doi:10.1093/scan/nsq094 SCAN (2012) 7, 173–183

Empathy matters: ERP evidence for inter-individual differences in social language processing

Daniëlle van den Brink,[1] Jos J.A. Van Berkum,[1,2,3] Marcel C.M. Bastiaansen,[1,2] Cathelijne M.J.Y. Tesink,[1] Miriam Kos,[1] Jan K. Buitelaar,[1] and Peter Hagoort[1,2]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, PO Box 9101, 6500 HB, Nijmegen, [2]Max Planck Institute for Psycholinguistics, PO box 310, 6500 AH, Nijmegen and [3]UiL OTS, Utrecht University, The Netherlands

When an adult claims he cannot sleep without his teddy bear, people tend to react surprised. Language interpretation is, thus, influenced by social context, such as who the speaker is. The present study reveals inter-individual differences in brain reactivity to social aspects of language. Whereas women showed brain reactivity when stereotype-based inferences about a speaker conflicted with the content of the message, men did not. This sex difference in social information processing can be explained by a specific cognitive trait, one's ability to empathize. Individuals who empathize to a greater degree revealed larger N400 effects (as well as a larger increase in γ-band power) to socially relevant information. These results indicate that individuals with high-empathizing skills are able to rapidly integrate information about the speaker with the content of the message, as they make use of voice-based inferences about the speaker to process language in a top-down manner. Alternatively, individuals with lower empathizing skills did not use information about social stereotypes in implicit sentence comprehension, but rather took a more bottom-up approach to the processing of these social pragmatic sentences.

Keywords: auditory sentence processing; stereotypes; sex differences; N400; gamma oscillations

INTRODUCTION

When a 6-year-old girl claims that she cannot sleep without her teddy bear, hardly anybody will look surprised. However, when an adult man says the same thing, this is bound to raise some eyebrows. Besides linguistic content, the voice also carries information about a person's identity relevant for communication, such as idiosyncratic features related to the gender and approximate age of the speaker (Campanella and Belin, 2007). The previous example illustrates that these context-bound aspects of language play a role in the interpretation of the spoken message. In linguistic theory, this is referred to as pragmatic aspects of language, involving the ability to attribute meaning to social cues. The present study used event-related brain potentials (ERPs) to investigate the issue of inter-individual differences in the cognitive processes that mediate the integration of social information in a linguistic context. More specifically, we were interested in inter-individual variability in the use of probabilistic inferences about the speaker while interpreting his/her message.

In language processing, a potential determinant of inter-individual variability is sex, with behavioural studies consistently reporting women, on average, to perform better on certain measures of verbal skills than men (Maccoby and Jacklin, 1974; Halpern, 1992; Herlitz *et al.*, 1997; Kimura, 1999), but see a meta-analysis of Hyde and Linn (1988). Recent neuroimaging studies provide further evidence for sex-based differences in language processing by revealing anatomical differences in the brain, as well as a stronger lateralization of language for men than for women (Shaywitz *et al.*, 1995; Pugh *et al.*, 1996; Jaeger *et al.*, 1998; Gur *et al.*, 2000; Kansaku *et al.*, 2000; Phillips *et al.*, 2001; Baxter *et al.*, 2003; Knaus *et al.*, 2004; Clements *et al.*, 2006; Hill *et al.*, 2006). In two recent studies by Daltrozzo *et al.* (2007) and Wirth *et al.* (2007) who employed the ERP technique in a lexical-semantic priming paradigm found substantiation for sex-based differences in semantic-language processing. The results showed that men and women differed in semantic processing as indicated by earlier onsets, as well as larger amplitudes of N400 effects for women as compared to men. It was suggested that these results indicated a more automated processing of semantic information in women than in men.

Specific to the current issue of inter-individual variability in pragmatic language functioning, recent work by Schirmer and colleagues also points to sex-based differences in the processing of a certain social aspect of language encapsulated in the speaker's voice: a person's affective state. In a line of studies investigating the neurophysiological correlates of vocal-emotion processing, they found differences in the way men and women process this form of social

Received 4 February 2010; Accepted 21 October 2010
Advance Access publication 8 December 2010
This work was supported by the Netherlands Organization for Scientific Research (NWO-VENI 451-04-107 to D.vdB.).

Correspondence should be addressed to Daniëlle van den Brink, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Radboud University Nijmegen, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands. E-mail: d.vandenbrink@donders.ru.nl

information. One ERP study revealed that women showed an increased mismatch negativity (MMN) to emotion compared to neutral vocalizations when presented outside the participants' attentional focus, whereas men did not (Schirmer *et al.*, 2005b). Another ERP study revealed that women, but not men, showed an enlarged N400 to words with incongruous emotional prosody when they were asked to focus on word meaning rather than emotional prosodic information (Schirmer and Kotz, 2003). In addition, an fMRI study using the same materials revealed that in women, the left inferior frontal gyrus (IFG) was more strongly activated in incongruous as compared to congruous trials. This congruence difference only appeared as a tendency for men (Schirmer *et al.*, 2004). Together, these studies confirm and extend earlier behavioural studies concerning sex-based differences favouring women in emotion perception and higher order language functions (Hall, 1978, 1984; Hampson and Kimura, 1992; Kimura, 1999; Hall *et al.*, 2000; Hall and Matsumoto, 2004).

In the present study, we use EEG to examine sex-based differences in a particular aspect of pragmatic language functioning, the ability to use voice-based inferences about the speaker. However, rather than simply focusing on sex differences *per se*, we also aim to assess to what extent such differences can be explained by a cognitive trait that on average tends to differ between men and women: the ability to empathize with another individual. The idea of considering empathy as a viable determinant of inter-individual differences in pragmatic language functioning follows from research in the realm of social cognition. The end result of social cognition has been defined as the accurate perception of the dispositions and intentions of other individuals (Brothers, 1990). Here, empathizing skills play a crucial role. In the literature, a distinction has been made between cognitive empathy (also known as mentalizing or Theory of Mind) and affective empathy (Mehrabian and Epstein, 1972; Davis, 1996; Decety and Jackson, 2004). Cognitive empathy involves the recognition and understanding of mental states of others, enabling us to predict their behaviour. Affective empathy, on the other hand, involves the experience of a similar emotion. Cognitive empathy is assumed to be of particular significance for pragmatic language functioning (Tager-Flusberg, 1993; Baron-Cohen, 1995), and as such may serve as a defining factor in inter-individual variability in the processing of social information. Studies using a wide variety of self report measures of empathy have consistently found that, on average, women score higher than men (for reviews see Eisenberg and Lennon, 1983; Davis, 1996). In addition, Hoffman (1977) found that women showed higher levels of empathy across a range of behavioural studies. Given this association of empathy with sex, sex-based differences in the processing of social information in a linguistic context may actually be the result of inter-individual differences in ability to empathize with another person.

The present study

To investigate inter-individual variation in social-language processing, we set up a study involving a particular form of social information processing conveyed by the human voice, namely the use of probabilistic inferences about the speaker while interpreting his/her message. In the current study, participants listened to sentences with conceptual messages that either did or did not match with stereotypical beliefs about the speaker, based on voice-based inferences about the speaker's age, sex or social economic status (e.g. 'I cannot sleep without my "teddy bear" in my arms' spoken by a 6-year-old boy *vs* an adult male speaker). The violations always emerged at a mid-sentence critical word, and up until the critical word the spoken sentence frames were fully compatible with voice-based assumptions about either speaker. In addition, participants heard sentences that were congruent or contained a lexical semantic violation (e.g. 'You wash your hands with "soap/horse" and water'). A previous ERP study using the same materials revealed that when listening to spoken sentences, voice-based inferences about the speaker are immediately used in utterance interpretation, in the same manner as lexical semantic information; semantic as well as pragmatic violations both elicited N400 effects (Van Berkum *et al.*, 2008). In the present study, we used the experimental materials and data of the Van Berkum *et al.* study to investigate inter-individual variability in these social pragmatic and lexical semantic N400 effects.

Consider the prominent difference between the two manipulations. In contrast to the semantic manipulation, the speaker identity incongruent items are not linguistic violations *per se*, as the sentences themselves are correct. Instead, it solely depends on the speaker whether a sentence is considered to be potentially anomalous or not. Violations are the result of a mismatch between the content of the message and stereotypical ideas about the speaker, based on the speaker's voice. In essence, the building of stereotypes helps simplify the complexity of perception by means of generalization and allows us to make predictions regarding our environment (Lee *et al.*, 1995). We hypothesized that individuals who empathize to a larger degree process social information in a top-down manner, i.e. use prior knowledge (in this case, stereotypical ideas about the speaker with regard to his/her sex, age or SES) to generate expectations about what the speaker will say, whereas individuals with low-empathizing skills process this information in a bottom-up manner i.e. do not make these predictions, but process the incoming signal and then relate it to the social information. This should result in a larger N400 effect in the pragmatic manipulation for individuals with high empathizing skills compared to those who exhibit low empathizing skills. Therefore, we explicitly investigated whether N400 effects in our study correlate with a measure for cognitive empathy, the self-reporting Empathizing Questionnaire (EQ; Baron-Cohen and Wheelwright, 2004).

In addition, since stereotypes are based on generalizations across individuals, experience with a given individual might be able to overwrite the stereotype effects. This we investigated by analysing how the speaker identity effects changed over the course of the experiment.

METHOD

Participants

The ERP experiment was conducted with 36 right-handed native speakers of Dutch, 18 males (18–26 years, mean age 20.8) and 18 females (18–35 years, mean age 21.8), 24 of which were included in the Van Berkum *et al.*'s (2008) study. None of the subjects reported having any neurological impairment, or having experienced any neurological trauma. All participants gave informed consent in writing according to the Declaration of Helsinki and were paid for their participation.

Materials

The experimental materials from a previous experiment (Van Berkum *et al.*, 2008) were used. For this experiment, 160 Dutch sentences had been constructed with a lexical content that either did or did not fit probabilistic inferences about the speaker's sex, age and socio-economic status, as could be inferred from the speaker's voice. Translated examples of speaker incongruent utterances are 'Before I leave I always check whether my "make up" is still in place', in a male voice, 'Every evening I drink some "wine" before I go to sleep' in a young child's voice and 'I have a large "tattoo" on my back' spoken in an 'upper-class' accent. In addition, participants heard sentences containing classic semantic anomalies which are pure linguistic violations matched with semantically congruent sentences (e.g. 'You wash your hands with "horse" and water' *vs* 'You wash your hands with "soap" and water'). For details, see Supplementary Data.

Procedure

After electrode application, participants were seated in a sound-attenuating booth and listened to 352 sentences, spoken by 21 different people, presented over audio speakers. Participants were asked to process each sentence for comprehension, and no additional task demands were imposed. After a short practice of 20 sentences, the trials were presented in five blocks of 10 min each, separated by rest periods. Each trial began with a fixation asterisk centred on the screen. After 1 s, the spoken sentence was played from file. The asterisk remained on the screen until 1 s after sentence offset, and was followed by a 3.6 s inter-trial interval. Participants were asked to avoid eye and other movements when the asterisk was visible, and to deliberately blink in the inter-trial interval. After the EEG experiment participants were asked to fill out Dutch translations of the Empathizing and Systemizing Questionnaires (Baron-Cohen *et al.*, 2003; Baron-Cohen and Wheelwright, 2004).

EEG recording and data analysis

EEG was recorded from 28 electrodes according to the extended International 10–20 system. Five additional electrodes were used to support in signal processing: one placed at the right mastoid, two at the outer left and right canthi, and two above and below the left eye (converted off-line to bipolar horizontal and vertical EOG signals, respectively). The recording reference was placed at the left mastoid. Electrode impedances were below 5 kOhm. Signals were recorded with a BrainAmps DC amplifier (BrainProducts, München) using a 200-Hz low-pass filter, a time constant of 10 s (0.016 Hz) and a 500-Hz sampling frequency.

After off-line re-referencing of the EEG signals to the mean of the left and right mastoid, they were filtered with a 30-Hz low-pass filter. Segments ranging from 200 ms before to 2000 ms after the acoustic onset of the critical word were baseline-corrected by subtracting mean amplitude from −200 to 0 ms pre-stimulus interval, and semi-automatically screened off-line for eye movements, muscle artefacts, electrode drifting and amplifier blocking. Segments containing such artefacts were rejected (on average 12.7%, with no asymmetry across conditions). The remaining EEG segments were averaged per participant and condition, and the associated mean amplitude values in the N400 latency range (300–600 ms) from 11 posterior electrodes were submitted to repeated-measures analyses of variance (ANOVAs) involving the within-subject factors Congruity (congruent, incongruent), and Electrode (CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2), and the between-subject factor Sex (male, female).

RESULTS

Behavioural data

In a previous rating experiment (Van Berkum *et al.*, 2008), 12 men and 12 women listened to the experimental stimuli and were asked to rate on a 5-point scale 'how normal or strange you think it is to have the speaker utter this particular sentence' (1 = completely normal, 5 = extremely odd). In the present study, we investigated possible sex-based variability in these data to assess whether at a behavioural level men and women rate the LS and SI violations differently. Table 1 reveals the results.

An ANOVA with the within-subject factors Congruity (congruent, incongruent), and Violation Type (lexical

Table 1 Off-line rating results for experimental materials

Condition	Mean rating (SD)	
	Men ($n = 12$)	Women ($n = 12$)
Lexical semantic congruent	1.4 (0.5)	1.5 (0.5)
Lexical semantic incongruent	4.6 (0.4)	4.6 (0.4)
Speaker identity congruent	1.5 (0.4)	1.6 (0.5)
Speaker identity incongruent	3.3 (0.8)	3.6 (0.9)

semantic, speaker identity) and the between-subject factor Sex (male, female) revealed a significant effect for Congruity [$F(1, 22) = 1706.90$, $P < 0.001$], and a significant effect for Violation Type [$F(1, 22) = 50.74$, $P < 0.001$], indicating that the consistent items were judged to be more normal compared to the inconsistent items, and the SI violations to be less severe than the LS violations. None of the interactions involving the factor Sex reached significance (all $P > 0.2$), indicating that at the behavioural level no sex differences were found.

ERP data

Sex differences

Figure 1 displays the grand average waveforms of the congruent and incongruent conditions of the lexical semantic (panel A) and the speaker identity (panel B) manipulations at seven posterior electrodes, time-locked to the onset of the critical word, for the male and female participants separately.

What can be seen is that, whereas the LS anomalies result in a large N400 effect for both men and women, the SI violations result in an N400 effect for women only. This effect is smaller than in the LS violations, but has a similar centro-parietal distribution as the classical lexical semantic N400 effect. No such N400 effect can be seen in men.

A repeated measures omnibus ANOVA on the mean amplitude values in the 300–600 ms latency with the within-subject factors Congruity (congruent, incongruent), Violation Type (LS, SI) and Electrode (CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2), and the between-subject factor Sex (male, female) revealed a significant main effect of Congruity [$F(1, 34) = 126.99$, $P < 0.001$], significant interactions of Congruity with Violation Type [$F(1, 34) = 73.68$, $P < 0.001$] and Congruity with Sex [$F(1, 34) = 7.94$, $P = 0.008$], in the absence of a three-way interaction between Congruity, Violation Type and Sex [$F(1, 34) < 1$, ns]. These results indicate that the size of the N400 effects differ between men and women (corresponding to an effect of 2.16 and 1.25 μV, collapsed across LS and SI), and that this sex difference in N400 effect size was present in both the LS and SI manipulations.

Adaptation effects

To test for adaptation effects across the experiment we performed a repeated measures omnibus ANOVA on the mean amplitude values in the 300–600 ms latency with the within-subject factors Congruity (congruent, incongruent), Violation Type (LS, SI), Half (first, second) and Electrode (CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, O2), and the between-subject factor Sex (male, female). A significant

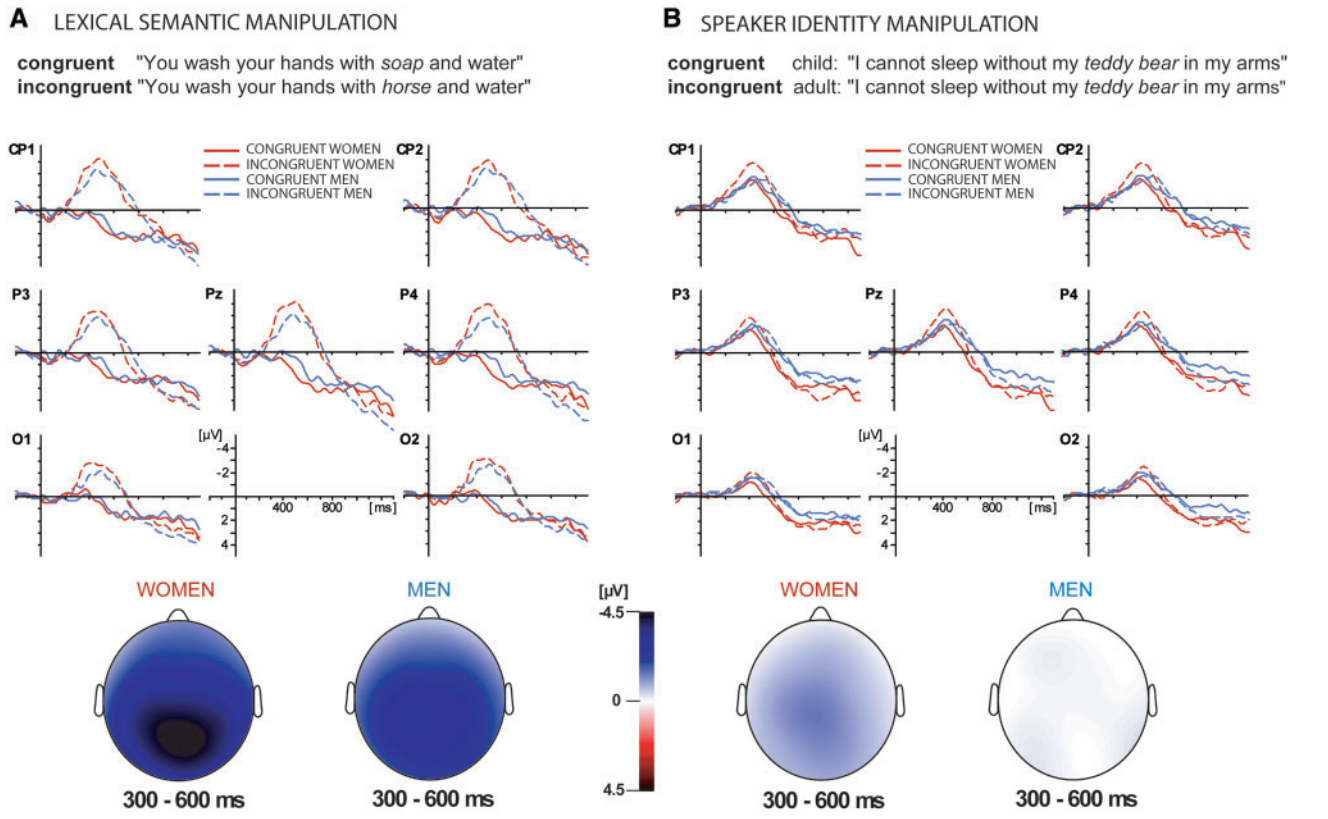

Fig. 1 ERP waveforms of men (blue) and women (red) at seven posterior sites and scalp distributions of N400 effects (incongruent minus congruent) per participant group for (**A**) Lexical Semantic manipulation and (**B**) Speaker Identity manipulation. Negative amplitudes are plotted upwards.

main effect of Congruity [$F(1, 34) = 104.52$, $P < 0.001$] was found, as well as significant interactions of Congruity with Violation Type [$F(1, 34) = 73.5$, $P < 0.001$], Congruity with Sex [$F(1, 34) = 7.18$, $P = 0.011$]. The Congruity by Half interaction did not reach significance [$F(1, 34) = 1.71$, $P = 0.2$]. However, analysis of the four-way interaction between these factors did reveal a trend [$F(1, 34) = 2.97$, $P = 0.094$].

Based on our *a priori* hypotheses that stereotypical ideas would be adjusted over the course of the experiment, we performed an additional analysis for the SI stimuli alone, where the effect is based on violations of stereotypical ideas about the speakers. Figure 2 displays the grand average waveforms of the speaker congruent and incongruent conditions for the male and female participants at seven posterior electrodes, time-locked to the onset of the critical word for the first half (panel A) and second half (panel B) of the experiment.

Figure 2 illustrates that in the first half of the experiment women, in contrast to men, show a large N400 effect to the SI stimuli, which completely disappears in the second half of the experiment. Instead, a posterior late-positive effect between 800 and 1100 ms emerges. Statistical analysis on the mean amplitude values in the 300–600 ms latency interval for SI items with the factors Congruity, Violation Type, Half and Electrode revealed a significant main effect of Congruity [$F(1, 34) = 9.06$, $P = 0.005$], as well as a significant three-way interaction between the factors Congruity, Half and Sex [$F(1, 34) = 5.74$, $P = 0.022$], indicating that the factor Half is of significance for the sex differences found.

When analysing the first and second half separately, results from the first half of the experiment revealed a significant interaction of Congruity by Sex [$F(1, 34) = 8.58$, $P = 0.006$]. Simple main effect analyses within each sex group revealed a main effect of Congruity for women, but not for men [$F(1, 17) = 17.15$, $P = 0.001$, corresponding to an effect of 1.56 μV, and $F(1, 17) < 1$, ns, corresponding to a mean difference of 0.01 μV, respectively]. Results from the second half revealed no main effects of Congruity [$F(1, 17) = 1.23$, $P = 0.275$], and no interaction of Congruity and Sex [$F(1, 34) < 1$, ns, with mean differences of 0.18 μV for women and 0.3 μV for men]. These results indicate that the Congruity by Sex interaction obtained in the whole-experiment analysis actually is the result of a (substantial) N400 effect in the female participant group in the first half of the experiment alone. In the second half of the experiment, no N400 effects were obtained for either the female or the male participant group.

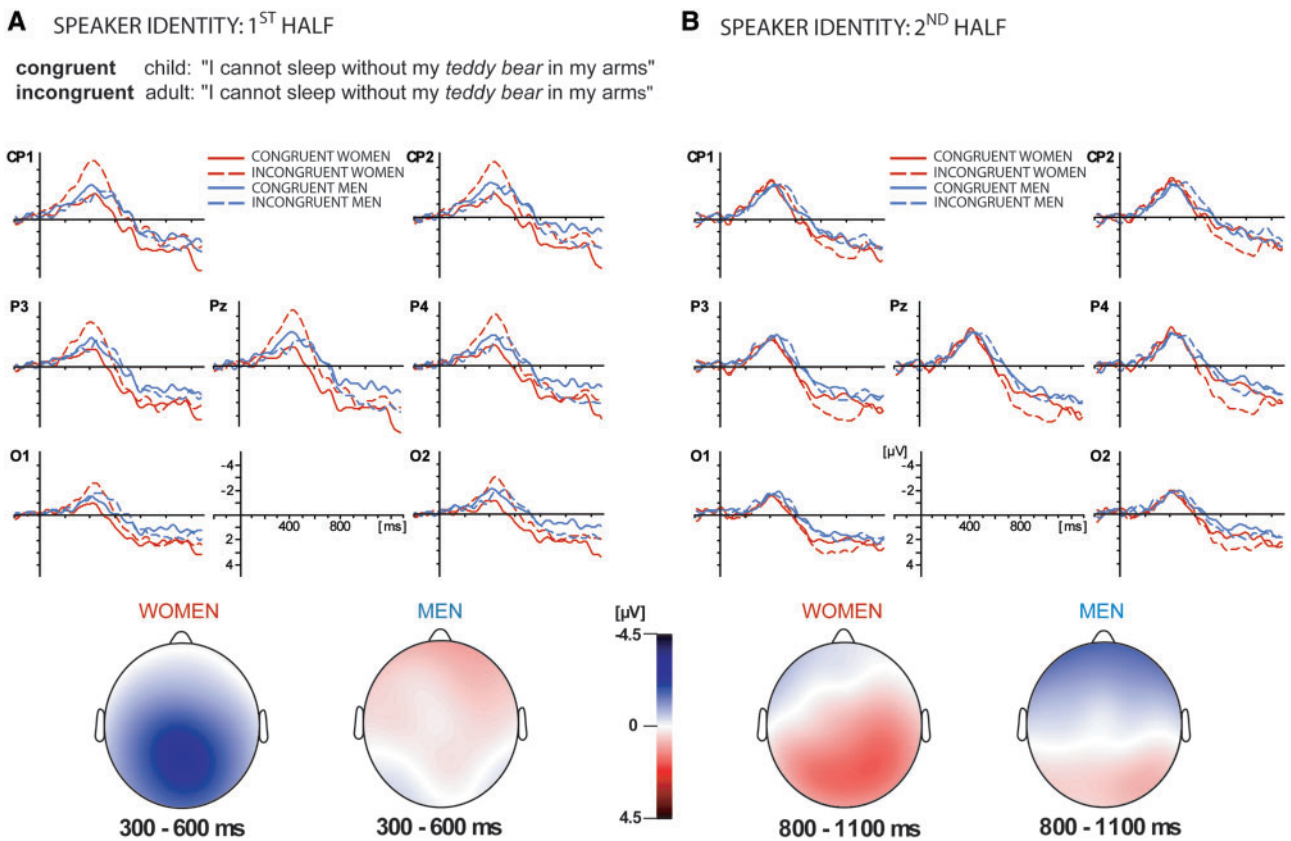

Fig. 2 ERP waveforms of men (blue) and women (red) at seven posterior sites for Speaker Identity manipulation for (**A**) first half of experiment showing scalp distributions of N400 effects (incongruent minus congruent) per participant group and (**B**) second half of experiment showing scalp distributions of Late Positive effects (incongruent minus congruent) per participant group.

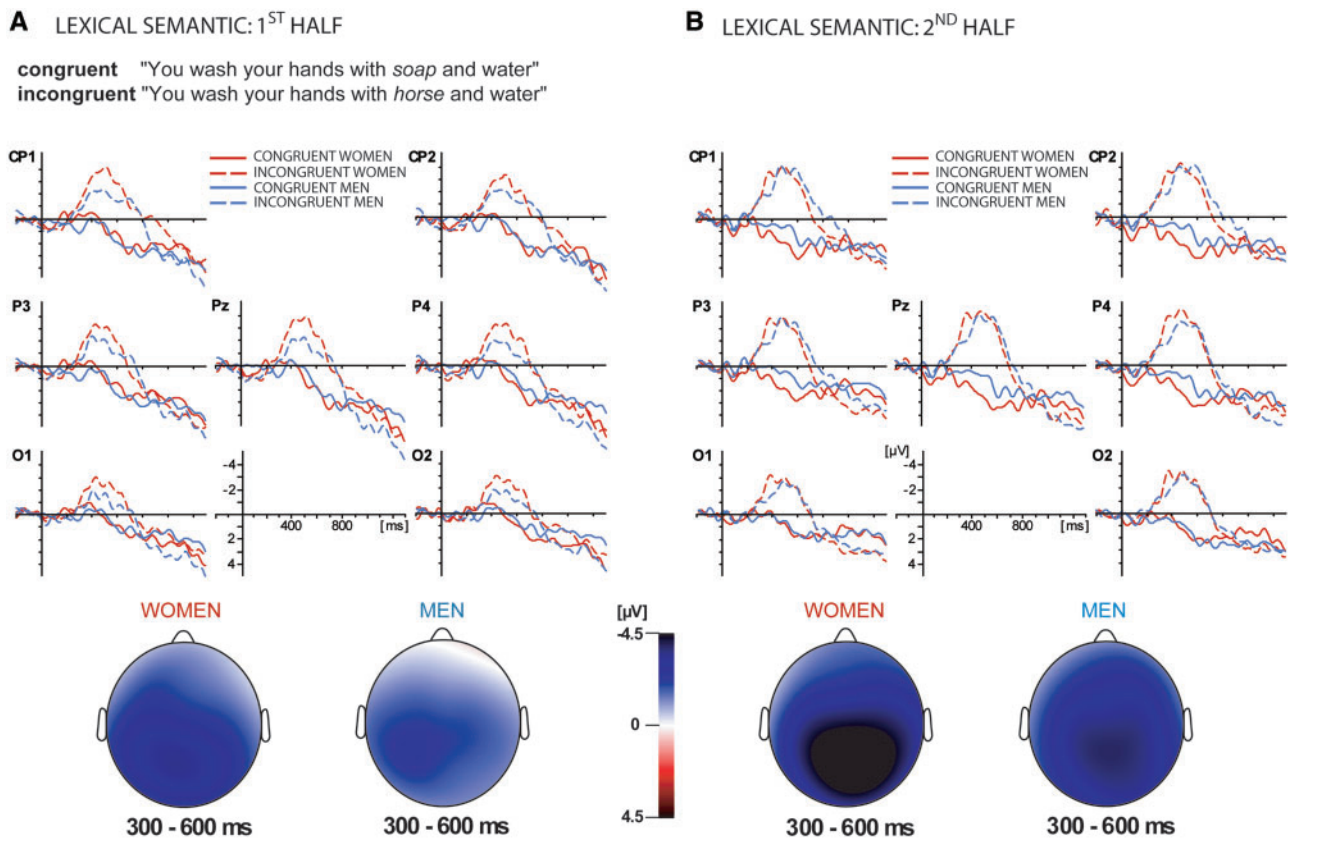

Fig. 3 ERP waveforms of men (blue) and women (red) at seven posterior sites and scalp distributions of N400 effects (incongruent minus congruent) per participant group for Lexical Semantic manipulation for (**A**) first half and (**B**) second half of experiment.

A repeated measures ANOVA on the mean amplitude values in the 800–1100 ms latency interval over the same 11 posterior electrodes for SI items with the factors Congruity, Sex and Half revealed a marginally significant three-way interaction between the factors Congruity, Half and Sex [$F(1,34) = 3.88$, $P = 0.057$], in the presence of a significant main effect of Congruity [$F(1,34) = 4.51$, $P = 0.041$]. When analysing the first and second half separately, only the second half of the experiment showed reliable ERP effects [$F(1,34) = 6.97$, $P = 0.012$], which did not differ between women and men [$F(1,34) = 2.11$, $P = 0.156$].

As the absence of an SI N400 effect in the second half of the experiment could be due to a general decline in amount of attention over the course of the experiment, we also analysed the first and second half of the experiment with respect to the LS items. Figure 3 displays the grand average waveforms of the semantically congruent and incongruent conditions for the male and female participants at seven posterior electrodes, time-locked to the onset of the critical word for the first half (panel A) and second half (panel B) of the experiment.

For the LS items, statistical analysis on the mean amplitude values in the 300–600 ms latency interval over the same 11 posterior electrodes with the factors Congruity, Sex and Half revealed a significant main effect of Congruity [$F(1,34) = 126{,}15$, $P < 0.001$], but not a three-way interaction between Congruity, Sex and Half [$F(1,34) < 1$, ns], indicating that LS violations result in significant N400 effects for both men and women in the first (with mean effects of 2.68 and 1.72 µV, respectively) and second half of the experiment (with mean effects of 4.24 and 3.07 µV, respectively). If anything, the LS violations elicited a larger N400 effect in the second half of the experiment as indicated by a significant Congruity by Half interaction [$F(1,34) = 7.63$, $P = 0.009$]. This suggests that the absence of an N400 effect for the SI items in the second half of the experiment is not due to a general decline in the amount of attention paid to the experimental stimuli, which would have resulted in smaller or absent N400 effects for the LS items in the second half of the experiment.[1]

[1]Analyses of the behavioural data, using the same order sequence of stimuli were performed. An ANOVA for SI conditions significant main effects of Con ($F(1,23) = 562.49$, $p < .001$), Half ($F(1,23) = 10.57$, $p = .004$) and a significant interaction between these two factors ($F(1,23) = 6.80$, $p = .016$), indicating that SI violations were not equally severe across the two halves of the experiment. However, the effect sizes reveal that the SI violations were actually more severe in the second half of the experiment (with effects of 1.79 and 2.01, respectively).

Analyses of the LS items revealed a significant main effect of Con ($F(1,23) = 1034.59$, $p < .001$), no effect for Half ($F(1,23) = 1.99$, $p = .171$) and no interaction between these two factors ($F(1,23) < 1$, ns), indicating, that LS violations were equally severe across the two halves of the experiment (with effects of 3.12 and 3.10, respectively). These results show that the electrophysiological differences between the two halves cannot be due to an uneven distribution of items across the two halves.

Correlations with empathy

Twenty-seven subjects (15 female) out of the 36 subjects who participated in the ERP experiment filled out a Dutch translation of the Empathizing Questionnaire (EQ; Baron-Cohen and Wheelwright, 2004). Baron-Cohen and Wheelwright have shown that consistent with other self-reporting questionnaires, on average, women score higher on the EQ than men. As a contrasting measure we also had the same 27 participants fill out another questionnaire shown to be sensitive to sex differences, the Systemizing Questionnaire (SQ; Baron-Cohen *et al.*, 2003), investigating the ability to understand and predict the workings of a system, which was assumed not to correlate with N400 effect size. Table 2 reveals the results. Concurrent with previous findings, on average, the female participants scored higher on the EQ than the male participants [$t(25) = -2.61$, $P = 0.015$], and men scored higher on the SQ than women [$t(25) = 2.86$, $P = 0.008$].

Table 2 Scores on empathizing and systemizing questionnaires

	Group	*N*	Min.	Max.	Mean (SD)
EQ score	Total	27	21	64	42.4 (12.0)
	Male	12	21	50	36.3 (9.4)
	Female	15	21	64	47.3 (11.8)
SQ score	Total	27	6	45	28.2 (9.4)
	Male	12	16	45	33.3 (8.7)
	Female	15	6	38	24.0 (7.9)

Next, we computed correlation scores for both the EQ and SQ scores in relation to the N400 effect sizes (i.e. mean amplitude difference between congruent and incongruent conditions). See Figure 4 for the corresponding scatter plots. For the LS stimuli we analysed the whole experiment as both halves elicited significant N400 effects; neither EQ nor SQ scores correlated with the LS N400 effect size [$r(25) = -0.14$, $P = 0.477$ and $r(25) = 0.04$, $P = 0.842$, respectively]. For the SI manipulation, we analysed the N400 effect size in the first half of the experiment; EQ scores correlated significantly with SI N400 effect size [$r(25) = -0.547$, $P = 0.003$]. Analysis of SQ scores did not reveal a significant correlation with SI N400 effect size [$r(25) = 0.022$, $P = 0.915$]. Individuals with high EQ scores revealed larger SI N400 effects than individuals with low EQ scores ($R^2 = 0.30$,

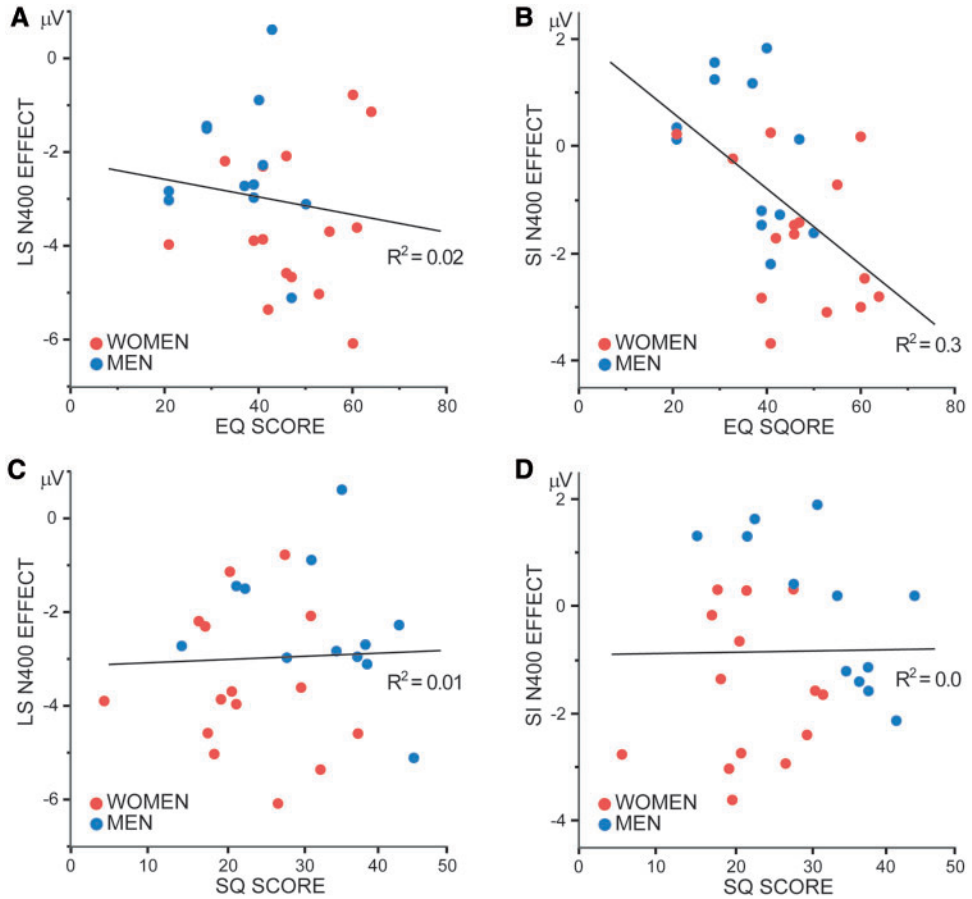

Fig. 4 Scatter diagrams showing correlations between (**A**) EQ score and mean Lexical Semantic N400 effect size; (**B**) EQ score and mean Speaker Identity N400 effect size; (**C**) SQ score and mean Lexical Semantic N400 effect size; (**D**) SQ score and mean Speaker Identity N400 effect size. Note that N400 effect is larger when values are more negative. Best fitting regression lines are also plotted.

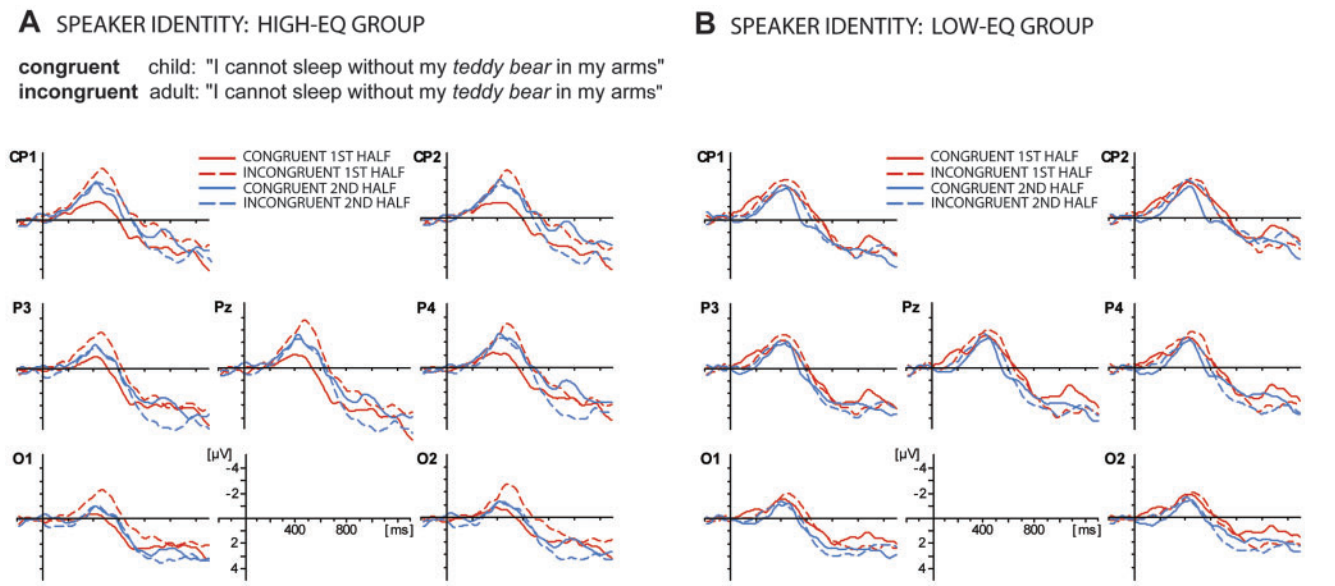

Fig. 5 ERP waveforms of Speaker Identity manipulation in first (red) and second (blue) half of the experiment for seven posterior sites for (**A**) a group of participants with high EQ scores and (**B**) a group with low EQ scores.

$P = 0.003$, SI N400 effect $= 2.03 - 0.07 \times$ EQ score). A multiple regression model (forward and backward) including Sex and EQ score revealed that EQ score was a significant predictor of the SI N400 effect ($\beta = -0.547$, $t(24) = -3.27$, $P = 0.003$), whereas Sex was not, ($\beta = 0.313$, $t(24) = 1.72$, $P = 0.098$). Based on these results, we performed a median split of the participants based on their EQ scores, resulting in a group of 14 individuals with low EQ scores (21–41, five women and nine men), and a group of 13 individuals with high EQ scores (42–64, 10 women and three men). Figure 5 reveals the SI ERP waveforms of the two EQ groups across the two halves of the experiment. What can be seen is that, whereas in the low-EQ group the waveforms for the congruent and incongruent conditions across both halves of the experiment fully overlap, the waveforms for the congruent and incongruent condition in the high-EQ group shows an N400 effect in the first half of the experiment, which is absent in the second half. Interestingly, this N400 effect in the first half of the experiment seems to be the result of a combination of a decreased N400 amplitude in the congruent condition, which significantly correlates with EQ [$r(25) = 0.391$, $P = 0.044$], and an increase in N400 amplitude to the incongruent condition (here, correlation with EQ does not reach significance, $r(25) = -0.266$, $P = 0.179$). The significant correlation of SI congruent N400 amplitude with EQ score indicates that, in individuals with higher EQ score, pragmatic processing is facilitated as a result of top-down processing.

DISCUSSION

Results from a previous ERP experiment have shown, that in sentences like 'I cannot sleep without my teddy bear in my arms', voice-based social information about who the speaker (e.g. a 6-year-old girl *vs* an adult male) is integrated into the preceding sentential context on-line, and in a similar fashion as lexical semantic information (Van Berkum *et al.*, 2008). Both social pragmatic and semantic information processing elicit a larger electrophysiological brain response, the N400 effect, in the incongruent condition relative to the congruent condition. The current study investigated two potential determinants of inter-individual differences, i.e. sex and empathy, in the processing of these particular aspects of information.

Results from an off-line behavioural experiment, where participants were asked to indicate how odd they thought a sentence was when the speaker was taken into account, indicated no sex-based differences in the processing of either pragmatic or semantic information (Table 1). In contrast, results from an on-line ERP experiment revealed that semantic information processing as well as voice-based speaker-identity processing were modulated by the listener's sex. In both types of information processing N400 effects were largest for women compared to men (Figure 1). The critical finding, however, is that although both types of information appear to be subject to sex-based variability, the sex difference in N400 sensitivity in the speaker identity manipulation can actually be reduced to individual differences in the ability to empathize with another person, whereas the sex difference in N400 sensitivity to the lexical semantic manipulation cannot (Figure 4). As indicated by a regression analysis, EQ score and not gender was found to be the sole determinant of inter-individual variability in the pragmatic N400 effect, indicating that cognitive style, rather than gender *per se* influences social language processing.

Sex differences in linguistic and verbal social information processing

Our findings of sex-based variability in language processing are in line with recent ERP findings by Daltrozzo *et al.* (2007), Wirth *et al.* (2007) and Schirmer *et al.* (Schirmer and Kotz, 2003; Schirmer et al., 2005b). Both the Daltrozzo *et al.* and the Wirth *et al.*'s study revealed that men and women differed in semantic processing as indicated by earlier and larger N400 effects for women as compared to men. The authors concluded that women process language more automatically, or more spontaneously conduct a deeper semantic analysis than men. Schirmer and Kotz (2003) revealed that only women showed an increased N400 to emotionally incongruent stimuli while performing a semantic task. They concluded that women seem to integrate emotional information from prosody more readily into ongoing semantic processing. These findings point to a higher sensitivity for women for language processing in general, whether it is semantic or social in nature.

A fundamental difference between our off-line rating experiment and the ERP experiment is that, whereas for reasons of ecological validity, we were interested in individual differences in N400 effects when language was processed implicitly, the behavioural scores of the rating study reflect explicit language processing (i.e. participants were required to judge how odd they thought the sentences were). This suggests that attention may be a crucial factor for individual differences in language processing. It appears that women exhibit a higher sensitivity in implicit language processing than men. However, in circumstances that require language to be processed explicitly, sex-based differences disappear. This proposal fits with previous findings of the influence of attention on sex-based differences in emotional speech processing (Schirmer *et al.*, 2002, 2005a). Schirmer and colleagues (2002), found that women integrated emotional prosody of a sentence (sad *vs* happy) and word valence (e.g. 'failure') earlier than men. However, when in a follow-up study participants were explicitly asked to judge whether emotional prosody matched with word meaning, sex-based differences at the electrophysiological level disappeared (Schirmer *et al.*, 2005a).

Empathy modulates top-down processing of social information processing

Interestingly, whereas in the present study the lexical semantic manipulation resulted in large N400 effects for both men and women, the speaker identity manipulation revealed a noticeable N400 effect only for the female participant group (Figure 1). In the male participant group, no ERP effect was discernable. Further analyses indicated that although both types of information appear to be subject to sex-based variability, there is another, more defining, factor influencing social information processing in particular. A regression analysis of the ERP data revealed that the individual's ability to empathize, as measured by scores on a self-reporting empathizing questionnaire, and not the listener's sex, was the sole predictor for the size of the pragmatic N400 effect. Individuals with an empathizing-driven cognitive style revealed larger N400 effects in the speaker identity contrast, indicating they are more sensitive to certain social aspects of language (Figure 4A). Since the ability to empathize, on average, differs between men and women, it is not inconceivable that Schirmer *et al.*'s findings, favouring women in emotional word processing, another form of social information processing, may also be explained by women having better empathizing skills.

These ERP results clearly indicate that there is a qualitative difference between the integration of 'semantic' and 'social' information into the linguistic context. Although both types elicit similar N400 effects, with similar onset latencies and topographical distributions, a person's ability to empathize correlates with social information processing but not lexical semantic processing (Figure 4). Note that this difference also appears to manifests itself in the oscillatory brain dynamics, where both types of information affect power changes in different frequency bands (see Supplementary Data for time–frequency analyses on the current data). Whereas lexical semantic violations elicited a theta-power increase across all participants, only individuals with an empathizing-driven cognitive style revealed a larger γ band (50–60 Hz) power increase to the speaker identity violations. Although these findings in the time-frequency domain are interesting, they merit adequate replication before conclusions can be drawn.

More evidence for qualitative differences between semantic and social information processing relates to an obtained adaptation effect in social information processing only. In contrast to the semantic manipulation, the speaker identity incongruent items were not linguistic violations, but mere violations of stereotypical ideas about the speaker, based on the speaker's voice. Figure 5 illustrates a modulation of the pragmatic N400 effect, with individuals with high-empathizing abilities eliciting a large N400 effect in the first half of the experiment, compared the second half, where the effect completely disappears. It is important to note that these results cannot be due to a general decline in attention over the course of the experiment. Statistical analysis of the lexical semantic stimuli revealed that, if anything, there was an increase in N400 effect size across the two halves of the experiment for these stimuli (Figure 3).

These results point to a relatively fast adaptation process for the pragmatic stimuli in individuals with high empathizing skills. Previous research has shown that individuals who empathize to a higher degree are better at predicting the actions of others and responding to them in appropriate ways (Saxe and Baron-Cohen, 2006). We suggest that, in the context of the present experiment, high empathizers initially used stereotype-based information to implicitly generate expectations about what the speaker might say. This resulted in a diminished N400 in sentences where the content

of message were in-line with stereotypical beliefs about the speaker and an increased N400 in cases of a mismatch between these two types of information (Figure 5). In the second half of the experiment high empathizers realized that stereotypical information did not apply to the individual speakers in this experiment, and therefore, was not a reliable cue. As a result the group of high empathizers did not use it any longer in the implicit processing of the message, resulting in N400 components of equal amplitude for both the congruent and incongruent stimuli in the second half of the experiment.

Note, that we do not mean to imply that empathy increases stereotyping, but rather that individuals who have better empathizing skills use social information to generate predictions about the actions of others. In the present study, group membership information was the only information available to our participants. However, we would predict that if other types of information are available to the listener (e.g. information about specific past behaviour or about the mood the speaker is in), high empathizers use this information as well in generating expectations about what the speaker might say next.

Importantly, from our results it appears that within our experimental setting, low empathizers consistently did not use information about social stereotypes in implicit sentence comprehension, but rather took a more bottom-up approach to the processing of these social pragmatic sentences. In contrast, the N400 effect in the lexical semantic stimuli was still present in the second half of the experiment, indicating that the hard wiring for integration of lexical semantic information into the linguistic context appears to be much stronger than for social information. Apparently, it is harder to consider contextual lexical semantic information as unreliable than social information concerning stereotypes derived from the context. These findings can be related to Bayesian inference models, where the priors based on stereotypical knowledge are quickly adapted if incoming information generates prediction errors. It points to the fact that some priors (e.g. lexical-semantic based priors) are less easily updated than other priors, such as those based on the pragmatic information about the speaker.

Interestingly, the ERPs to the speaker identity manipulation in the second half of our experiment are reminiscent of findings of an ERP study by Lattner and Friederici (2003) investigating the processing of stereotypical information. In their study, participants heard utterances spoken by four male and four female speakers, half of which were incongruent with stereotypical assumptions about the speaker, always hinging on the speaker's sex. In contrast to our main finding of an N400 effect to violations of stereotypical assumptions, Lattner and Friederici obtained a late positive effect leading them to conclude that speaker identity information is processed at a later stage than semantic information. Statistical analysis in a later time window indicated that in the second half of our experiment the speaker identity manipulation also elicited a late positive effect (Figures 3 and 5). The specific design of the Lattner and Friederici's study may inadvertently have resulted in it being prone to adaptation processes, with participants rapidly realizing that stereotype-based information was not a reliable cue, thereby eliminating effects specifically related to stereotypical expectations (i.e. N400 effects interacting with participants' sex), and eliciting a later positive effect instead, possibly related to more reflective processes, following on-line integration as reflected by the N400 (Daltrozzo *et al.*, 2007).

CONCLUSION

The present electrophysiological study testifies to the importance of the investigation of inter-individual differences in cognitive functioning. Rather than considering them as noise that obscures underlying regularities in the data, individual differences may help to reveal the nature of underlying mechanisms (Underwood, 1975; Kosslyn *et al.*, 2002). Our results reveal individual differences in verbal social-information processing that are explained by individuals' cognitive styles. Inter-individual differences in a specific cognitive trait, the ability to empathize, modulate ERP effects (as well as oscillatory dynamics, see Supplementary Data online), in social information processing in a linguistic context. Whereas women were more likely to show brain reactivity to social pragmatic violations than men, an individual's ability to empathize was the sole reliable determinant of this aspect of social pragmatic language functioning. Individuals with an empathizing-driven cognitive style revealed larger N400 effects (and a larger increase in γ band power in the oscillatory domain) compared to low-empathizing individuals. This finding suggests that individuals with higher empathizing abilities are able to rapidly integrate information about the speaker with the content of the message as they appear to make use of these social aspects of language to process language in a top-down manner.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

REFERENCES

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* Cambridge, MA: MIT Press.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1430), 361–74.

Baron-Cohen, S., Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–75.

Baxter, L.C., Saykin, A.J., Flashman, L.A., et al. (2003). Sex differences in semantic language processing: a functional MRI study. *Brain and Language*, *84*, 264–72.

Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–61.

Campanella, S., Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–43.

Clements, A.M., Rimrodt, S.L., Abel, J.R., et al. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language*, *98*(2), 150–8.

Daltrozzo, J., Wioland, N., Kotchoubey, B. (2007). Sex differences in two event-related potentials components related to semantic priming. *Archives of Sexual Behavior*, *36*(4), 555–68.

Davis, M.H. (1996). *Empathy: A Social Psychological Approach*. Colorado: Westview Press.

Decety, J., Jackson, P.L. (2004). The functional architecture of human empathy. *Behavioral Cognitive Neuroscience Reviews*, *3*(2), 71–100.

Eisenberg, N., Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, *94*, 100–31.

Gur, R.C., Alsop, D., Glahn, D., et al. (2000). An fMRI study of sex differences in regional activation to a verbal and a spatial task. *Brain and Language*, *74*, 157–70.

Hall, J.A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*, 845–58.

Hall, J.A. (1984). *Nonverbal Sex Differences: Communication Accuracy and Expressive Style*. Baltimore: Johns Hopkins University Press.

Hall, J.A., Carter, J.D., Horgan, T.G. (2000). Gender differences in the nonverbal communication of emotion. In: Fischers, A.H., Fischers, A.H., editors. *Gender and Emotion: Social Psychological Perspectives*. Paris: Cambridge University Press, pp. 97–117.

Hall, J.A., Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*(2), 201–6.

Halpern, D. (1992). *Sex differences in cognitive abilities*. Lawrence Erlbaum, Hillsdale (NJ).

Hampson, E., Kimura, D. (1992). Sex differences and hormonal influences on cognitive function in humans. In: Becker, J., Breedlove, S., Crewss, D., editors. *Behavioral Endocrinology*. Cambridge, MA: MIT Press, pp. 357–98.

Herlitz, A., Nilsson, L., Backman, L. (1997). Gender differences in episodic memory. *Memory & Cognition*, *25*(6), 801–11.

Hill, H., Ott, F., Herbert, C., Weisbrod, M. (2006). Response execution in lexical decision tasks obscures sex-specific lateralization effects in language processing: evidence from event-related potential measures during word reading. *Cerebral Cortex*, *16*(7), 978–89.

Hoffman, M.L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin*, *84*, 712–22.

Hyde, J.S., Linn, M.C. (1988). Gender differences in verbal ability: a meta-analysis. *Psychological Bulletin*, *104*(1), 53–69.

Jaeger, J.J., Lockwood, A.H., Van Valin, R.D.J., Kemmerer, D.L., Murphy, B.W., Wack, D.S. (1998). Sex differences in brain regions activated by grammatical and reading tasks. *Neuroreport*, *9*(12), 2803–7.

Kansaku, K., Yamaura, A., Kitazawa, S. (2000). Sex differences in lateralization revealed in the posterior language areas. *Cerebral Cortex*, *10*(9), 866–72.

Kimura, D. (1999). *Sex and Cognition*. Cambridge, MA: MIT Press.

Knaus, T.A., Bollich, A.M., Corey, D.M., Lemen, L.C., Foundas, A.L. (2004). Sex-linked differences in the anatomy of the perisylvian language cortex: a volumetric MRI study of gray matter volumes. *Neuropsychology*, *18*(4), 738–47.

Kosslyn, S.M., Cacioppo, J.T., Davidson, R.J., et al. (2002). Bridging psychology and biology - the analysis of individuals in groups. *American Psychologist*, *57*(5), 341–51.

Lattner, S., Friederici, A.D. (2003). Talker's voice and gender stereotype in human auditory sentence processing - evidence from event-related brain potentials. *Neuroscience Letters*, *339*(3), 191–4.

Lee, Y.T., Jussim, L.J., McCauley, C.R. (1995). *Stereotype Accuracy: Toward Appreciating Group Differences*. Washington, DC: American Psychological Association.

Maccoby, E.E., Jacklin, C.N. (1974). *The psychology of sex differences*. Stanford: Stanford University Press.

Mehrabian, A., Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*(4), 525–43.

Phillips, M.D., Lowe, M.J., Lurito, J.T., Dzemidzic, M., Mathews, V.P. (2001). Temporal lobe activation demonstrates sex-based differences during passive listening. *Radiology*, *220*(1), 202–7.

Pugh, K.R., Shaywitz, B.A., Shaywitz, S.E., et al. (1996). Cerebral organization of component processes in reading. *Brain*, *119*, 1221–38.

Saxe, R., Baron-Cohen, S. (2006). The neuroscience of theory of mind. *Social Neuroscience*, *1*, 1–9.

Schirmer, A., Kotz, S.A. (2003). ERP evidence for a sex-specific stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, *15*(8), 1135–48.

Schirmer, A., Kotz, S.A., Friederici, A.D. (2005a). On the role of attention for the processing of emotions in speech: sex differences revisited. *Cognitive Brain Research*, *24*(3), 442–52.

Schirmer, A., Kotz, S.A., Friederici, A.D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, *14*(2), 228–33.

Schirmer, A., Striano, T., Friederici, A.D. (2005b). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, *16*(6), 635–9.

Schirmer, A., Zysset, S., Kotz, S.A., Yves von Cramon, D. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *NeuroImage*, *21*(3), 1114–23.

Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature*, *373*(6515), 607–9.

Tager-Flusberg, H. (1993). What language reveals about the understanding of minds in children with autism. In: Baron-Cohen, S., Tager-Flusberg, H., Cohens, D.J., editors. *Understanding Other Minds: Perspectives from Autism*. Oxford: Oxford University Press, pp. 138–57.

Underwood, B.J. (1975). Individual-differences as a crucible in theory construction. *American Psychologist*, *30*(2), 128–34.

Van Berkum, J.J.A., Van den Brink, D., Tesink, C.M.J.Y., Kos, M., Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–91.

Wirth, M., Horn, H., Koenig, T., et al. (2007). Sex differences in semantic processing: event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. *Cerebral Cortex*, *17*(9), 1987–97.