Can context affect gender processing? ERP differences between definitional and stereotypical gender

4 authors, including:

Hamutal Kreiner
Ruppin Academic Center

Sibylle Mohr
University of Glasgow

Klaus Kessler
Aston University

Can context affect gender processing?

ERP differences between definitional and stereotypical gender

Hamutal Kreiner, Sibylle Mohr,

Klaus Kessler, Simon Garrod

Department of Psychology
University of Glasgow

Abstract

Previous research using Eye-Tracking and ERP has shown that readers experience processing difficulty when an anaphor (*herself*) refers to a gender-mismatching antecedent (*minister*). The mismatch-effect is due to a clash between the gender of the pronoun and the gender of the co-referential noun. We report two EEG experiments using anaphora (Experiment 1) and cataphora (Experiment 2) sentences, designed to investigate the processing differences between stereotypical (*minister*) and definitional (*king*) gender role nouns. Consistent with previous findings (Osterhout et al., 1997), our results reveal similar mismatching effects for these noun types in anaphora. Critically, however, in cataphora, where the pronoun precedes its co-referring noun, diverging ERP signatures are revealed. Differences in an early interval suggest fast "gender-coercing" for stereotypical nouns while effects in a later interval are likely to reflect gender mismatch processing in definitional nouns. These findings suggest that discourse constraints modulate the processing of these two noun types in different ways.

1. Introduction

Although previous research has shown that gender information is represented at different linguistic levels such as morphology, syntax, and pragmatics, gender agreement is typically considered as a grammatical process. In English, however, there is no regular morphological marking for gender and the computation of gender-agreement may rely on non-grammatical information. For example, in words such as *boy* gender is part of the lexical entry whereas words like *soldier* are assumed to be definitionally neutral but their interpretation is stereotypically biased by world knowledge towards either male or female. There are two theoretical accounts that make different assumptions about the processing of definitional and stereotypical gender role nouns. According to the lexical view, gender is a lexical feature for both noun types, whereas the inferential view assumes that stereotypical gender has to be inferred from world knowledge or discourse context (Carreiras et al., 1996; Garnham et al., 2002). In this paper, we address these contrasting views and investigate whether the processing of gender is controlled by different processes in these two noun types.

Several Eye-Tracking (ET) studies on gender agreement demonstrate that readers immediately slow down when an anaphor (*herself*) refers to an antecedent (*minister*) that mismatches its stereotypical gender (Carreiras et al., 1996; Kennison, 2003; Sturt, 2003; Duffy & Keir, 2004). A corresponding mismatch-effect was also shown in ERP (Osterhout et al., 1997). These effects have been attributed to a clash between the gender of the pronoun and the conflicting gender of the co-referring noun. According to the inferential account, inferences are made immediately when a stereotypical noun is processed, leading to processing difficulty when the anaphor mismatches the inferred gender. According to a lexical view, stereotypical gender is a lexico-semantic feature, just like definitional gender, and the mismatch-cost reflects a feature clash. Most of the previous research trying to clarify this issue used anaphora sentences, where the antecedent precedes the referring pronoun (but cf. Banaji & Hardin, 1996). This allows time for inferences to take place before the referring expression is encountered, making it impossible to say whether the processing difficulty occurs at the discourse level due to an inference conflict or at the grammatical level due to a feature clash.

However, a recent ET study by Kreiner et al. (2008) compared anaphora [1a–1d] and cataphora [2a–2d] sentences contrasting the processing of definitional (*king*) versus stereotypical (*minister*) role nouns. Note that in cataphora, the reflexive pronoun precedes the co-referring role noun and unambiguously determines the gender of the relevant discourse referent before the critical noun is actually processed. Crucially, the lexical and the inferential views make different predictions for the processing of stereotypical and definitional role nouns in *cataphora* sentences: according to the inferential view, gender is lexically underspecified and acquired via inferences for stereotypical nouns. Consequently, in cataphora, where the pronoun has already assigned a gender to the discourse referent before the relevant role noun is encountered, the gender inference is constrained and the agreement violation is prevented. Hence, stereotypical violations of gender agreement in cataphora are not expected to exhibit a mismatch effect. By contrast, the lexical view presumes no qualitative difference in the processing of stereotypical and definitional role nouns and consequently predicts similar mismatch effects in violations of gender agreement for both noun types. The results of the ET study by Kreiner et al. (2008) showed indeed that in cataphora, where a gender-marked pronoun precedes its co-referential noun, the mismatch-cost for stereotypical gender is eliminated [2d vs. 2c], suggesting that unlike definitional-gender [2a, 2b] the processing of stereotypical gender is modulated by discourse constraints.

These findings are inconsistent with previous findings that showed that both noun types elicited similar ERP components (P600) in anaphora sentences, which suggested a syntactic rather than a discourse clash between the pronoun and the antecedent (Osterhout et al., 1997). To resolve these conflicting findings, we designed two EEG experiments using anaphora (Experiment 1, [1a–1d]) and cataphora sentences (Experiment 2, [2a–2d]) that contrast stereotypical and definitional role nouns bound by matching or mismatching reflexive pronouns. In these experiments, we examined the hypotheses that discourse constraints may modulate the processing of stereotypical gender whereas definitional gender that is lexically determined would be insensitive to such constraints. If distinct processes control gender agreement with stereotypical compared to definitional role nouns, we would expect the different nouns to elicit distinct ERP signatures during the processing of gender agreement.

2. Experiments

2.1. Method

Participants. For each experiment, we recruited twenty native speakers of English from the Glasgow University community who were paid for participation. Only participants that had not taken part in Experiment 1 were allowed to participate in Experiment 2.

Materials. Experiment 1 used 160 anaphora sentences such as [1a–1d] that employed a role noun as antecedent of a matching or mismatching co-referring pronoun (*himself/herself*; pronoun and role noun are both underlined in the examples) that served as the target word in the sentence. Each item had four versions, representing the four experimental conditions of a 2x2 design manipulating Gender Type (definitional [1a, 1b] versus stereotypical [1c, 1d]) and Matching (match [1a, 1c] versus mismatch [1b, 1d]). The selection of stereotypical role nouns was based on a norming study with 50 participants from Glasgow University who did not take part in the other experiments.

[1a] The <u>king</u> left London after reminding <u>himself</u> about the letter.

[1b] The <u>king</u> left London after reminding <u>herself</u> about the letter

[1c] The <u>minister</u> left London after reminding <u>himself</u> about the letter

[1d] The <u>minister</u> left London after reminding <u>herself</u> about the letter

Experiment 2 was similar to Experiment 1, except for the experimental sentences. We used 160 cataphora sentences such as [2a–2d] which were derived from the materials used in Experiment 1. The linear order of pronoun and noun was reversed, i.e. the referring pronoun occurred prior to the critical noun, hence the noun served as the target word. As in Experiment 1, each item appeared in four versions, representing the four experimental conditions.

[2a] After reminding himself about the letter, the king left London.

[2b] After reminding herself about the letter, the king left London.

[2c] After reminding himself about the letter, the minister left London.

[2d] After reminding herself about the letter, the minister left London.

Procedure. EEG was recorded while participants read silently sentences presented word by word on a computer screen. A BIOSEMI Active-Two amplifier system was used for continuous recording of EEG activity from 72 electrodes at a sampling rate of 256 Hz. Each trial started with a sentence title including 2–5 words presented on the screen until the participant was ready and pressed a key to continue. This was followed by a 500 ms blank screen and then by a fixation cross in the centre of the screen. The sentence was then presented word-by-word. The words appeared in the centre of the screen for a duration of 300 ms each, and with an inter-stimulus interval of 200 ms. In addition to the 160 experimental items 160 filler sentences were included in each experiment. Finally, 25% of the stimuli (both experimental and fillers) were followed by a simple comprehension question that required a yes/no reply via button press.

Data analysis. BESA Version 5.2 was used to filter frequency bands of 0.3–25.0 Hz. and correct or reject ERP artifacts. Participants for which more than 12.5% of the trials of a particular experimental condition were rejected were excluded from the analysis (all together 4 participants in each experiment). EEG was time-locked to the onset of the target word (i.e., the pronoun in Experiment 1 and the role noun in Experiment 2). We computed average waveform in this critical time window for each participant in each experimental condition, after normalizing the waveforms of the individual trials relative to a 150 ms pre-stimulus baseline interval preceding the critical word using the electrodes' average as reference.

For the purpose of statistical analysis, the electrodes were grouped into five Regions Of Interest (ROI): Left Anterior (LA: Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, C1, C3, C5, T7); Left Posterior (LP: PO9, O1, PO7, P7, CP3, CP1, TP7, CP5, P1, P3, P5, PO3); the corresponding Right Anterior (RA) and Right Posterior (RP) regions; and the Central region (Pz, CPz, Fz, FCz, Cz). ERPs were averaged over electrodes within each ROI. At each relevant epoch, we performed hierarchical statistical analysis, starting with a within subject 2x2x2x2 design reflecting (2 hemispheres) x (2 anterior-posterior ROIs) x (2 types of role nouns – definitional versus stereotypical) x (2 matching conditions – match versus mismatch). The Central ROI was analyzed

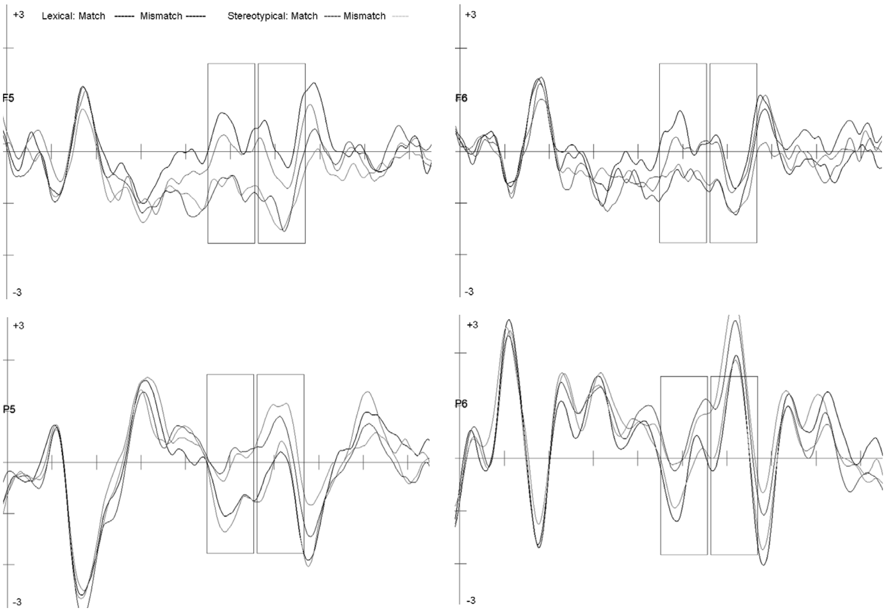

Figure 1. Grand averaged ERPs time-locked to the onset of the target word (the reflexive pronoun *herself/himself*). The darkest lines represent the matching conditions with the darkest representing the Definitional-Matching and the slightly lighter line the Stereotypical-Matching. The lighter lines represent the mismatching conditions with the slightly darker representing the Definitional-Mismatching and the lightest line the Stereotypical-Mismatching. Four example electrodes that represent the four different ROIs and show reliable effects are presented.

separately. Subsequently, for ROIs in which reliable effects were observed, we report repeated measures ANOVA testing the effects of experimental factors for each electrode.

2.2. Experiment 1: Results and discussion

Figure 1 presents grand averages for four example electrodes from the LA, LP, RA, and RP ROIs. As can be clearly seen in this figure, matching and mismatching conditions show diverging patterns from around 450 ms after stimulus onset. Based on this visual observation, we performed statistical analysis on the epochs of 450–550 ms and 550–650 ms from pronoun onset.

In the epoch of 550–650 ms from the target (pronoun) onset, we found a reliable effect for Matching ($F(1, 19) = 26.24$, $p < .0001$) modulated by a reliable interaction with Hemisphere ($F(1, 19) = 5.07$, $p < .05$) and with the Anterior-Posterior factor ($F(1, 19) = 35.83$, $p < .0001$), as well as a reliable 3-way interaction between Matching, Noun Type and the Anterior-Posterior factor ($F(1, 19) = 7.29$, $p < .05$). Subsequent 2x2 analysis of each ROI shows a reliable Matching effect in all ROIs except the Central

(LA: $F(1, 19) = 45.89$, $p < .0001$; LP: $F(1, 19) = 6.63$, $p < .05$; RA: $F(1, 19) = 9.13$, $p < .005$; RP: $F(1, 19) = 39.29$, $p < .0001$). This was modulated by an interaction with the Noun Type reliable only in the Left ROIs (LA: $F(1, 19) = 6.63$, $p < .05$; LP: $F(1, 19) = 7.91$, $p < .05$). Inspections of the epoch mean amplitudes at each region reveals a positivity Matching effect for both LP (definitional match –0.13 vs. mismatch 0.80; stereotypical match 0.09 vs. mismatch 0.57) and RP (definitional match 0.47 vs. mismatch 1.63; stereotypical match 0.68 vs. mismatch 1.38). The reverse pattern, namely a negativity effect, was observed for LA (definitional match –0.15 vs. mismatch –1.12; stereotypical match –0.41 vs. mismatch –0.87) and RA (definitional match –0.04 vs. mismatch –0.51; stereotypical match –0.12 vs. mismatch –0.42). Although the interactive pattern, showing a slightly smaller difference between the matching and mismatching conditions of stereotypical compared to definitional gender, is only reliable for the posterior ROIs, a similar trend is observed in the Anterior ones. Detailed analysis of the single electrode data reveals a consistent pattern with reliable Matching effect in several electrodes in LA (Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3), LP (CP5, CP3, P1, P3, P5, P7, PO7, PO3, O1), RA (Fp2, AF8, AF4, F2, F4, F6, F8, FT8, FC6), RP (TP8, CP6, CP4, CP2, P2, P4, P6, P8, PO8, PO4, O2) and in the Central ROI (Iz, Oz, POz, Pz, CPz, Fpz, AFz, Fz). Only a few electrodes showed reliable interaction (La: Fp1, F7, FC5, FC3; LP: PO7, O1; RP: P8; Central: Pz).

Analysis of the earlier epoch of 450–550 ms from the target onset reveals that the Matching effect is reliable at this epoch. Thus, a reliable effect is shown for Matching ($F(1, 19) = 18.65$, $p < .0005$) modulated by a reliable interaction with the Anterior-Posterior factor ($F(1, 19) = 17.16$, $p < .001$). Subsequent 2x2 analysis of each ROI shows a reliable Matching effect in all ROIs (LA: $F(1, 19) = 15.64$, $p < .001$; LP: $F(1, 19) = 13.05$, $p < .01$; RA: $F(1, 19) = 7.38$, $p < .01$; RP: $F(1, 19) = 17.50$, $p < .0001$; Central: $F(1, 19) = 26.83$, $p < .0001$). In this epoch, unlike in the 550–650 epoch, this effect was modulated by an interaction with the Noun Type only at the Central ROI ($F(1, 19) = 14.28$, $p < .001$). As can be seen in the data from the example electrodes (Figure 1) at this epoch, the Matching effect is more negative for the posterior ROIs and relatively positive for the anterior ROIs. Detailed analysis of the single electrode data reveals that the Matching effect is reliable in several electrodes in LA (Fp1, AF7, AF3, F3, F5, F7, FT7, FC5, FC3, FC1), LP (CP5, CP3, CP1, P1, P3, P5, P7, PO7, PO3, O1), RA (Fp2, AF8, AF4, F6, F8, FT8, FC6, FC2, C2, C4), RP (TP8, CP6, CP4, CP2, P2, P4, P6, P8, PO8, PO4, O2) and in the Central ROI (POz, Pz, CPz, Fpz, AFz, FCz, Cz).

In general, the results are consistent with previous findings that showed a P600-like Matching effect for both noun types modulated by an interaction with the Noun Type (Osterhout et al., 1997) and with EM findings showing a corresponding mismatch-cost for both Noun Types.

2.3. Experiment 2: Results and discussion

Figure 2a and 2b presents grand averages for four example electrodes from the LA, LP, RA, and RP ROIs. A brief look at the figure reveals that the ERP patterns elicited by the cataphora sentences are very different from those elicited by anaphora. The most striking difference is that the Gender Matching factor does not seem to result in a very clear Matching effect across the two Noun Types as demonstrated in Experiment 1. Based on our previous eye-movement findings that showed that the gender matching effect for cataphora is delayed compared to anaphora, and on visual inspection of the wave forms, we performed statistical analysis on the epochs of 650–800 ms and 250–400 ms from pronoun onset.

In the epoch of 650–800 ms from the target (role noun) onset, the general 4-way analysis did not reveal any reliable effects. However, subsequent 2x2 ANOVAs for the individual ROIs revealed that the effect of Noun Type reached significance in the RA ROI ($F(1, 19) = 5.35$, $p < .05$). Whereas none of the ROIs exhibited a reliable effect of Matching, the interaction between Matching and Noun Type was reliable in both LA ($F(1, 19) = 8.49$, $p < .01$) and RP ($F(1, 19) = 6.98$, $p < .05$) ROIs. Inspections of the epoch mean amplitudes clarifies the pattern of simple effects underlying this interaction: a reliable Matching effects for the definitional (RP: match –0.08 vs. mismatch –0.14; LA: match –0.20 vs. mismatch 0.06) but not for the stereotypical (RP: match 0.04 vs. mismatch 0.09; LA: match –0.10 vs. mismatch –0.15) conditions. Detailed analysis of the single electrode data in ROIs that showed reliable effects reveals a reliable interaction in several electrodes in the LA ROI (F1, F3, F5, FC5, FC3, FC1) and only in one electrode (P6) in the RP ROI. The effect of Noun Type reached significance in 3 RA electrodes (F8, FT8, T8). While this pattern of results is not as clear as the results shown for anaphora sentences, it is generally consistent with our previous EM findings for cataphora sentences that indicated an interaction between Noun Type and Matching whereby violation of definitional gender agreement, unlike stereotypical, results in a processing disruption.

In the 450–550 ms and 550–650 ms from the target onset, we see some reliable effects for the Noun Type. Such effects may reflect differences between these noun types (e.g. morphology) that are not the main focus of this paper (as long as they do not modulate the Matching effect). Hence, the related analyses are not reported here. Surprisingly, in the earlier epoch of 250–400 ms from target onset, the general 4-way analysis reveals a reliable 3-way interaction between the Matching, the Noun Type and the Hemispheres ($F(1, 19) = 5.10$, $p < .05$). None of the main effects or other interactions has reached significance. Subsequent 2x2 ANOVAs for the individual ROIs revealed that the effect of Noun Type reached significance in the RP ROI ($F(1, 19) = 5.04$, $p < .05$). Whereas the Matching effect was not reliable, the interaction between Matching and Noun Type was reliable in both LP ($F(1, 19) = 8.24$, $p < .01$) and RA ($F(1, 19) = 8.34$, $p < .01$) ROIs. Inspections of the epoch mean amplitudes clarifies the pattern of simple effects underlying this interaction, showing Matching effects for the stereotypical (RP: match 0.07 vs. mismatch 0.11; LA: match –0.08 vs. mismatch –0.13) but not for the definitional (RP: match –0.08 vs. mismatch 0.005; LA: match –0.01 vs. mismatch –0.05) conditions. Detailed analysis of the single electrode data in ROIs that showed reliable effects reveals a reliable interaction in several electrodes in the LP (P5, P7, PO7, O1) and RP (FC6, C6, T8) ROIs. These relatively early effects for stereotypical but not definitional gender agreement do not correspond to our previous EM findings for cata-

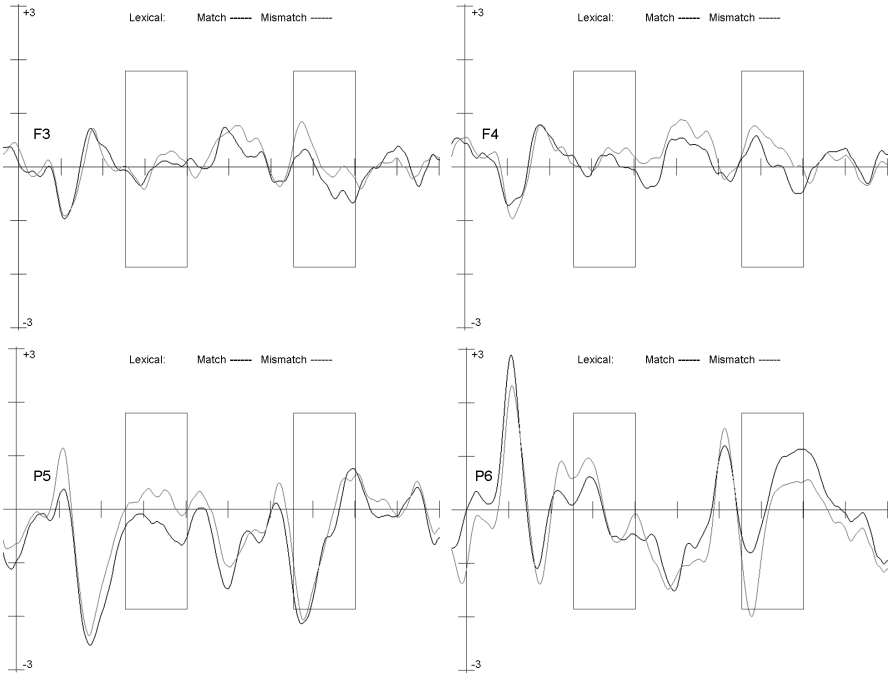

Figure 2a. Grand averaged ERPs time-locked to the onset of the target word (the role noun *king/minister*) in the Definitional condition. The darker line represents the matching condition and the slightly lighter line the Mismatching condition. Four example electrodes that represent the four different ROIs and show reliable effects are presented.

phora sentences. It seems, therefore that the ERP effects at this epoch reflect processes that are not captured by the EM measures. One possible interpretation of this effect is that it reflects the process underlying the coercion of stereotypical gender to discourse constraints (we discuss this interpretation in more detail in the Conclusion).

2.4. Oscillatory brain activity

Recent research suggests that transforming the EEG signal into the time-frequency domain may provide a more sensitive analysis, since changes in oscillatory activity are largely cancelled out due to signal averaging during ERP analysis. Therefore we also looked for Event-related Phase Synchronization in a frequency range of 2–60 Hz to determine whether stimulus-specific oscillatory modulations are linked to gender-agreement violations during the anaphoric processing of definitional and stereotypical gender nouns. Power estimates for each subject and each condition were obtained from wavelet-based time-frequency (TF) representations using the BESA® 5.1 software package. To isolate induced-activity oscillations, we subtracted evoked activity oscillations

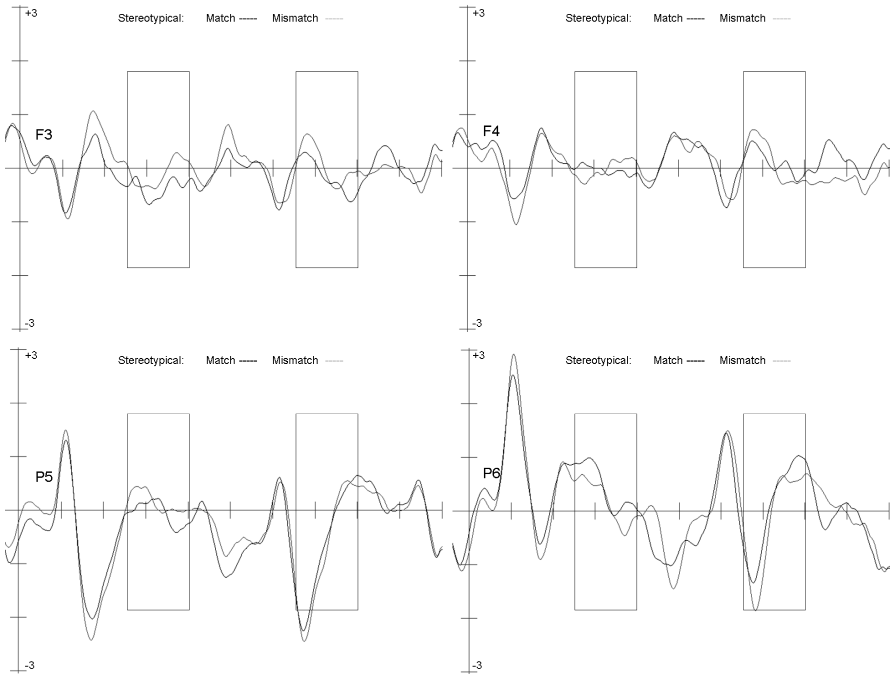

Figure 2b. Grand averaged ERPs time-locked to the onset of the target word (the role noun *king/minister*) in the Stereotypical condition. Legend as for Figure 2a.

by means of regression (cf. Tallon-Baudry, 2003). Using FieldTrip (a Matlab® toolbox developed at the FC Donders Centre for Cognitive Neuroscience, Nijmegen) the TF representations were then averaged over trials for each subject and each condition. To normalize for inter-individual differences, we set a baseline interval –5.7 s to –5.2 s prior to stimulus onset. We analyzed the time window of 0–1000 ms from stimulus onset applying a randomization procedure, a cluster-based permutation test (Monte Carlo method, 1000 randomizations, $\alpha = 0.05$) with a cluster growing approach that corrects for the multiple comparisons problem (Maris et al., 2007). This method identifies time-frequency-channel clusters, which show reliable power changes between the mismatching and matching conditions. This analysis revealed significant differences only for the alpha and gamma frequency bands.

At the alpha frequency range (8–12 Hz) the comparison between mismatch vs. match conditions shows a decrease of power in both experiments for both definitional and stereotypical role nouns (one negative cluster each, $p < .05$, left posterior electrodes for both experiments, e.g. Pz, PO7, PO3, PO1, P7, P5, P3, O1, 0–1000 ms). This result indicates that while the alpha band activity is sensitive to gender matching, it does not reveal any distinction between the different noun types we have examined.

In contrast, the gamma frequency range (30–60 Hz for all channels) appears to be more sensitive to the processing differences between these two noun-types. In the cataphora sentences (Experiment 2) it reveals a positive cluster (p = 0.02) indicating a strikingly larger power increase for the mismatch compared with match conditions for definitional but not for stereotypical role nouns. The anaphora sentences of Experiment 1 do not reveal such a pattern. Subsequent comparison of hemispheres showed that the effect is confined to central electrodes of the right hemisphere (F6, F8, FC2, FC4, FC6, FT8, C2, C4, C6, T8, CP4, CP6, TP8) and peaks around 300–900 ms.

Currently, little is known about the neuro-cognitive functions underlying oscillatory brain activity during language comprehension. Therefore, we can only propose preliminary and rudimentary interpretations of these findings. Modulations in the alpha frequency range have been interpreted as reflecting attention (Klimesch, 1999), working memory (Jensen et al., 2002), semantic processing (Rohm et al., 2001), and visual processing involved in reading words (Bastiaansen & Hagoort, 2003). The differentiated alpha modulations in matching compared to mismatching conditions observed in both experiments are consistent with higher level functions such as working memory or semantic processing. Gamma rhythms have been linked to attention and alertness (Tallon-Baudry, 2003) and may reflect underlying feature-binding and memory processes ("binding" gamma rhythm; Engel & Singer, 2001). Although this feature-binding interpretation was proposed in the context of attentional-perceptual integration, it may have implications for linguistic integration. The widespread gamma modulation for the definitional mismatch condition in cataphora, which is hardly observable in anaphora, may suggest differences in the process of linguistic binding underlying reference resolution in anaphora compared with cataphora (see Kazanina et al. (2007) for a discussion of the processing differences between anaphora and cataphora). Future research will shed more light on these somewhat speculative interpretations and on the potential contributions of lexical syntactic and discourse factors to such gamma modulations.

3. Conclusion

The results of Experiment 1 clearly show a mismatch effect in anaphora sentences for both stereotypical and definitional gender role-nouns, and this effect is quantitavely modulated by an interaction with the noun type. This pattern replicates Osterhout et al.'s (1997) findings, and is seemingly consistent with the lexical view that the representation and processing of gender in these two noun types are qualitatively similar. The quantitative differences between them may reflect graded differences in the gender representation rather than distinct processing mechanisms. However, as argued above, the critical test for the hypothesis that the gender processing in these noun types is controlled by different processes is the cataphora experiment.

The cataphora sentences used in Experiment 2 elicited strikingly different waveforms and ERP components, plausibly related to general differences (discussed below) between these structures. However, the critical finding in addressing the definitional-stereotypical processing difference hypothesis is that in cataphora, unlike anaphora, these noun types elicit distinct ERP signatures. Thus, while in anaphora both noun types exhibited a similar mismatch-effect, in cataphora differentiated patterns were observed: initially, at the 250–400 ms, a mismatch effect was observed only for stereotypical role nouns; later at the 650–800 ms epoch, a mismatch-effect was observed only for definitional and not for stereotypical gender. Similarly, the TF analysis at the gamma band revealed a mismatch effect only for definitional gender. The latter definitional mismatch-effects correspond to the ET findings of Kreiner et al. (2008), indicating that violation of gender agreement in cataphora sentences results in reading difficulty only for definitional gender. Both the inferential and the lexical views assume that definitional gender is part of the lexical entry. Thereby, it is accessed by default, leading to a gender clash when the dependency is resolved and agreement violation is detected both in anaphora and in cataphora. Crucially, however, the inferential view, unlike the lexical view, predicts a different pattern in cataphora. Based on the inferential view, we initially predicted that in cataphora stereotypical role nouns would not exhibit mismatch effects because the discourse constraints set by the preceding gender marked pronoun make gender inferences unnecessary and may therefore prevent the gender clash. The stereotypical mismatch effect revealed in the relatively early (250–400 ms) epoch is inconsistent with this prediction. To account for this finding we propose a slightly different interpretation of the inferential view. Namely, stereotypical gender inferences may not be prevented by discourse constraints. Rather, they are automatically activated but coerced by discourse constraints in case of conflict. According to this account, the early and short-lived mismatch-effect exhibited by the stereotypical role nouns may be interpreted as reflecting the process of coercion, i.e. shifting from the world-knowledge based gender bias to the discourse constrained bias. Since such coercion cannot occur in the case of definitional role nouns, this early mismatch-effect is exhibited selectively by stereotypical role nouns. The notion of automatic activation of gender stereotype is consistent with previous findings both from anaphora resolution (Oakhill et al., 2005; Reynolds et al., 2006) and from priming (Banaji & Hardin, 1996) studies.

Owing to the different linear order of the cataphora sentences, there is another major difference between the stimuli used in the two experiments. In cataphora, the earliest point at which agreement can be computed is when the role noun is encountered; inevitably then the target word is the role noun. Thus, whereas the target words in anaphora are high frequency function words and morphologically marked for gender (*himself/herself*), in cataphora the targets are of lower frequency and not marked with a regular morphological gender marker[1]. It is plausible that the different wave forms and ERP components observed reflect the different processes involved in processing pronouns compared with role nouns. However, this divergence cannot account for the differences in the mismatch effect in stereotypical compared to definitional nouns.

Taken together, the findings from the two experiments reported here are consistent with the inferential view. In showing distinct ERP signatures for reference resolution

1 Some of the definitional role nouns have affixes such as *-ess* in *waitress*. However, since these are not regular gender markers in English, we tend to consider them as part of the lexical entry rather than a morpho-grammatical feature.

with stereotypical compared with definitional role nouns, they support the view that the underlying mechanisms of gender processing in these noun types are different.

4. Acknowledgements

This work is part of the ESRC-funded Gender project (grant number: RES-000-22-1924).

References

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7*(3), 136–141.

Bastiaansen, M., & Hagoort, P. (2003). Event-induced theta responses as a window on the dynamics of memory. *Cortex, 39*, 967–992.

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from english and spanish. *Quarterly Journal of Experimental Psychology, 49*, 639–663.

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551–559.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences, 5*(1), 16–25.

Garnham, A. Oakhill, J., & Reynolds, D. (2002). Are inferences from stereotyped role names to characters' gender made elaboratively? *Memory & Cognition, 30*(3), 439–446.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex, 12*(8), 877–882.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language, 56*, 384–409.

Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research, 32*(3) 355–378.

Klimesch, W. (1999) EEG alpha and thete oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29: 169–195.

Kreiner, H., Sturt, P., & Garrod, S. C. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language, 58*, 239–261.

Maris, E., Schoffelen, J.-M., Fries, P. (2007). Nonparametric statistical testing of coherence differences. *Journal of Neuroscience Methods, 163*, 161–175.

Oakhill, J. V, Garnham, A., & Reynolds, D. J. (2005). Immediate activation of stereotypical gender information in reading. *Memory & Cognition, 33*(6), 972–983.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25*, 273–285.

Reynolds, D. J., Garnham, A., & Oakhill, J. (2006), Evidence of immediate activation of gender information from a social role name. *The Quarterly Journal of Experimental Psychology, 59*(5), 886–903.

Rohm, D., Klimesch, W., Haider, H., & Doppelmayr, M. (2001). The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters, 310*(2–3), 137–140.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language, 48*(3), 542–562.

Tallon-Baudry, C. (2003). Oscillatory synchrony and human visual cognition. *J. Physiol. Paris* 97 (2–3), 355–363.