in the MPFC both in the current study and in previous studies (Amodio and Frith, 2006; D'Argembeau et al., 2007; Goldberg et al., 2006; Johnson et al., 2002; Kelley et al., 2002; Ochsner et al., 2005; Oddo et al., 2010; Summerfield et al., 2009). In contrast, the disagreement condition activated only the left fusiform gyrus due to word recognition or the unfamiliar stimulus rather than in response to the implicit intention of the task. As illustrated in Figure 6, there was a clear distinction in the activated brain regions between the two conditions, and the activations that were related to cognitive functions increased only when the subjects agreed with the sentence contents (before seeing the end of the sentences).

Comparisons of BOLD responses in the ending block

In the sentence-ending block, there were only two kinds of sentence endings: affirmative or negative. Unlike the contents block, we found several clusters of brain areas that survived a threshold of $p < .001$ that was uncorrected for multiple comparisons with a minimum cluster size of 20 voxels. Table 5 shows the regions of group activation in these extracted clusters in which the peak intensities were thresholded at 3.6105. As is also described in Figure 6, the comparison between agreement versus disagreement showed regions of activation that were mostly found in the right temporal regions: the superior temporal gyrus and the middle temporal gyrus. In the human brain, these areas are part of the temporal cortex, which performs functions that are dedicated to high-level visual processing, recognition memory, auditory processing, and language (Colom et al., 2009). In other words, these regions were not associated with cognitive processes, such as decision-making and rational thinking, and this was similar to the regions of activation that were observed in the contents block.

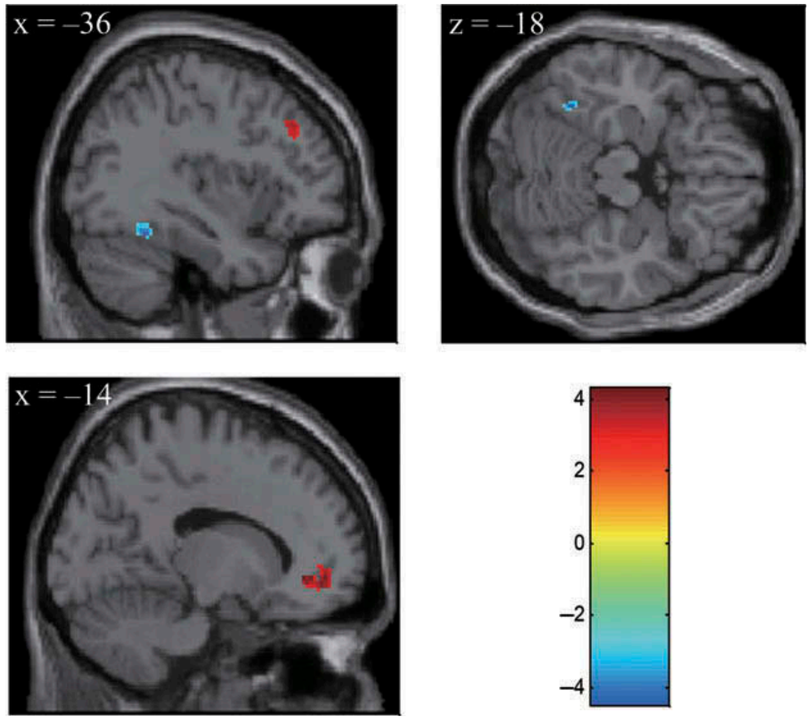

Figure 5. Regions of group activation that were specifically associated with agreement and disagreement. In the upper left panel, small regions of the left medial frontal gyrus show more activation for agreement (red) compared to those of the left fusiform gyrus that are more active for disagreement (blue) in the sagittal view ($x = -36$). In the upper right panel, activation that is associated with disagreement in the left fusiform gyrus is found in the transverse view ($z = -18$). In the lower left panel, the left anterior cingulate regions show more activation for agreement (red) in the sagittal view ($x = -14$).

TABLE 4
Group activations associated with agreement and disagreement

Peak anatomical label	*Number of voxels*	*Peak intensity (T)*	*Peak MNI coordinate* *x*	*y*	*z*
Contents in Agreement—Contents in Disagreement					
L Superior frontal gyrus	43	4.27	−38	34	36
L Anterior cingulate	105	4.19	−14	48	−6
R Anterior cingulate	30	3.82	4	40	8
R Cingulate gyrus	53	3.78	12	4	30
R Paracentral lobule	50	3.62	8	−38	76
R Supplementary motor area	36	3.58	2	−20	68
L Postcentral gyrus	35	3.34	−32	−46	70
R Paracentral lobule	24	3.25	12	−36	52
Contents in Disagreement—Contents in Agreement					
L Fusiform gyrus	28	4.41	−36	−50	−18

Notes: The contrasts were thresholded at an uncorrected p-value $< .005$, which corresponded to a t-statistic of 2.8784 and a cluster size of 20 voxels. L, left; R, right; MNI, Montreal Neurological Institute.

However, in the opposite comparison, we made a meaningful discovery of human intention. Brodmann area 10, which is the so-called frontopolar cortex (FPC), is known to be involved in human executive function. Furthermore, it has recently been well demonstrated that the FPC is associated with switching behaviors in simple decision-making tasks (Boorman, Behrens, Woolrich, & Rushworth, 2009; Christoff & Gabrieli, 2000; Koechlin & Hyafil, 2007). New insights of FPC function involving the selection of an optimal behavior among multiple behavioral alternatives have been suggested. In this experiment, the disagreement condition implied that the subject should select his explicit response by switching the intention according to the polarity of sentence ending, such as a "no" response to an affirmative ending and a "yes" response to a negative ending. The FPC might play a crucial role in the evaluation of the internally generated information, which is what we defined as implicit intention.

In addition, we observed a difference in the activation that occurred between two responses regardless of sentence polarity, as shown in the bottom two

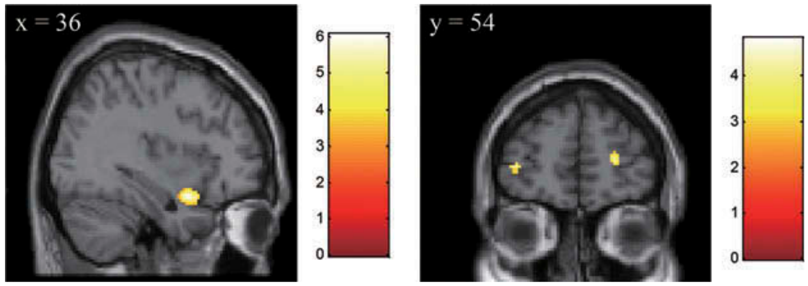

Figure 6. Regions of group activation that were associated with agreement and disagreement in the ending block. The left panel shows regions of the right temporal gyrus that were more active for agreement compared to that for disagreement in the sagittal view (x = 36). The right panel shows that both sides of the frontal gyrus (Brodmann area 10) were more active for disagreement in the coronal view (y = 54).

TABLE 5
Regions of group activation that were associated with agreement and disagreement in the sentence-ending block

Peak coordinate region	*Number of voxels*	*Peak intensity*	*Peak MNI coordinate* x	y	z
Agreement > Disagreement					
R superior temporal gyrus	122	6.0744	36	6	−18
L Subgyral	37	5.1834	−24	36	4
R Subcallosal gyrus	24	4.7444	20	16	−16
L Corpus callosum	41	4.6133	−12	32	6
R Middle temporal gyrus	105	4.5876	54	10	−18
Disagreement > Agreement					
L Inferior frontal gyrus	26	4.8291	−38	50	4
R Superior frontal gyrus	23	4.6085	24	56	8
Yes response > No response					
No suprathreshold activity					
No response > Yes response					
R Extranuclear	37	5.2103	26	−22	18
R Subgyral	21	4.8242	30	−66	24
R Lingual gyrus	30	4.2477	8	−76	−4

Notes: The contrasts were thresholded at an uncorrected p-value < .001, which corresponded to a t-statistic of 3.6105 and a cluster size of 20 voxels. L, left; R, right; MNI, Montreal Neurological Institute.

contrasts of Table 5. The contrast between a “yes” response and a “no” response had no significant activity in the ending block, and no frontal activity was found for the reverse contrast.

DISCUSSION

When decision-making is related to a personally relevant issue, we assumed that a subject’s real intention might be revealed with fMRI. We defined implicit intention as an intention that can be observed in neural activities but that is not yet explicitly expressed. The results showed that, when a subject felt that the contents of a stimulus were self-relevant, activation was found in the MPFC, which is known to be involved in rational thinking or decision-making, while only the left fusiform gyrus was activated when the contents of a stimulus were not self-relevant. However, in the ending block, we found that the patterns of functioning were very different compared to the previous block.

Compared to lie-detection studies, the most distinct difference in our study was that the subjects’ responses were regarded as all truth. Unlike lie detection, whether the subject tells the truth or a lie is not our concern. In many fMRI studies (Abe et al., 2008; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Kozel et al., 2005; Langleben et al., 2002; Phan et al., 2005), brain activity comparisons between deceptive and truthful responses have been found to be different. The anterior cingulate activation that was found by Langleben et al. (2002) and Kozel et al. (2005) was consistent with our findings. The MPFC activations that were found by Ganis et al. (2003), Phan et al. (2005), and Abe et al. (2008) were also consistent with many self-relevance studies, as described in the introduction. However, the activation in the right superior frontal gyrus that was reported by Abe et al. (2008) is not consistent with the findings of self-relevance studies and with our findings.

These findings provide an extended dimension to cross-cultural studies. This experimental design can be applied only to native speakers of SOV-structured languages. Furthermore, the expression of agreement and disagreement is different between nations. Let us imagine again that the sentence, “I have never been to Paris,” is presented on the screen. If the subject is a Korean and he has been there before, “No” is his proper answer. However, if he is an American, then he will say "Yes." In this case, agreement is no longer a combination of a yes response to an affirmative sentence or a no response to a negative sentence but instead consists of a yes response to both sentence types. Agreement and disagreement is only dependent upon the answer's polarity, regardless of the contents' polarity. When we grouped each condition according to the English standard, we did not find any regions of activation that were commonly associated with yes responses to both sentence types, while grouping according to the Korean convention revealed crucial regions of activation that were related to cognitive processes. We expect that this might also be observable in English-speaking subjects' responses. This could form the basis of further study and enable us to find cultural differences in implicit intention between Korean-speaking and English-speaking people.

For this study, we must note that we had two different dimensions for characterizing the types of implicit intention in response to the self-relevant contents. All of the self-relevance studies that have been conducted thus far have contrasted brain activations in response to self-relevant stimulation and neutral (or control) conditions. However, there should be a subtle distinction according to the stimuli, even though they were all used for self-relevant stimulation. Thus, unlike previous studies, we only used self-relevant stimulation and investigated the differences in the brain activation according to the participants' answers. A major limitation of our study compared to real settings was that all of the participants' answers were regarded as truth. Although we tried to instruct them to answer honestly, the participants probably were not willing to do so for some sentences. In other words, a lie in response to one sentence could result in unexpected brain activation. Therefore, an accurate and precise paradigm will likely have to account for the variations that arise from the different contents of sentences.

In addition, we used contact measurements in a limited environment during our experiment. There is currently no method available to measure neural activities in a real environment, such as moving situations, for tasks with a high cognitive demand. Human–computer interfaces should understand real intentions by using only simple measurements in moving situations, such as the recently developed speech recognition systems that are available in smartphones. The measuring devices would ideally be smaller and more convenient to handle. There are two possible solutions. One is to make portable devices that can measure one's neural activities. The other solution is to investigate the relationship between neural activities and easily measured outputs (e.g., speech and video) so as to predict implicit intention by using only noncontact measurements in a natural situation. With this approach, the neural substrates that we found can be used as a basis of implicit intention. Studies of implicit intention are also related to studies of mind reading, which enable the control of machines with thinking. In the not too distant future, mind reading could also be a general aspect of future machines. Implicit intention studies are the forerunner of mind-reading research. Further improvements are required to make human–computer interactions more natural.

CONCLUSION

This report is a first attempt to demonstrate that feelings of agreement and disagreement in response to self-relevant contents were associated with different patterns of brain activation. The determination of relevance was made by the subject's own will. As a result, self-relevant contents affected the reader's implicit intention generation even before reading the complete sentence, as shown by fMRI. We can conclude that implicit intentions that are related to the self-relevant information are generated in advance and can thus be used to predict how a subject is going to answer.

Original manuscript received 26 June 2014
Revised manuscript accepted 25 May 2015
First published online 10 July 2015

REFERENCES

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., … Fujii, T. (2008). Neural Correlates of true memory, false memory, and deception. *Cerebral Cortex*, *18*, 2811–2819. doi:10.1093/cercor/bhn037

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–277. doi:10.1038/nrn1884

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, *26*, 313–331. doi:10.1016/j.evolhumbehav.2004.08.015

Boorman, E. D., Behrens, T. E., Woolrich, M. W., & Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*, *62*, 733–743. doi:10.1016/j.neuron.2009.05.014

Buchweitz, A., Mason, R. A., Tomitch, L. M. B., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology and Neuroscience*, *2*, 111–123. doi:10.3922/j.psns.2009.2.003

Card, S., Moran, T., & Newell, A. (1986). *The Psychology of Human-Computer Interaction*. Oxford: Taylor & Francis.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1999). Time course of fmri-activation in language and spatial networks during sentence comprehension. *NeuroImage*, *10*, 216–224. doi:10.1006/nimg.1999.0465

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*, 975–979.

Christensen, K. R. (2009). Negative and affirmative sentences increase activation in different areas in the brain. *Journal of Neurolinguistics*, *22*, 1–17. doi:10.1016/j.jneuroling.2008.05.001

Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, *28*, 168–186.

Colom, R., Haier, R. J., Head, K., Álvarez-Linera, J., Angeles Quiroga, M., Shih, P. C., & Jung, R. E. (2009). Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence*, *37*, 124–135. doi:10.1016/j.intell.2008.07.007

D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Balteau, E., Luxen, A., … Salmon, E. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Journal of Cognitive Neuroscience*, *19*, 935–944. doi:10.1162/jocn.2007.19.6.935

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D., Acharyya, M., Loughead, J. W., … Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, *28*, 663–668. doi:10.1016/j.neuroimage.2005.08.009

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*, 254–262. doi:10.1016/j.tics.2011.04.003

den Ouden, D.-B., Fix, S., Parrish, T. B., & Thompson, C. K. (2009). Argument structure effects in action verb naming in static and dynamic conditions. *Journal of Neurolinguistics*, *22*, 196–215. doi:10.1016/j.jneuroling.2008.10.004

Dong, S. Y., & Lee, S. Y. (2012) Understanding human implicit intention based on frontal electroencephalography (EEG), *Proc. IJCNN*, pp. 1–5.

Dong, S. Y., & Lee, S. Y. (2013a) Recognition of human implicit intention based on fMRI and EEG. Presented at Neuroinformatics 2013.

Dong, S. Y., & Lee, S. Y. (2013b, February 16–20). Understanding human implicit intention while reading self-relevant sentences: An fMRI study. Presented at the 19th Annual Meeting of the OHBM, Seattle, WA.

Dong, S. Y., Kim, B. K., & Lee, S. Y. (2013). Decoding and predicting implicit agreeing/disagreeing intention based on electroencephalography (EEG). *Iconip*, *2*, 587–594.

Dunbar, R. (1998). Theory of mind and the evolution of language. In J. A. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 92–110). Cambridge: Cambridge University Press.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Amsterdam: Elsevier B.V.

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, *28*, 531–547. doi:10.1111/psyp.1991.28.issue-5

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*, *62*, 562–574. doi:10.1016/j.neuroimage.2012.05.003

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex*, *13*, 830–836. doi:10.1093/cercor/13.8.830

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*, 292–304. doi:10.1111/tops.2009.1.issue-2

Goldberg, I. I., Harel, M., & Malach, R. (2006). When the brain loses its self: Prefrontal inactivation during sensorimotor processing. *Neuron*, *50*, 329–339. doi:10.1016/j.neuron.2006.03.015

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, *40*, 216–224. doi:10.1016/S0022-1031(03)00092-1

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage*, *15*, 647–660. doi:10.1006/nimg.2001.1001

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain*, *125*, 1808–1814. doi:10.1093/brain/awf181

Just, A., & Marcel, S. (2009). A comparative study of two state-of-the-art sequence processing techniques for hand gesture recognition. *Computer Vision and Image Understanding*, *113*, 532–543. doi:10.1016/j.cviu.2008.12.001

Kaliouby, R. E., & Robinson, P. (2005). Real-time inference of complex mental states from facial expressions and head gestures. In B. Kisacanin, V. Pavlovic, & T. Huang (Eds.), *Real-Time Vision for Human-Computer Interaction* (pp. 181–200). New York, NY: Springer.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785–794. doi:10.1162/08989290260138672

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science*, *318*, 594–598. doi:10.1126/science.1142995

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, *58*, 605–613. doi:10.1016/j.biopsych.2005.07.040

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., … Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage*, *15*, 727–732. doi:10.1006/nimg.2001.1003

Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givon & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Amsterdam: Benjamins.

Morewedge, C. K., Preston, J., & Wegner, D. M. (2007). Timescale bias in the attribution of mind. *Journal of Personality and Social Psychology*, *93*, 1–11. doi:10.1037/0022-3514.93.1.1

Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D., Kihsltrom, J. F., & D'Esposito, M. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, *28*, 797–814. doi:10.1016/j.neuroimage.2005.06.069

Oddo, S., Lux, S., Weiss, P. H., Schwab, A., Welzer, H., Markowitsch, H. J., & Fink, G. R. (2010). Specific role of medial prefrontal cortex in retrieving recent autobiographical memories: An fMRI study of young female subjects. *Cortex*, *46*, 29–39. doi:10.1016/j.cortex.2008.07.003

Pammer, K., Hansen, P. C., Kringelbach, M. L., Holliday, I., Barnes, G., Hillebrand, A., … Cornelissen, P. L. (2004). Visual word recognition: The first half second. *NeuroImage*, *22*, 1819–1825. doi:10.1016/j.neuroimage.2004.05.004

Phan, K. L., Magalhaes, A., Ziemlewicz, T. J., Fitzgerald, D. A., Green, C., & Smith, W. (2005). Neural correlates of telling lies: A functional magnetic resonance imaging study at 4 Tesla. *Academic Radiology*, *12*, 164–172. doi:10.1016/j.acra.2004.11.023

Rosenfeld, J. P. (2002). Event-related potentials in detection of deception, malingering, and false memories. In M. Kleiner (Ed.), *Handbook of Polygraphy* (pp. 265–286). New York, NY: Academic Press.

Smith, P. K. (1996). Language and the evolution of mind-reading. In P. Carruthers & P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 344–354). Cambridge: Cambridge University Press.

Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Cortical midline involvement in autobiographical memory. *NeuroImage*, *44*, 1188–1200. doi:10.1016/j.neuroimage.2008.09.033

Szpunar, K. K., Chan, J. C., & McDermott, K. B. (2009). Contextual processing in episodic future thought. *Cerebral Cortex*, *19*, 1539–1548. doi:10.1093/cercor/bhn191

Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., & Moro, A. (2008). Negation in the brain: Modulating action representations. *NeuroImage*, *43*, 358–367. doi:10.1016/j.neuroimage.2008.08.004

Wolpe, P. R., Foster, K. R., & Langleben, D. D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *The American Journal of Bioethics*, *5*, 39–49. doi:10.1080/15265160590923367

Yokoyama, S., Watanabe, J., Iwata, K., Ikuta, N., Haji, T., Usui, N., … Kawashima, R. (2007). Is Broca's area involved in the processing of passive sentences? An event-related fMRI study. *Neuropsychologia*, *45*, 989–996. doi:10.1016/j.neuropsychologia.2006.09.003

Zhao, K., Wu, Q., Zimmer, H. D., & Fu, X. (2011). Electrophysiological correlates of visually processing subject's own name. *Neuroscience Letters*, *491*, 143–147. doi:10.1016/j.neulet.2011.01.025