



## Name agreement in picture naming: An ERP study

Xiaorong Cheng<sup>a</sup>, Graham Schafer<sup>a,\*</sup>, Elkan G. Akyürek<sup>a,b</sup>

<sup>a</sup> School of Psychology and Clinical Language Sciences, University of Reading, Reading, Harry Pitt Building, Earley Gate, Whiteknights Road, Reading, Berkshire, RG6 6AL, UK

<sup>b</sup> Department of Psychology, University of Groningen, Groningen, The Netherlands

### ARTICLE INFO

#### Article history:

Received 6 May 2009

Received in revised form 5 March 2010

Accepted 6 March 2010

Available online 12 March 2010

#### Keywords:

ERP

Naming

Lexical processing

Semantics

Categorization

### ABSTRACT

Name agreement is the extent to which different people agree on a name for a particular picture. Previous studies have found that it takes longer to name low name agreement pictures than high name agreement pictures. To examine the effect of name agreement in the online process of picture naming, we compared event-related potentials (ERPs) recorded whilst 19 healthy, native English speakers silently named pictures which had either high or low name agreement. A series of ERP components was examined: P1 approximately 120 ms from picture onset, N1 around 170 ms, P2 around 220 ms, N2 around 290 ms, and P3 around 400 ms. Additionally, a late time window from 800 to 900 ms was considered. Name agreement had an early effect, starting at P1 and possibly resulting from uncertainty of picture identity, and continuing into N2, possibly resulting from alternative names for pictures. These results support the idea that name agreement affects two consecutive processes: first, object recognition, and second, lexical selection and/or phonological encoding.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

A picture of an object may be a good (i.e., faithful) depiction of that object, or it may not. The extent to which different people agree that a given picture depicts an object with a particular name is termed that picture's *name agreement*. Name agreement has robust behavioral effects in cognitive tasks. Naming a picture that has only one name is faster and more accurate than naming a picture that has multiple names. As summarized by Alario et al. (2004), name agreement is a significant predictor of picture naming latency for: American English (Snodgrass and Yuditsky, 1996), Welsh (Barry et al., 1997), British English (Ellis and Morrison, 1998), French (Bonin et al., 2002; Bonin et al., 2003), Spanish (Cuetos et al., 1999) and Italian (Dell'Acqua et al., 2000).

Name agreement appears to have its effect on naming latency independent from frequency or age of acquisition (AoA) effects (Lachman et al., 1974; Vitkovitch and Tyrell, 1995). Alario et al. (2004) reported that name agreement was a stronger predictor of picture naming latency than eight other image attributes, namely: visual complexity, image agreement, imageability, AoA, word frequency, concept familiarity, number of syllables and number of phonemes.

The present study is a preliminary investigation of event-related potential ('ERP') phenomena associated with naming of pictures with

different levels of name agreement. In the future, such studies may elucidate processes involved in lexical access. Our present and more limited goal is to describe ERP effects of name agreement, in the context of neural models of word production.

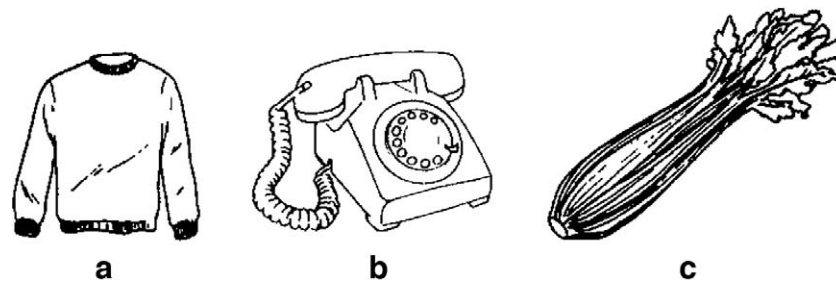
Vitkovitch and Tyrell (1995) identified three different types of 'name disagreement':

- (1) The use of multiple names, for example, in Fig. 1a, the object depicted is a jumper and also called 'sweater', 'pullover', 'jersey', 'sweatshirt'.
- (2) The use of appropriate abbreviations (or elaborations), for example, in Fig. 1b, an acceptable abbreviation for telephone is 'phone'.
- (3) Incorrect names. In Fig. 1c, the picture is celery, but some people misidentify it, giving incorrect names such as 'rhubarb', 'Chinese leaves', 'cabbage', or 'marrow'.

In their first experiment, Vitkovitch and Tyrell (1995) compared the naming latencies for the three kinds of low name agreement images with corresponding high name agreement images (matched for frequency of the object's name in print, word length, complexity, familiarity, image agreement, and AoA). Low name agreement due to multiple or incorrect names (Fig. 1a and c) increased naming latency, while correct abbreviations or elaborations did not (Fig. 1b). An object–decision task was conducted to see if the two kinds of disagreement which had been found to delay object naming took effect at the same stage of processing. Participants had to decide whether a picture represented a real object or not. In such a task, it is not necessary to retrieve the lexical representation of the name of a

\* Corresponding author.

E-mail address: [G.W.Schafer@reading.ac.uk](mailto:G.W.Schafer@reading.ac.uk) (G. Schafer).



**Fig. 1.** Three kinds of name disagreements: (a) the use of multiple alternative names ('AN' in the text); (b), the use of correct abbreviations (or elaborations); (c), incorrect names ('PU' – 'picture uncertainty' in the text).

Pictures were from Snodgrass and Vanderwart (1980).

picture, so semantic<sup>1</sup> processing or name selection are not required, and presumably therefore any increase in latency suggests an effect of name agreement occurring before semantic processing. Object–decision times for pictures with low name agreement due to multiple names did not significantly differ from those for control pictures, whereas object–decision times for pictures with low name agreement due to incorrect names were longer than those for control pictures. Vitkovitch and Tyrell concluded that the naming delay due to incorrect names (picture uncertainty) had its locus at the stage of object recognition. Further, they argued that the increased latency to name pictures with multiple names happens only at the name retrieval stage. They identified two possible representational structures which might lead to such a delay: (1) a single conceptual representation linked to multiple downstream semantic pathways, or (2) distinct conceptual representations underlying each alternative name.

Studies of predictors of picture naming have generally been conducted within the framework of stage models (e.g., Bonin et al., 2002; Glaser, 1992; Humphreys et al., 1988; Levelt, 1989, 1991; Levelt et al., 1999). One influential stage model is the LRM model, named after Levelt et al. (1999). In this model, the process of picture naming involves four sequential stages: (1) object recognition (to access the stored structural representation from the visual representation of an object); (2) concept activation (to activate concepts related to the object); (3) lexicalization (to retrieve the name of the object); and (4) articulation. Lexicalization in this model is further assumed to proceed in two steps: selection of the *lemma* (word meaning and relation to any other words in a potential utterance) and phonological encoding (e.g., Dell, 1986; Kempen and Huijberts, 1983; Levelt, 1989, 1991). To parameterize the stages of word production, Indefrey and Levelt (2004) conducted a meta-analysis of 82 studies of word production and concluded that, in picture naming, conceptually driven lexical selection takes place between 175 ms and 250 ms post-stimulus-onset; lexical phonological code retrieval between 250 ms and 330 ms post-stimulus-onset; and syllabification between 330 ms and 455 ms post-stimulus-onset.

However, it is by no means clear that cognitive processes in picture naming take place sequentially. There is debate about the extent to which one process must be complete before the next can begin. Although serial computation is possible (see e.g., Alario et al., 2004, Fig. 1), processes might alternatively be conducted in a cascade, with partially-computed output from one step forming the input to the next (Humphreys et al., 1988). In principle, name agreement effects provide a way to study these issues in detail, because in low name agreement words, a *single* concept may activate *multiple* words. An important first step then is to establish whether, and at what point(s) in time, the brain responds differently to high and low name agreement images. In the experimental work described in this paper

we use event-related potential ('ERP') methods to investigate the time point(s) at which name agreement has its effect(s).

Consistent with Vitkovitch and Tyrell's (1995) study, other researchers have differentiated between effects of name agreement during object recognition and lexical processing stages. Alario et al. (2004) distinguished between 'image agreement' and 'name agreement' effects. Image agreement is the extent to which a picture matches images generated by participants in response to that picture's name (i.e., the extent to which a picture matches people's stored canonical representations of that item). Alario et al. proposed that image agreement has its effect in the visual recognition system, while true name agreement effects occur between the conceptual stage and the lexical stage: low name agreement pictures evoke more candidates for the name of the depicted object than do high name agreement pictures; and therefore, compared with high name agreement pictures, it takes longer to eliminate competitors and select one specific name for such low name agreement pictures. Such reasoning is in line with Johnson's (1992) proposition that name agreement effects arise after object identification, again on the basis that low name agreement increases naming time, but leaves object–decision response times unaffected. Johnson, Paivio, and Clark suggested that name agreement effects occur '[during] name retrieval, response generation, or both' (Johnson et al., 1996, p119). Bonin et al. (2002) also agreed with Vitkovitch and Tyrell (1995) that there are two main sources of name disagreement, namely uncertainty of pictures and alternative names of depicted objects, and propose that in the case of picture uncertainty, name agreement effects occur while accessing stored structural knowledge, whereas if a picture has more than one alternative name, name agreement effects occur after conceptual access.

ERP and MEG methods have been used with success in recent years to uncover cognitive processes involved in picture naming (e.g., Abdel Rahman et al., 2002; Eulitz et al., 2000; Greenham et al., 2000; Koester and Schiller, 2008; Laganaro et al., 2009; Maess et al., 2002; Strijkers et al., 2010). The sequence of cognitive processes in picture naming have been addressed in some ERP studies using lateralized readiness potential ('LRP') methods in two-choice go/no-go tasks (Rodriguez-Fornells et al., 2002; Schmitt et al., 2001; van Turennout et al., 1997). The LRP is a slow, negative-going potential developing in advance of the execution of a voluntary movement (see Jescheniak et al., 2002, pp. 591–2, for a brief description and evaluation of the LRP technique in this context). By manipulation of go/no-go tasks requiring semantic or phonological processing, it has been reported that semantic processing occurs about 120 ms earlier than phonological processing (van Turennout et al., 1997) and syntactic processing about 40 ms earlier than phonological processing (van Turennout et al., 1998). Schmitt et al.'s (2001) findings are also consistent with this estimate of semantic processing beginning approximately 80 ms before syntactic processing.

However, ERP paradigms using the lateralized readiness response measure are potentially problematic because the cognitive processes required to perform the task (explicit monitoring of phonemes, and explicit categorization judgments) may be very different from the

<sup>1</sup> We use the term 'semantic' to refer to knowledge of word meaning. We use the term 'conceptual' to refer to knowledge about an *object*. The former entails but is not limited to lexical knowledge; the latter may or may not entail linguistic knowledge.

(fast, implicit) processes invoked in 'pure' picture naming, or in word production more generally. To address this issue, more recent ERP studies have employed a range of techniques. One such study involved auditorily-presented target words in a delayed picture naming paradigm (for example, Jescheniak et al., 2003; Jescheniak et al., 2002) to investigate the relative order of activation of semantic and phonological information during speech planning. Jescheniak et al. (2002) recorded ERPs to auditorily-presented words which were either semantically or phonologically related to the name of priming images presented immediately beforehand. Participants performed a linguistic task (delayed naming). Jescheniak et al. found an early (250–400 ms) effect of phonological relatedness, but not semantic relatedness, and a late (400–800 ms) effect for both kinds of relatedness. In contrast, when participants were performing a nonlinguistic task (size judgment), there were no early effects, and only a late effect of semantic relatedness. It therefore appears that both phonological and semantic codes are activated by a linguistic task, but only semantic codes by a nonlinguistic one, suggesting that in word production tasks evoked by picture naming, phonological codes are activated downstream of—or at least, later than—semantic ones.

Another ERP paradigm which uses picture naming successfully has involved the use of 'covert', or silent, naming (e.g., Greenham et al., 2000). In the present study we also applied a covert naming paradigm to investigate name agreement effects in the time-course of picture naming. We used two groups of pictures, each having differing levels of name agreement (high vs. low), and matched on other picture attributes, namely: objective AoA, rated frequency, familiarity, visual complexity, picture–name agreement, number of phonemes, and number of syllables (Vitkovitch and Tyrell, 1995). We recorded ERPs while participants named pictures covertly and compared ERP for pictures with high name agreement and ERPs for pictures with low name agreement.

Other ERP studies have used an immediate overt picture naming method, by analyzing artifact-free ERPs before overt naming (Koester and Schiller, 2008; Strijkers et al., 2010). The difference between overt (spoken) naming and covert (silent) naming has been directly investigated using ERP methods, by Eulitz et al. (2000). ERPs were recorded when participants passively viewed pictures, and were compared with ERPs taken when participants viewed the same pictures and either overtly or covertly named them. It was concluded that up to 400 ms post-stimulus-onset, there is little difference between these two methods of presentation, in terms of the ERP components observed during picture naming (p. 2096).

What features might we expect to observe in our ERP? Greenham et al. (2000) studied ERP responses in a picture naming paradigm in which participants covertly named words and pictures. Words and pictures were presented either individually or in superimposed word–picture pairs. When pictures were presented individually, the averaged distribution of ERP responses included occipito-parietal P240 and P390 components, and a frontal N450.

To permit some exploration of the relative effects of the source of any observed name agreement effects, participants in our study were

asked in a post-test to judge the source of name disagreement for each picture in the study. Our purpose was to establish whether, for each individual, low name agreement of a particular picture was caused by 'picture uncertainty' or 'alternative names'. This allowed low name agreement pictures to be divided into two subgroups, by participant.

## 2. Method

### 2.1. Participants

Twelve females and 7 males participated in return for course credit. Their mean age was 20.4 years (SD = 3.52, 18–33). All were right-handed (by self report) native speakers of English, educated to at least high-school level. Another four participants' data were discarded because over 20% trials in a given condition (10 out of 50 trials) were of poor quality.

### 2.2. Stimuli

Norms for name agreement, as well as for other attributes for the pictures and picture names were obtained from the database of Morrison et al. (1997). Images were selected to be of high or low name agreement whilst matched on other attributes (see Table 1). Stimuli were selected to yield 50 low name agreement items, 50 high name agreement items, eight practice trials and 30 fillers (see Appendix for a list of words used in the experiment). Corresponding pictures were chosen, 114 from original pictures of Snodgrass and Vanderwart (1980), eight from Snodgrass and Vanderwart pictures redrawn by Morrison et al. (1997) and 16 from additional drawn pictures by Morrison et al. (1997). The original pictures of Snodgrass and Vanderwart (1980) were digitally scanned into TIF format documents and processed into clear line-drawings with high contrast. Morrison et al.'s additional pictures were downloaded from Morrison's online source, <http://www.cf.ac.uk/uwcc/psych/morrison/> (no longer available at this address). All pictures were saved as bitmaps (see Fig. 1 for examples). On-screen size of all pictures was smaller than 3 cm horizontally and 3.5 cm vertically. The DPI (dots per inch) value for all pictures was 72 pixels/inch.

From the original database (Morrison et al., 1997), the 50 low name agreement ('LNA') items had name agreement percentages between 50% and 87% (Mean = 76.0%, SD = 10.2%). The 50 high name agreement ('HNA') items all had 100% name agreement. Independent *t*-tests revealed that they differed significantly in name agreement, but not in other attributes (see Table 1). We additionally checked whether the three sets of images (LNA, HNA, fillers) differed from each other in terms of low-level picture attributes. Using data from a previous study by Laws and Gale (2002) we established two intrinsic measures for each image: (1) the proportion of black pixels, and (2) the internal complexity of each image—the latter measure Laws and Gale (2002) found to vary systematically for different categories of the Snodgrass pictures. Data were available for the majority of our pictures (LNA 42/50, HNA 37/50, fillers 24/30). Using MANOVA to

**Table 1**

Attributes of stimulus pictures and picture names (for details, see Morrison et al., 1997). SD for each attribute is presented in brackets.

	Low name agreement images and their names	High name agreement images and their names	Independent <i>t</i> -tests: two-tailed significance
Name agreement	76.0% (0.102)	100% (0)	.001
Objective AoA (months)	65.0 (27.9)	59.3 (24.6)	.28
Rated AoA	2.70 (0.606)	2.77 (0.626)	.56
Word frequency	2.54 (0.708)	2.56 (1.01)	.90
Familiarity	2.85 (0.834)	2.73 (0.808)	.50
Visual complexity	3.03 (0.713)	2.95 (0.741)	.57
Picture–name agreement	4.52 (0.277)	4.59 (0.226)	.15
Number of phonemes	4.40 (1.55)	4.50 (1.55)	.75
Number of syllables	1.64 (0.693)	1.64 (0.693)	.99

Note. AoA = age of acquisition.

compare the two image attributes simultaneously across the three picture groups, we found no difference between the image sets, Wilk's Lambda  $F(4, 198) = 1.06, p = .378$ . We concluded that the low-level attributes of the image sets did not differ systematically.

Eight practice trials were devised, four with low name agreement images and four with high name agreement images. A further 30 filler (catch) trials contained images in which the names and pictures presented did not match.

## 2.3. Procedure

### 2.3.1. ERP experiment

Stimuli were presented using the E-prime package (version 1.2, Psychology Software Tools, Inc., Pittsburgh, PA) on a 17 in. CRT screen. The visual angles were moderated to be less than 7° horizontally and 8° vertically (Holcomb and McPherson, 1994). Participants sat in front of the monitor and read the instructions, then pressed the space bar to initiate each trial. A practice session with eight trials was followed by an experimental session.

Each trial commenced with the presentation of a small black fixation cross in the center of the monitor (duration 1500 ms). Subsequently, a picture was shown on the screen for 1000 ms and participants were asked to name the picture covertly (i.e. silently, to themselves) as soon as it was presented. After picture offset, the most common name for the picture was presented visually, with the question 'Same name?' underneath, which, as the participants had previously been informed, prompted them to decide whether the name on the screen exactly matched the word in their heads. Participants responded 'Yes' with their left hand or 'No' with their right hand by keypress. Participants were asked not to blink their eyes or to move any part of their body during the time the pictures were on the screen.

All trials (50 HNA items, 50 LNA items and 30 fillers) were presented once in random order. For high and low name agreement items, it was highly likely that the expected name shown on the screen would be the same as that the participants had named covertly, so most answers would be 'Yes'. For this reason, filler (catch) trials were introduced, so that participants had to think before responding.

Participants pressed the space bar to start each trial and they could take a break whenever necessary by not pressing the space bar. The whole recording session took approximately 10 min.

### 2.3.2. EEG recording

Electrophysiological (EEG) signals were collected from the scalp with an Electrical Geodesics GSN 200 sensor net system with 128 channels (Electrical Geodesics, Inc., Eugene, OR), amplified by the EGI NetAmps 200 high impedance amplifier with a bandpass of 0.1–100 Hz, and digitized at a sample rate of 250 Hz. The threshold for impedance was set at 50 k $\Omega$  and all sites were recorded with a vertex reference.

Electrophysiological signals were filtered with a 40 Hz low-pass filter. In order to correct the polar average reference effect (PARE), a PARE-corrected reference was used (Junghofer et al., 1999), which was computed from the average of the entire surface of the scalp. Subsequently, individual trials were labeled as bad because of eye blinks or movements (EOG was recorded from 6 electrodes: 8, 26, 125, 126, 127, 128; an eye-blink or eye-movement is identified when amplitude is over 70  $\mu$ V in any of the six electrodes) or because of more than 10 bad channels (defined as having an average amplitude over 200  $\mu$ V or surpassing a differential threshold of over 100  $\mu$ V, compared to a 10-sample running average). In the final data analysis, 95.4% of trials were good in the HNA group and 95.2% in the LNA group (good trials rates ranged from 90% to 100%). Bad channels were replaced by the interpolation of good channels in proximity to the bad ones according to the spherical spline algorithm (Srinivasan et al., 1996). After bad channel replacement, all segments for each condition of each participant were averaged individually. Finally, ERPs were baseline corrected using a 200 ms pre-stimulus interval. The epoch length was 1200 ms.

### 2.3.3. Post-test for source of name disagreement

The purpose of the post test, which took place immediately after ERP recording, was to classify images according to the source of name disagreement. Accordingly, only low name agreement pictures were used. Participants were instructed to press a key to indicate the relationship between a picture and a written name. In each trial, a low name agreement picture was presented with its name (according to Morrison et al., 1997), and two choices were given: (1) 'This is ONE OF the names of the object' (i.e., the disagreement source is alternative names for the object depicted (AN subgroup); or (2) 'The picture COULD BE the object', or the picture COULD show something else' (i.e., the disagreement source is the uncertainty of the picture, the PU subgroup). Participants were given four practice trials, followed by the experimental session.

## 2.4. Identification of ERP components

Previous studies have shown that cognitive processes which might be affected by name agreement can be localized to several brain regions (e.g., Levelt et al., 1998). However, we chose not select a small and spatially specific number of electrodes for analysis because EEG has a relatively low spatial resolution. Instead, we chose to focus on the temporal dimension and settled for grouping sets of electrodes into six broad clusters, distributed across the scalp. The bilateral clusters were labeled 'frontal', 'parietal', and 'occipital', corresponding to their approximate location on the scalp (see Fig. 2). Note that we do not mean to claim that activity observed in a particular cluster necessarily reflects a generator source directly beneath, or even in the brain region with the same name (e.g., parietal cortex).

ERPs to high and low name agreement pictures, averaged over each cluster, are shown in Fig. 3. As can be seen in Fig. 3, in parietal and occipital clusters, following picture onset ERP waveforms reached a positive peak around 120 ms (P1), followed by a negative peak around 170 ms (N1). After N1, ERP waveforms quickly went positive, peaking around 220 ms (P2), followed by a second small negative peak around 290 ms (N2). After N2, ERPs in occipital clusters showed a P3 peaking at around 400 ms. This P3 was also discernable in parietal clusters, but the duration was longer and the peak was less distinct than that in occipital cluster, although mean amplitude across the time window was higher in the former. In frontal clusters, ERP waveforms were reversed from those seen in parietal and occipital clusters, with peaks with opposite polarity occurring at approximately the same time points as those in parietal and occipital clusters. These components seem compatible with those observed by Greenham et al. (2000). Peaks corresponding to our P1, P2 and N2 are also present in the Greenham data, with very similar latencies to those observed in the present study. Peaks of these identified ERP components are also well in line with the time-course of picture naming described by Levelt and colleagues (e.g., Indefrey and Levelt, 2004; Levelt et al., 1998, 1999).

## 2.5. Analysis of ERP components

For early ERP components (P1, N1 and P2), mean amplitudes for all electrodes in each cluster in a 50 ms time window around each peak were extracted, that is, 100–150 ms for P1, 150–200 ms for N1 and 200–250 ms for P2. For later ERP components (N2, P3 and the late time window), mean amplitudes in 100 ms time windows were extracted, namely 250–350 ms for N2, 350–450 ms for P3. Finally, a window of 800–900 ms was included to cover the late, sustained difference that was observed between HNA and LNA conditions.

The analytic approach was as follows. In each time window, ERP mean amplitudes were compared using repeated-measures analyses of variance with three variables: Name Agreement (high, low), Hemisphere (left, right) and Cluster (frontal, parietal and occipital). If this three-way ANOVA showed any reliable main effect or interaction involving Name Agreement (with alpha set to .05), further ANOVAs

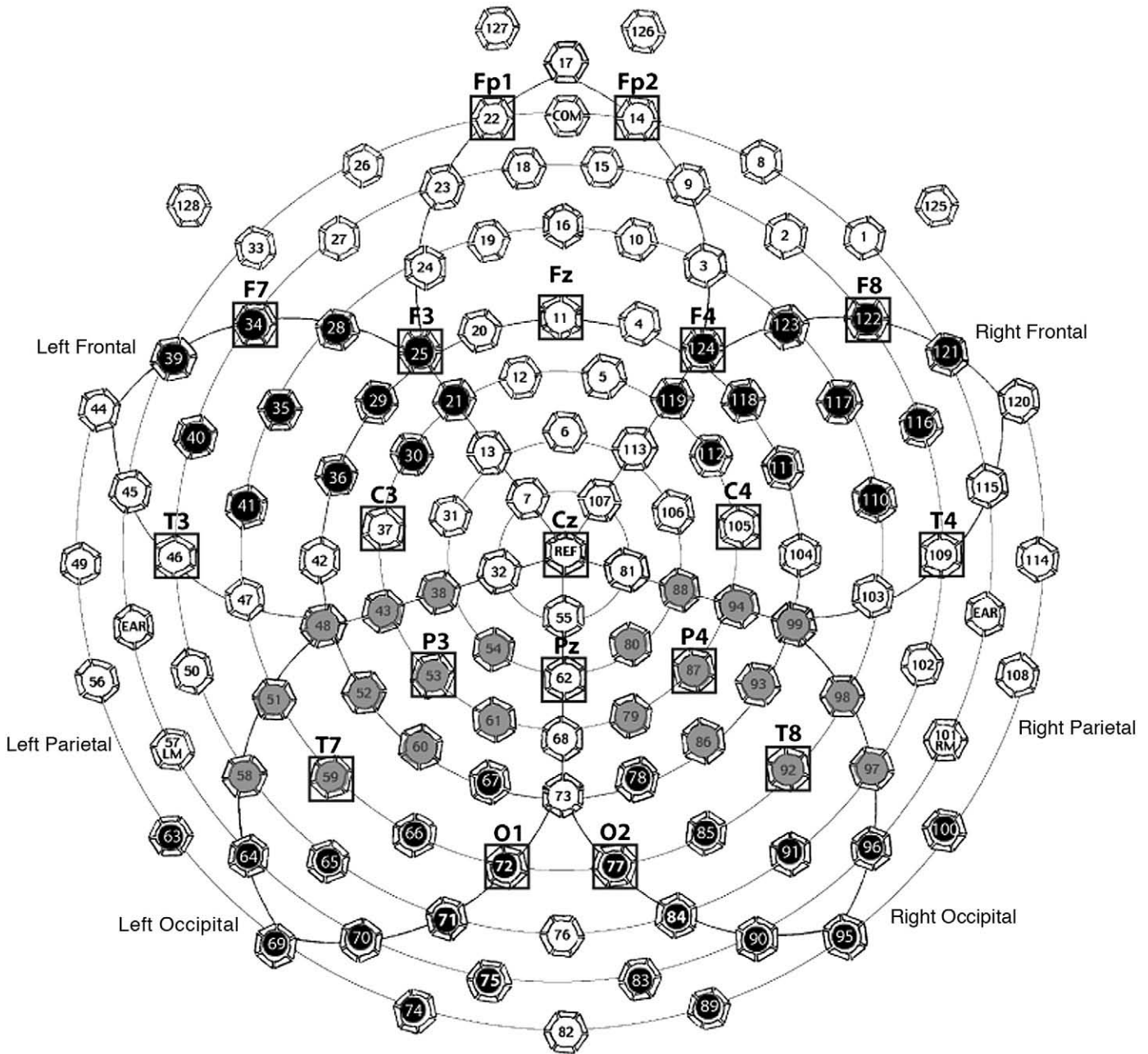


Fig. 2. Electrode clusters in the 128 electrode array. Approximate corresponding locations in the 10–20 system are shown. Six clusters of electrodes were selected a priori as regions of interest for the data analyses, namely left and right frontal, left and right parietal, and left and right occipital: these are shown in black on the diagram.

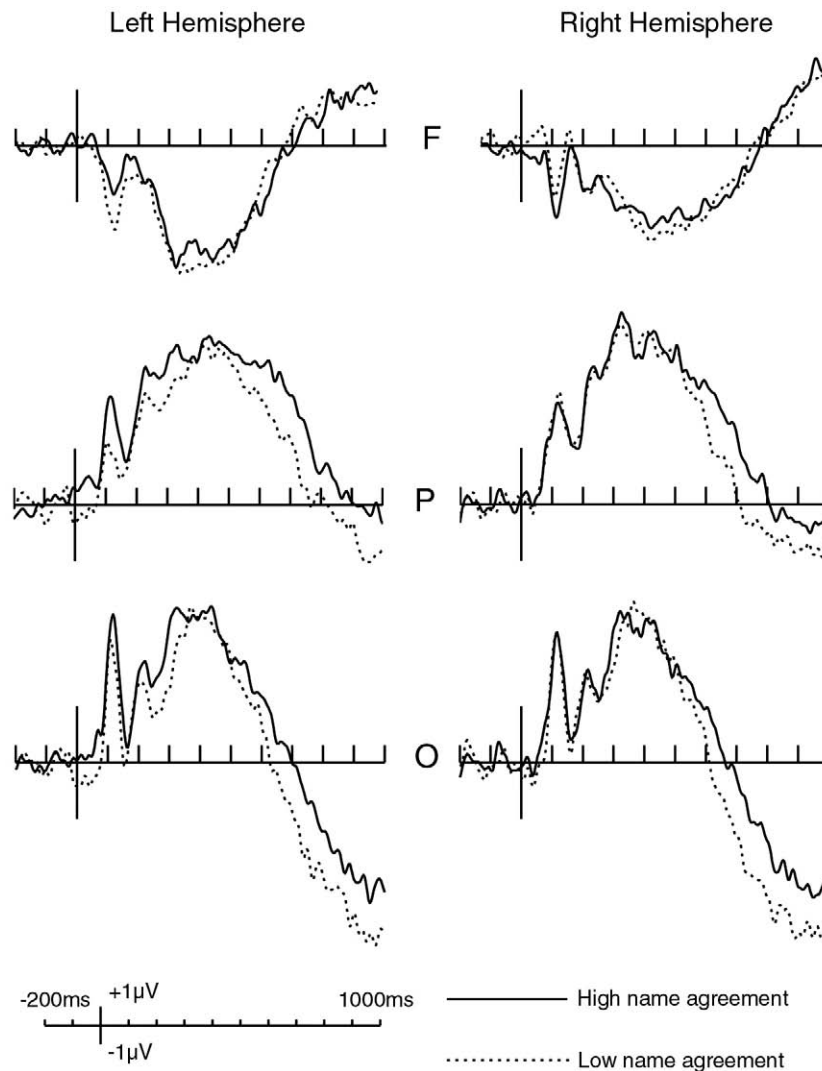
were carried out with fewer variables (i.e., Name Agreement and Hemisphere in three separate analyses for Cluster, or Name Agreement and Cluster in two separate analyses for Hemisphere). To check simple main effects of Name Agreement, paired *t*-tests comparing ERPs in the two Name Agreement conditions were carried out. All repeated-measures ANOVAs were subjected to Greenhouse-Geisser adjustment. Alpha values in all post-hoc comparisons arising from one-way ANOVA were adjusted by Bonferroni correction.

### 3. Results and discussion

#### 3.1. Behavioral data

In all valid datasets, there were only two wrong responses in filler trials (one each from two participants). In high name agreement trials, the mean number of 'no' responses (i.e., that the name we

provided did not match the internally produced name) was 2.37 ( $SD = 1.95$ , 0–5). In low name agreement trials, the mean number of 'no' responses was 11.1 ( $SD = 5.18$ , 0–20). 'No' responses were significantly less likely to high name agreement ('HNA') pictures than to low name agreement ('LNA') pictures,  $t(19) = -9.18$ ,  $p < .001$ . We calculated the actual name agreement for each picture in the current study according to participants' responses in the post-test. The average actual name agreement for HNA pictures was 95.4% (6.78%, 78.9%–100%) and for LNA pictures was 77.9% (19.3%, 31.6%–100%). Although the average actual name agreement for HNA pictures was not 100% as it had been from the Morrison et al. (1997) database, an independent *t*-test, adjusted for inequality of variance, confirmed that name agreement was significantly higher for HNA than LNA pictures,  $t(61.0) = 6.05$ ,  $p < .001$ , which was statistically consistent with the planned name agreement comparison. To avoid unbalanced numbers of pictures in the two groups and different items across participants,



**Fig. 3.** ERPs to pictures with different levels of name agreement (see text). The waveforms in this figure show the average ERP in each cluster, averaged over electrodes in the cluster. Vertical whiskers at the start of each epoch represent 1  $\mu$ V.

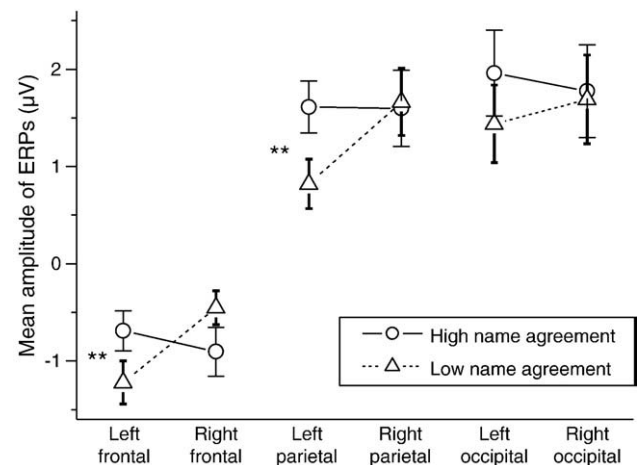
ERPs to all pictures were used in data analyses, irrespective of behavioral responses.

### 3.2. ERP data

#### 3.2.1. P1 (100–150 ms from picture onset)

For the P1 component, the three-way ANOVA showed a significant main effect of Name Agreement,  $F(1, 18) = 6.47, p < .03$ , and a significant main effect of Cluster,  $F(1.18, 21.3 [2, 36]) = 12.5, p < .002, \epsilon = .592$ . A two-way interaction between Name Agreement and Hemisphere was significant,  $F(1, 18) = 12.9, p < .003$ . The three-way interaction was also significant, Name Agreement  $\times$  Hemisphere  $\times$  Cluster,  $F(2, 36) = 3.94, p < .04$ . No other significant main effects or interactions were found ( $ps > .09$ ) (Fig. 4).

In frontal clusters, main effects of Name Agreement and Hemisphere were both non-significant,  $ps > .05$ . The interaction between Name Agreement and Hemisphere was significant,  $F(1, 18) = 14.1, p < .004$ . In paired  $t$ -tests, ERPs to high name agreement pictures ( $-0.689 \mu$ V) were more positive than ERPs to low name agreement pictures ( $-1.22 \mu$ V) in the left frontal cluster,  $t(18) = 3.03, p < .03$  (corrected for 4 comparisons), while they did not differ in the right frontal cluster,  $p > .20$  (corrected for 4 comparisons). In addition, ERPs



**Fig. 4.** The interaction between Name Agreement, Hemisphere and Cluster in P1. Amplitudes of P1 to High Name Agreement ('HNA') pictures were more positive than those to Low Name Agreement ('LNA') pictures in left frontal and parietal clusters, but not in the right hemisphere. Amplitude of P1 to HNA pictures in occipital clusters did not differ from amplitude of P1 to LNA pictures in either hemisphere. Error bars denote one standard error and \*\*\* denotes that significance is less than 0.05.

**Table 2**  
Summary of ANOVA effects on ERP components involving Name Agreement.

ERP component	Hypothesized processing stage (according to Indefrey and Levelt, 2004; see main text)	ANOVA effects involving Name Agreement	Simple effects (significant after correction)
P1	Object recognition	NA main effect; NA×Hemisphere; NA×Hemisphere×Cluster	High NA>low NA in left F and P; left<right for low NA in F and P
N1	Lemma selection	–	–
P2	Lemma selection	–	–
N2	Lemma selection/phonological encoding	NA main effect; NA×Hemisphere (P only)	High NA>low NA over left P left<right for low NA in P
P3	Phonological encoding	–	–
Late wave	Phonological encoding	NA main effect	High NA>low NA

Notes: NA: name agreement; F: frontal; P: parietal; O: occipital.

to LNA pictures were less positive in the left frontal ( $-1.22 \mu\text{V}$ ) cluster than in the right frontal cluster ( $-0.453 \mu\text{V}$ ),  $t(18) = -3.80$ ,  $p < .005$  (corrected for 4 comparisons), while ERPs to HNA pictures in the left frontal cluster and in the right frontal cluster were not different from each other,  $p > .90$  (corrected for 4 comparisons).

In parietal clusters, the main effect of Name Agreement was significant,  $F(1, 18) = 4.77$ ,  $p < .05$ , and the main effect of Hemisphere was non-significant,  $p > .09$ . The interaction between Name Agreement and Hemisphere was significant,  $F(1, 18) = 11.9$ ,  $p < .004$ . In paired  $t$ -tests, ERPs to HNA pictures ( $1.61 \mu\text{V}$ ) were more positive than ERPs to LNA pictures ( $0.819 \mu\text{V}$ ) in the left parietal cluster,  $t(18) = 4.75$ ,  $p < .005$  (corrected for 4 comparisons), while they did not differ in the right parietal cluster,  $p > .90$  (corrected for 4 comparisons). In addition, ERPs to LNA pictures were less positive in the left parietal cluster ( $0.819 \mu\text{V}$ ) than in the right parietal cluster ( $1.66 \mu\text{V}$ ),  $t(18) = -3.21$ ,  $p < .03$  (corrected for 4 comparisons), while ERPs to HNA pictures in the left parietal cluster and in the right parietal cluster were not different from each other,  $p > .90$  (corrected for 4 comparisons).

In occipital clusters, main effects of Name Agreement and Hemisphere were both non-significant,  $ps > .15$ . The interaction between Name Agreement and Hemisphere was significant,  $F(1, 18) = 5.70$ ,  $p < .03$ . However, in paired  $t$ -tests, ERPs to HNA pictures did not significantly differ from ERPs to LNA pictures in either occipital cluster,  $ps > .15$  (corrected for 4 comparisons), and ERPs in the left occipital cluster did not differ from ERPs in the right occipital cluster,  $p > .90$  (corrected for 4 comparisons).

In summary, for P1, effects of Name Agreement were confined to frontal and parietal clusters. Over both these areas, differences between ERPs to high and to low name agreement pictures were confined to the left hemisphere; and differences between hemispheres were confined to low name agreement pictures (see Table 2).

### 3.2.2. N1 (150–200 ms from picture onset)

For the N1 component, the three-way ANOVA did not show any significant main effects or interactions,  $ps > .05$ . However, the lack of a topographical specificity of this component might have been caused by the use of electrode cluster averages in the analyses.

### 3.2.3. P2 (200–250 ms from picture onset)

For the P2 component, the three-way ANOVA showed a significant main effect of Cluster,  $F(1.22, 22.0 [2, 36]) = 16.7$ ,  $p < .001$ ,  $\epsilon = .612$ . No other main effects or interactions were significant,  $ps > .06$ .

### 3.2.4. N2 (250–350 ms from picture onset)

For the N2 component, the three-way ANOVA showed three significant main effects: Name Agreement,  $F(1, 18) = 6.43$ ,  $p < .03$ , Hemisphere,  $F(1, 18) = 13.7$ ,  $p < .003$ , and Cluster,  $F(1.30, 23.5 [2, 36]) = 31.4$ ,  $p < .001$ ,  $\epsilon = .652$ , respectively. Most importantly, mean amplitude to HNA pictures ( $1.19 \mu\text{V}$ ) was less negative than that to LNA pictures ( $0.91 \mu\text{V}$ ). Two two-way interactions were significant, Name Agreement×Hemisphere,  $F(1, 18) = 5.37$ ,  $p < .04$ , and Hemi-

sphere×Cluster,  $F(2, 36) = 9.44$ ,  $p < .002$ . The first of these interactions is depicted in Fig. 5. The Name Agreement×Cluster interaction was non-significant.

In frontal clusters, only the main effect of Hemisphere was significant,  $F(1, 18) = 21.9$ ,  $p < .001$ .

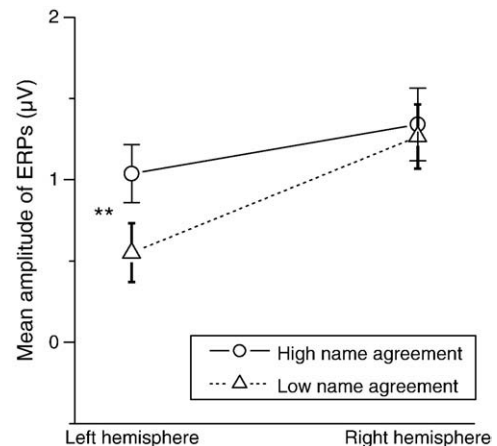
In parietal clusters, the main effect of Hemisphere was significant,  $F(1, 18) = 8.57$ ,  $p < .01$ , and the interaction between Name Agreement and Hemisphere was significant,  $F(1, 18) = 4.44$ ,  $p < .05$ . In paired  $t$ -tests, ERPs to HNA pictures ( $2.51 \mu\text{V}$ ) were less negative than to LNA pictures ( $1.96 \mu\text{V}$ ) in the left parietal cluster,  $t(18) = 5.01$ ,  $p < .004$  (corrected for 4 comparisons), while they did not differ in the right parietal cluster,  $p > .90$  (corrected for 4 comparisons). In addition, ERPs to LNA pictures were more negative in the left parietal ( $1.96 \mu\text{V}$ ) than in the right parietal cluster ( $2.76 \mu\text{V}$ ),  $t(18) = -3.29$ ,  $p < .03$  (corrected for 4 comparisons), while ERPs to HNA pictures in the left parietal cluster and in the right parietal cluster were not different from each other,  $p > .40$ .

In occipital clusters, there were no significant main effects or interactions,  $ps > .07$ .

In summary, for N2, effects of Name Agreement were confined to parietal clusters. Over this area, and similarly to P1 effects, differences between ERPs to HNA and to LNA pictures were confined to the left hemisphere; and differences between hemispheres were confined to low name agreement pictures (see Table 2).

### 3.2.5. P3 (350–450 ms from picture onset)

For the P3 component, the three-way ANOVA showed a significant effect of Cluster,  $F(1.21, 21.9 [2, 36]) = 29.1$ ,  $p < .001$ ,  $\epsilon = .607$ , and a significant two-way interaction between Hemisphere and Cluster,



**Fig. 5.** The interaction between Name agreement and Hemisphere in N2. Amplitudes of N2 to High Name Agreement pictures were less negative than those to Low Name Agreement pictures in the left hemisphere, while they did not show a reliable difference in the right hemisphere. Error bars denote one standard error.

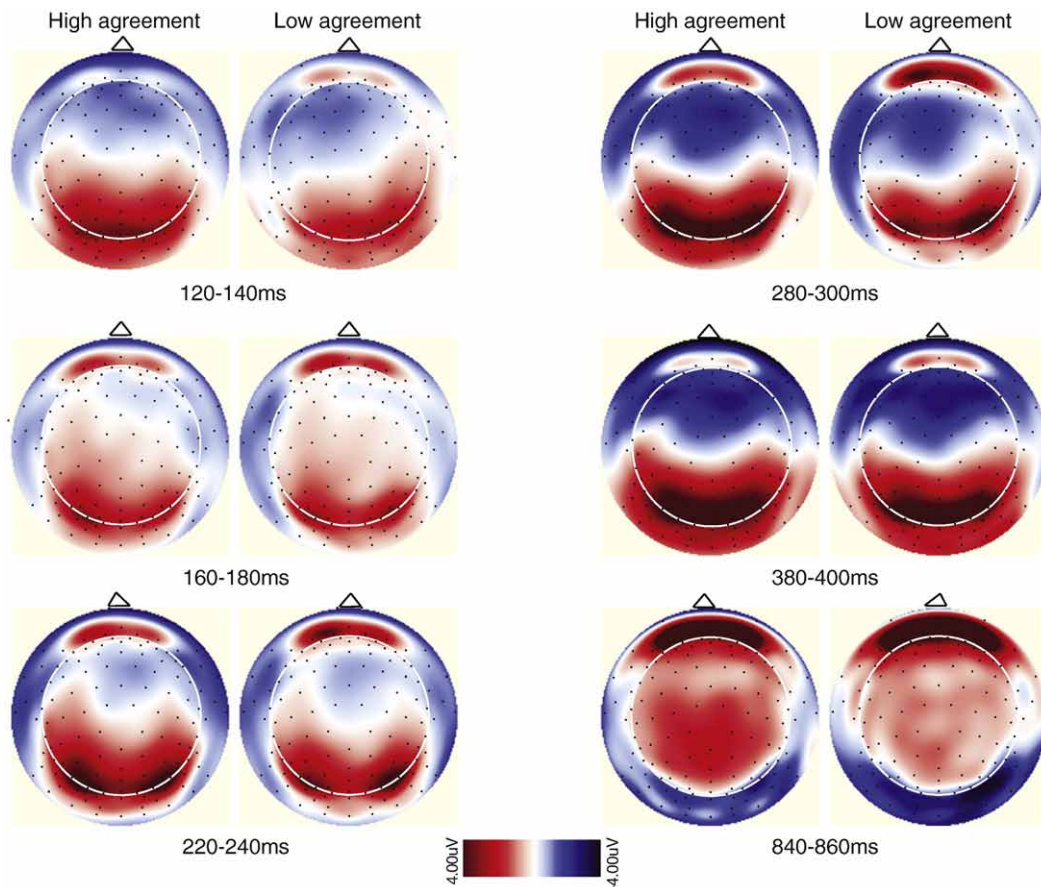


Fig. 6. A topographic map of ERP response to covert picture naming, as it unfolds following picture onset.

$F(2, 36) = 5.41, p < .02$ . No other main effects or interactions were significant,  $p > .07$ .

### 3.2.6. 800–900 ms after picture onset

The three-way ANOVA showed a significant main effect of Name Agreement,  $F(1, 18) = 6.94, p < .02$ , and a significant main effect of Cluster,  $F(1.50, 27.4 [2, 36]) = 23.3, p < .001, \epsilon = .760$ . Mean amplitude to HNA pictures ( $-0.32 \mu\text{V}$ ) was more positive than that to LNA pictures ( $-0.78 \mu\text{V}$ ). No other main effects or interactions were significant.

The results of these analyses are summarized in Table 2. A topographic map based on ERP potentials to covert picture naming, as it unfolds following picture onset, is presented in Fig. 6. In the topography, some positivity is shown on electrodes along the forehead, which can be attributed to the eyes. Importantly, this positivity was dissociated from our significant effects because the topography shows that the positivity for LNA pictures was stronger than for HNA pictures, but on the electrodes included in our analyses the mean amplitudes of P1 and N2 for HNA pictures were more positive than those for LNA pictures.

It remains to consider the source of the name agreement effects identified in Table 2. Recall, that after the ERP experiment, we asked participants to allocate low name agreement items to one of two subgroups: 'Picture Uncertainty' ('PU') and 'Alternative Names' ('AN'). Participants differed between each other in their responses in the post-test for the source of name disagreement; that is, in the way they allocated items to the PU and AN subgroups. These subgroups therefore varied across participants, precluding their use in the foregoing analysis. The mean number of items allocated to the PU subgroup was 17 ( $SD = 5.07, 6-24$ ), compared with 33 for the AN subgroup ( $SD = 5.07, 26-44$ ). Although these numbers differ, and items in each subgroup

necessarily vary between participants, the subgroups may nonetheless be relevant in the preliminary investigation of the source of a P1 and N2 name agreement difference observed in frontal and parietal clusters. Using planned comparisons, we compared mean P1 amplitude for each stimulus subgroup (Picture Uncertainty, 'PU'; Alternative Names, 'AN') with P1 amplitude to the HNA pictures<sup>2</sup>. In the left frontal cluster, P1 amplitude to HNA pictures was more positive than to the PU subgroup ( $-0.689 \mu\text{V}$  vs.  $-1.41 \mu\text{V}$ ),  $F(1, 18) = 9.78, p < .02$ , but did not differ reliably from the AN subgroup ( $-1.15 \mu\text{V}$ ),  $F(1, 18) = 4.71, p > .05$ . In the left parietal cluster, P1 amplitude to HNA pictures was more positive than to the PU subgroup ( $1.61 \mu\text{V}$  vs.  $0.699 \mu\text{V}$ ),  $F(1, 18) = 18.7, p < .001$ , and more positive than the amplitude to the AN subgroup ( $1.61 \mu\text{V}$  vs.  $0.868 \mu\text{V}$ ),  $F(1, 18) = 15.6, p < .002$ . For the N2 component, in the left parietal cluster, N2 amplitude to HNA pictures did not differ from the PU subgroup ( $2.51 \mu\text{V}$  vs.  $2.24 \mu\text{V}$ ),  $F(1, 18) = 1.63, p > .40$ , but was more positive than that to the AN subgroup ( $2.51 \mu\text{V}$  vs.  $1.86 \mu\text{V}$ ),  $F(1, 18) = 11.6, p < .01$ .

## 4. General discussion

The present study is the first electrophysiological investigation of the effect of name agreement during covert picture naming (in which

<sup>2</sup> We recognize that this analysis involves averages computed for different numbers of epochs. Nonetheless, the averages constitute valid estimates of the peak amplitudes to the stimuli, albeit prone to noise and to confounds introduced by unbalancing the careful matching of stimulus attributes described in the Method section. We offer these comparisons as additional data, rather than as definitive results.



images are seen and named silently 'in the head'). We found that, following picture onset, a P1 occurs at around 120 ms, followed by an N1 at around 170 ms, a P2 at around 220 ms, followed by an N2 at around 290 ms; finally a P3 showed at around 400 ms in clusters over the parietal and occipital areas. The peaks of the ERP waveform thus appear very similar to the average ERP waveform reported in the study of Greenham et al. (2000) where a similar paradigm, covert picture naming, was used.

Our prime motivation was to observe and report on differences between ERP waves during covert naming of pictures with low and with high name agreement. We were also interested in whether the apparent timing of name agreement phenomena might suggest where further targeted work should be done.

We observed an early main effect of Name Agreement, and an early interaction between Name Agreement, Hemisphere, and Cluster in the P1 time window. In the left frontal and parietal clusters, P1 to HNA pictures was more positive than P1 to LNA pictures, while there was no such difference in the right parietal cluster. For LNA pictures, amplitude in the left frontal and parietal clusters was lower than for HNA pictures. Somewhat later, another main effect of Name Agreement and an interaction with Hemisphere was found in the N2 time window. In the left parietal cluster, N2 to HNA pictures was less negative than N2 to LNA pictures, while no such difference was observed in the right parietal cluster. Similarly, N2 amplitude to LNA pictures was lower in the left parietal cluster than in the right parietal cluster. Finally, we observed a main effect of Name Agreement in the late time window. We will first consider potential low-level attentional explanations for these data, before discussion of mechanisms more specific to the lexical–semantic system.

The two early ERP components, P1 and N1, have been well documented in previous literature on attention (e.g., reviewed in Herrmann and Knight, 2001; Hillyard and Anllo-Vento, 1998). In the visual attention literature, P1 typically peaks around 120 ms after stimulus presentation (Heinze et al., 1994). It may reflect a facilitation of early sensory processing for stimuli presented at an attended location (cited in Herrmann and Knight, 2001, p469; Luck et al., 1990). In the present study, the task demanded that attention should invariably be located on the single presented picture, but such early sensory processing might conceivably have differed between high and low name agreement pictures, given our finding of a name agreement effect in P1. Hillyard and Anllo-Vento (1998) have demonstrated stronger P1 at posterior sites, and N1 at frontal sites, to attended stimuli as opposed to unattended stimuli. Similarly, in the present study, the P1/N1 results may reflect deployment of more attention to LNA than HNA pictures. This is plausible in light of our finding that for all participants, some LNA pictures were of uncertain identity, and hence might require immediate attentional resources for identification. Although some ERP studies of visual attention suggest that early effects may relate to low-level sensory processing, e.g. P1 and N170 modulated by luminance detectability (such as, Luck et al., 1994) and image contrast (Itier and Taylor, 2004), because we found no evidence for a systematic bias in the low-level properties of the pictures in the two groups, this low-level sensory processing account for our data is excluded.

In terms of mechanisms specific to the lexical–semantic system, one potential mechanism for ERP differences between HNA and LNA pictures relates to how many concept networks are activated by a given picture. For the former, a single candidate object is presumably activated. In contrast, for the latter, the picture's ambiguity resides in the fact that it might depict more than one object, or may have more than one name, and thus the LNA images may each activate more than one conceptual representation. It may therefore be tentatively argued that P1 differences over frontal and parietal clusters in the present study arise because of differences in the number of concept networks activated by given images. Limited support for this idea comes from Abdel Rahman and Sommer (2008), who recently reported that in an overt object naming task, a stronger P1 was evoked when participants

named pictures of which they had minimal knowledge, compared with pictures of objects of which they had in-depth knowledge. Perhaps the early P1 over frontal areas in the present study indexes early conceptual knowledge effects in object identification, i.e., differential top-down effects for low versus high name agreement images. We must acknowledge that this is unlikely to be the whole story, however, because the study of Abdel Rahman and Sommer (2008) showed the P1 effect over occipital, rather than—as in our case—frontal and parietal sites. Abdel Rahman and Sommer concluded that their data might be explained by either (1) top-down effects or (2) categorization effects emerging in perception. If our finding of early frontal and parietal effects are taken at face value, and if they reflect anterior rather than posterior cerebral activation, then top-down effects are the more likely explanation, but more work clearly needs to be done before this conclusion can be drawn with certainty.

Following the P1 and N1 was the P2 component, which is unfortunately not well documented in previous literature. Therefore we hesitate to link it to a specific cognitive process in the present study; and in any case we did not observe any modulation of P2 by name agreement.

The other component reliably modulated by name agreement in the present study was the N2. In the present study, the N2 peaked at 290 ms, that is, at the very beginning of the phonological encoding stage of the LRM model (Levelt et al., 1999). The stage that this ERP name agreement effect occurred seems later than the lexical selection stage, 150–275 ms post picture onset suggested in behavioral studies (reviewed in Indefrey and Levelt, 2004) and 150–225 ms post picture onset suggested in a MEG study (Maess et al., 2002). Maess et al.'s (2002) MEG study investigated the timing of neuronal processes related to category-based lexical competition. In an overt picture naming task, participants named a series of pictures which were either within a category (transport, clothing, etc.) or from different categories. In behavioral tasks, participants show longer naming latencies for same-category pictures than for ones from different categories (Kroll and Stewart, 1994). In Maess et al.'s study, the MEG difference between the two types of picture sequences peaked between 150 ms and 225 ms, apparently reflecting competition amongst co-activated semantic networks. However, in the study of Maess et al. (2002), participants study all pictures with their names before testing sessions and they saw each picture for 48 times during the testing session, while participants in the present study saw each picture once only and were asked to silently name pictures immediately. Therefore, the earlier effect in the study of Maess et al. (2002), compared with the effect in the present study, may reflect a strong repetition effect.

Another possibility for the relatively late effects is that name agreement effects pose a particularly difficult problem for the lexical–semantic selection system. For example, the difference between 'train' and 'bicycle' (or their German equivalents) on the one hand (as in the Maess et al.'s study), and 'cooker' and 'stove', or 'mushroom' and 'toadstool' on the other (as in our study), is considerable. That is, in our study, the conceptual networks associated with each lemma are much more similar to each other than those in studies such as that of Maess et al. (2002). Thus, lemma selection may be particularly delayed in the present study, which might in principle allow phonological processes, if they have commenced in parallel with lemma selection processes, to play a part in selection. In naming pictures with low name agreement, the strong similarity of concepts which are, *ex hypothesi*, in competition with one another appears to have utility in stressing the lexical–semantic system in interesting ways.

In support of the idea that name agreement effects may stress the lexical–semantic system and delay semantic information until the start of phonological encoding, van Turennout et al. (1997) noted that although semantic processing precedes phonological processing, 'there may be some overlap between the final part of the semantic stage and the start of phonological encoding' (p. 802). In a similar go/no-go LRP design to that employed by van Turennout et al. (1997),

Abdel Rahman and Sommer (2003) were able to show that, if the semantic system is stressed—by making semantic classification relatively hard—semantic and phonological processing appear to overlap (see also Abdel Rahman et al., 2003).

Although the name agreement effects in the present study showed at the beginning of phonological encoding stage of the LRM model, the cognitive processes invoked by LNA and HNA pictures might be expected to be different well in advance of phonological encoding: This is true for both kinds of low name agreement image, i.e., those which have low name agreement because of picture uncertainty ('PU'), and those which have low name agreement due to alternative names ('AN'). To choose a specific name for an LNA picture from the PU subgroup, competition between a number of different *concepts* must be resolved. To choose a specific name for an LNA picture from the AN subgroup, competition between a number of different *lemmas* must be resolved<sup>3</sup>. Because N2 peaks in the phonological encoding stage of the LRM model, the effect of name agreement arguably persists into the phonological encoding stage. One account for this pattern might be that, in the lemma selection stage, several lemmas are activated but only during the phonological encoding stage is the system finally forced to select a single name from all activated lemmas. But this account is at odds with the mechanism suggested by the LRM model, in which lexical selection takes place during lemma selection, rather than phonological encoding. Analysis of later parts of the waveform, during the P3 and during the 800–900 ms window, did not provide much further insight. While a sustained effect of Name Agreement was present in the latter window, the P3 failed to show any effect of this kind.

As discussed above, our experiment was carefully set up to balance attributes of LNA and HNA pictures, and not primarily to examine the cognitive *source* of the name agreement effect(s). However, the PU and AN subgroups may nonetheless be relevant to the source of the N2 difference we observed in frontal and parietal clusters. The lack of difference between N2 amplitude for HNA and PU pictures suggests that the cognitive system may have resolved issue of picture identity by this point, leaving no residual difference between these two image types. Our tentative finding—that the difference in N2 amplitude between HNA and AN pictures was reliable, but the difference in N2 amplitude between HNA and PU pictures was not—tends to support the idea that it is lemma selection which is driving the N2 component. Of course, confirmation would require more careful manipulation of the source of name agreement than in the present study. However, the idea that our N2 component is related to lemma selection is consistent with the finding of Jescheniak et al. (2002) that phonological codes for a depicted object's name (or in our case, possible names) are not automatically activated, but are activated only when its name is to be produced.

In our treatment of the differential effects of the source of name agreement effects on N2 we appear to have generated a paradox: in our consideration of the timing of the N2 peak, we argued that effects of name agreement differences appeared to persist into the period of, and might (*pace* LRM) be resolved by phonological encoding processes. But Jescheniak et al. (2002) concluded that phonological codes for a depicted object's name are only activated if that object is to be named. These assertions can only *both* be correct if multiple phonological codes *are* obligatorily activated in the case where the picture is unambiguous ('uni-categorical') but nonetheless has multiple lemmas associated with it. Again, more definite data are required, but the utility of studying name agreement in this context is clear.

Furthermore, in parietal clusters, hemispheric asymmetry of P1 appears to have been caused by both uncertainty of identification of the images themselves and by those pictures having more than one name (ERPs to HNA pictures were more positive than both subgroups in the left parietal cluster). In contrast, in the left frontal cluster, P1

amplitude to HNA pictures was more positive than for the PU, but not the AN subgroup. To the extent that the AN pictures differed from PU pictures—in that the former each represent a single concept with multiple names—the difference in P1 effects in frontal versus parietal areas might suggest that different cognitive processes were measured by these clusters. Neuroimaging with greater spatial resolution than ERP is required to investigate this further. However, our results for the first three 'stages' of processing are in line with the behavioral data of Vitkovitch and Tyrell (1995) that name agreement effects tend to occur because of differences between the high name agreement pictures and those from a 'Picture Uncertainty' subgroup in the early stage of picture naming, and then between the high name agreement pictures and those from an 'Alternative Names' subgroup at or after lexical selection.

Although we have found results which are consistent with previous studies, a methodological issue in the present study is worth noting. We matched the two groups of pictures very carefully on several dimensions i.e., high and low name agreement pictures did not differ in Objective AoA (months), Rated AoA, Word frequency, Familiarity, Visual complexity, Picture–name agreement, Number of phonemes, Number of syllables, or on the two low-level image attributes of proportion of black pixels, and internal complexity of each image. However, some studies have shown that other dimensions of a word may affect word production as well, such as, neighborhood frequency (e.g., Vitevitch and Sommers, 2003), morphology (e.g., Roelofs and Baayen, 2002), and semantic transparency (e.g., Dohmes et al., 2004). Therefore it is still possible that besides name agreement effects, other effects from unknown attributes may affect the present findings. Future experiments concerning effects from a specific attribute of a word should consider as many attributes as possible.

In conclusion, in the present study we obtained significant effects of name agreement on ERP responses, while participants covertly named images presented one at a time on a screen in front of them. During covert picture naming, in the first stage, ascribed to processes of object recognition and concept activation, we found name agreement effects, and also a hemispheric asymmetry, associated with differences between ERPs to pictures with low name agreement and pictures with high name agreement, for P1s in frontal and parietal clusters. A strong name agreement effect was also found in N2s, tentatively ascribed to the beginning of the phonological encoding stage. It may be concluded that there are two stages during picture naming in which name agreement effects occur, one in object recognition, due to the activation of one or more concept networks, and the other either towards the end of the lemma selection process, or at the start of phonological encoding, due to several activated lemmas from a single activated concept network.

Additional (and necessarily provisional) analyses suggested that the early name agreement effect and hemispheric asymmetry predominantly originated from images in the Picture Uncertainty subgroup, rather than images in the Alternative Names subgroup. In contrast, it is probable that pictures in Alternative Names subgroup contributed to the strong name agreement effect in N2. To confirm the modulation by name agreement of more than one cognitive process in picture naming, a partial replication of the present study is required, in which pictures in the PU and AN groups are at least as closely matched as were the pictures in the LNA and HNA groups in the present study.

## Acknowledgements

This work was supported by an Overseas Studentship from the University of Reading to the first author.

We thank Prof. Keith Laws, University of Hertfordshire, for kindly providing low-level image data derived from Snodgrass and Vanderwart's images. We also thank three anonymous reviewers for their valuable comments on previous manuscripts.

<sup>3</sup> Of course, as pointed out by Vitkovitch and Tyrell (1995), it is also possible that each lemma has its own associated conceptual network, 'upstream' from lemma selection (i.e., distinct conceptual representations underlying each alternative name).

## Appendix A. Experimental words and filler pictures lists

### Low name agreement words

Ant	Barn	Bear	Bee	Beetle
Boat	Bow	Broom	Brush	Bus
Cooker	Deer	Diamond	Dice	Dustbin
Eagle	Flask	Glasses	Gorilla	Gun
Handbag	Helicopter	Hen	Jacket	Jumper
Leopard	Lettuce	Lightbulb	Lips	Lorry
Mitten	Monkey	Mouse	Mushroom	Necklace
Needle	Nut	Ostrich	Peach	Peg
Pineapple	Pliers	Potato	Rocket	Scales
Sledge	Suitcase	Tiger	Web	Wheel

### High name agreement words

Acorn	Anchor	Arrow	Axe	Bell
Book	Cake	Camel	Cannon	Carrot
Coat	Crown	Dress	Drum	Elephant
Flag	Grapes	Guitar	Harp	Iron
Jelly	Mermaid	Microphone	Nun	Onion
Scarecrow	Screwdriver	Pipe	Plug	Pumpkin
Seahorse	Shirt	Snail	Snake	Squirrel
Stool	Strawberry	Sun	Swan	Telescope
Toaster	Torch	Tortoise	Van	Violin
Waistcoat	Whale	Whistle	Windmill	Yo-yo

### Fillers

Accordion	Ball	Bowl	Butterfly	Cat
Chair	Comb	Dog	Ear	Finger
Flower	Fox	Hand	House	King
Ladybird	Lemon	Nurse	Owl	Peacock
Pencil	Pepper	Pram	Ruler	Scissors
Snowman	Sword	Tractor	Vase	Witch

## References

- Abdel Rahman, R., Sommer, W., 2003. Does phonological encoding in speech production always follow the retrieval of semantic knowledge?: Electrophysiological evidence for parallel processing. *Cognitive Brain Research* 16 (3), 372–382.
- Abdel Rahman, R., Sommer, W., 2008. Seeing what we know and understand: how knowledge shapes perception. *Psychonomic Bulletin & Review* 15 (6), 1055–1063.
- Abdel Rahman, R., Sommer, W., Schweinberger, S.R., 2002. Brain-potential evidence for the time course of access to biographical facts and names of familiar persons. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28 (2), 366–373.
- Abdel Rahman, R., Van Turenout, M., Levelt, W., 2003. Phonological encoding is not contingent on semantic feature retrieval: an electrophysiological study on object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29 (5), 850–860.
- Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U.H., Segui, J., 2004. Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers* 36 (1), 140–155.
- Barry, C., Morrison, C.M., Ellis, A.W., 1997. Naming the Snodgrass and Vanderwart pictures: effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology* 50A, 560–585.
- Bonin, P., Chalard, M., Meot, A., Fayol, M., 2002. The determinants of spoken and written picture naming latencies. *British Journal of Psychology* 93, 89–114.
- Bonin, P., Peereman, R., Malardier, N., Meot, A., Chalard, M., 2003. A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers* 35, 158–167.
- Cuetos, F., Ellis, A.W., Alvarez, B., 1999. Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers* 31, 650–658.
- Dell, G.S., 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93, 283–321.
- Dell'Acqua, R., Lotto, L., Job, R., 2000. Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers* 32, 588–615.
- Dohmes, P., Zwitserlood, P., Bölte, J., 2004. The impact of semantic transparency of morphologically complex words on picture naming. *Brain and Language* 90, 203–212.
- Ellis, A.W., Morrison, C.M., 1998. Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 24, 515–523.
- Eulitz, C., Hauk, O., Cohen, R., 2000. Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. *Clinical Neurophysiology* 111 (11), 2088–2097.
- Glaser, W.R., 1992. Picture naming. *Cognition* 42, 61–105.
- Greenham, S.L., Stelmack, R.M., Campbell, K.B., 2000. Effects of attention and semantic relation on event-related potentials in a picture–word naming task. *Biological Psychology* 55, 79–104.
- Heinze, H.J., Mangun, G.R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T.F., Gos, A., Scherg, M., Johannes, S., Hundeshagen, H., et al., 1994. Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* 372, 543–546.
- Herrmann, C.S., Knight, R.T., 2001. Mechanisms of human attention–event-related potentials and oscillations. *Neuroscience and Biobehavioral Reviews* 25 (465–476).
- Hillyard, S.A., Anlo-Vento, L., 1998. Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the USA* 95, 781–787.
- Holcomb, P., McPherson, W., 1994. Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition* 24, 259–276.
- Humphreys, G.W., Riddoch, M.J., Quinlan, P.T., 1988. Cascade processes in picture identification. *Cognitive Neuropsychology* 5, 67–103.
- Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Itier, R.J., Taylor, M.J., 2004. Face recognition memory and configural processing: a developmental ERP study using upright, inverted, and contrast-reversed faces. *Journal of cognitive neuroscience* 16 (3), 487–502.
- Jeschienak, J.D., Schriefers, H., Garrett, M.F., Friederici, A.D., 2002. Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *Journal of Cognitive Neuroscience* 14 (6), 951–964.
- Jeschienak, J.D., Hahne, A., Schriefers, H., 2003. Information flow in the mental lexicon during speech planning: evidence from event-related brain potentials. *Cognitive Brain Research* 15, 261–276.
- Johnson, C.J., 1992. Cognitive components of naming in children: effects of referential uncertainty and stimulus realism. *Journal of Experimental child psychology* 53 (1), 24–44.
- Johnson, C.J., Paivio, A., Clark, J.M., 1996. Cognitive components of picture naming. *Psychological Bulletin* 120, 113–139.
- Junghofer, M., Elbert, T., Tucker, D.M., Braun, C., 1999. The polar average reference effect: a bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology* 110, 1149–1155.
- Kempen, G., Huijberts, P., 1983. The lexicalization process in sentence production and naming: indirect election of words. *Cognition* 14, 185–209.
- Koester, D., Schiller, N.O., 2008. Morphological priming in overt language production: electrophysiological evidence from Dutch. *NeuroImage* 42 (4), 1622–1630.
- Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language* 33 (2), 149–174.
- Lachman, R., Shaffer, J.P., Hennrikus, D., 1974. Language and cognition: effects of stimulus codability, name–word frequency and age of acquisition on lexical reaction time. *Journal of verbal learning and verbal behavior* 13, 613–625.
- Laganaro, M., Morand, S., Schneider, A., 2009. Time course of evoked-potential changes in different forms of anomia in aphasia. *Journal of cognitive neuroscience* 21 (8), 1499–1510.
- Laws, K.R., Gale, T.M., 2002. Category-specific naming and the 'visual' characteristics of line drawn stimuli. *Cortex* 38, 7–21.
- Levelt, W.J.M., 1989. *Speaking: From Intention To Articulation*. MIT Press, Cambridge.
- Levelt, W.J.M., 1991. Lexical access in speech production: stages vs cascading. In: Peters, H.F.M., Hulstijn, W., Startweather, C.W. (Eds.), *Speech Motor Control and Stuttering*. Elsevier Science, New York.
- Levelt, W.J.M., Praamstra, P., Meyer, A., Helenius, P., Salmelin, R., 1998. An MEG study of picture naming. *Journal of Cognitive Neuroscience* 10, 553–567.
- Levelt, W.J.M., Roelofs, A., Meyer, A., 1999. A theory of lexical access in speech production. *Behavioral & Brain Sciences* 22, 1–75.
- Luck, S.J., Heinze, H.J., Mangun, G.R., Hillyard, S.A., 1990. Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and clinical neurophysiology* 75 (6), 528–542.
- Luck, S.J., Hillyard, S.A., Mouloua, M., Woldorff, M.G., Clark, V.P., Hawkins, H.L., 1994. Effects of spatial cuing on luminance detectability: psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance* 20 (4), 887–904.
- Maess, B., Friederici, A.D., Damian, M.F., Meyer, A., Levelt, W., 2002. Semantic category interference in overt picture naming: sharpening current density localization by PCA. *Journal of cognitive neuroscience* 14 (3), 455–462.
- Morrison, C.M., Chappell, T.D., Ellis, A.W., 1997. Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 50A (3), 528–559.
- Rodriguez-Fornells, A., Schmitt, B.M., Kutas, M., Munte, T.F., 2002. Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia* 40, 778–787.
- Roelofs, A., Baayen, H., 2002. Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin & Review* 9 (1), 132–138.
- Schmitt, B.M., Schiltz, K., Zaake, W., Kutas, M., Munte, T.F., 2001. An electrophysiological analysis of the time course of conceptual and syntactical encoding during tacit picture naming. *Journal of cognitive neuroscience* 13 (4), 510–522.

- Snodgrass, J.G., Vanderwart, M., 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory* 6, 174–215.
- Snodgrass, J.G., Yuditsky, T., 1996. Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments & Computers* 28, 516–536.
- Srinivasan, R., Nunez, P.L., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1996. Spatial sampling and filtering of EEG with spline-Laplacians to estimate cortical potentials. *Brain Topography* 8, 355–366.
- Strijkers, K., Costa, A., Thierry, G., 2010. Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex* 20 (4), 912–928.
- van Turennout, M., Hagoort, P., Brown, C.M., 1997. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition* 23, 787–806.
- van Turennout, M., Hagoort, P., Brown, C.M., 1998. Brain activity during speaking: from syntax to phonology in 40 ms. *Science* 280, 572–574.
- Vitevitch, M.S., Sommers, M.S., 2003. The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition* 31 (4), 491–504.
- Vitkovitch, M., Tyrell, L., 1995. Sources of disagreement of object naming. *The quarterly journal of experimental psychology* 48A (4), 822–848.