Guessing What's on Your Mind: Using the N400 in Brain Computer Interfaces

Marijn van Vliet, Christian Mühl, Boris Reuderink, and Mannes Poel

University of Twente, Human Media Interaction, Enschede 7522 NB, NL
w.m.vanvliet@student.utwente.nl

Abstract. In this paper, a method is proposed for using a simple neurophysiological brain response, the N400 potential, to determine a deeper underlying brain state. The goal is to construct a BCI that can determine what the user is 'thinking about', where 'thinking about' is defined as being primed on. The results indicate that a subject can prime himself on a physical object by actively thinking about it during the experiment, as opposed to being shown explicit priming stimuli. Probe words are presented that elicit an N400 response which amplitude is modulated by the associative relatedness of the probe word to the object the user has primed himself on.

1 Introduction

Brain Computer Interfaces (BCI) are devices that let a user control a computer program, without any physical movement. A BCI measures the activity within the brain directly, interprets it and sends a control signal to a computer. By actively or passively changing his brain activity, the user can send different control signals and by doing so, operate the system. The effectiveness of a BCI depends highly on the ability to measure relevant processes within the brain and the performance of the classification of the signal.

1.1 Low Level Brain Responses versus High Level Cognitive Processes

Without significant progress in recording and signal processing technology, brain-computer interfaces (BCI) that rely on electroencephalography (EEG) recordings only have access to basic, neurophysiological responses. Examples include the P300 response, event related (de)synchronization (ERD/ERS) and steady state visually evoked potentials (SSVEP). These directly measurable phenomena can be regarded as low level responses. These are manifestations of higher level cognitive processes that are more complex and cannot be directly measured, such as object recognition, intention of movement and visual processing. When measuring the low level responses, most information about specifics of the higher level processing is lost. However, the low level responses can give insight into the higher level brain processes by using probes in a search scheme. By using probes, different possibilities for the high level brain state can be tested until the correct one is found.

Y.Y. Yao et al. (Eds.): BI 2010, LNAI 6334, pp. 180–191, 2010.

One example of existing BCI that try to determine a deeper underlying brain state are those that try to determine the memories of the user, usually by exploiting the P300 effect [9]. The P300 potential is linked to an oddball task: when multiple stimuli are presented to the subject, the task related interesting stimulus will elicit a bigger P300 then a uninteresting one, due to a response triggered by increased attention due to recognition. This method allows the user to both consciously control which stimulus to select by focussing on a task and unconsciously be probed by stimuli that trigger a recognition response. The unconscious probes could for instance be used to determine whether a subject looked into a box containing some objects, or not. He would be instructed to look for images of birds in a collection of various photographs. Photographs of birds and those of objects in the box both elicit an enlarged P300 potential in relation to irrelevant photographs [1].

In this paper, the N400 potential is used as the low level brain response to determine a search scheme using probes to determine the high level brain state. While the P300 is related to attention, the N400 has been associated with semantic processing [6].

1.2 N400

The N400 potential was first discovered by Kutas et al. [5], who were analyzing the Event Related Potential (ERP) of subjects that were reading sentences. They studied the effect of adding words that did not make sense given the preceding ones in a sentence in order to get an insight into brain activity during semantic parsing. To this end, a set of sentences was created of which half the sentences ended in a semantically congruent word (e.g. *I drink coffee with milk and sugar*) and half the sentences ended incongruent (e.g. *I drink coffee with milk and socks*).

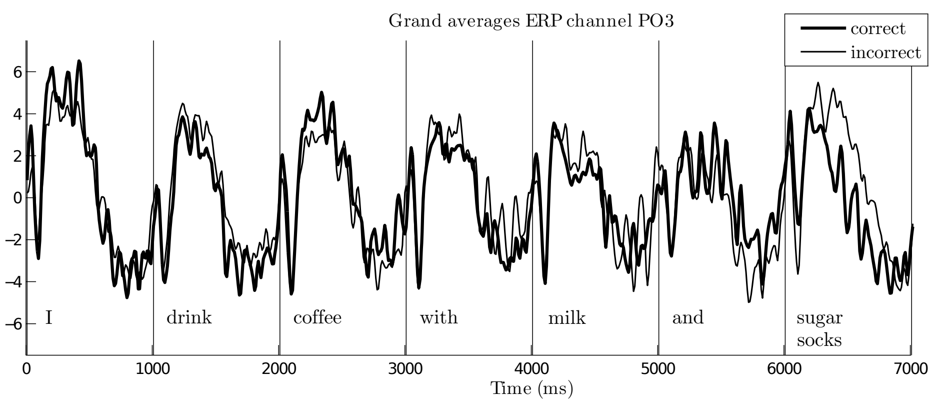


Fig. 1. Plot of an ERP recorded on a subject which was shown 7 word sentences. When the last word is shown, there is a distinctive difference between a word that lies in the line of expectation (labeled as *correct*) and a word that does not (labeled as *incorrect*).

The grand average ERP's, recorded at position PO3, of both classes are shown in fig. 1 for a subject, recorded at position PO3. Each second, a word is flashed on the screen and a recognizable series of ERP components reappears after the onset of each word. One of these components, which appears around 400ms after the word is shown, changes amplitude when the word is unrelated to the rest of the sentence and is called the N400 potential.

The N400 has been linked to the concept of priming. Priming is an improvement in performance in a perceptual or cognitive task, relative to an appropriate baseline, which is caused by previous, related experience. In semantic priming, the ability to decode a stimulus carrying meaning is improved by previous exposure to a stimulus with a related meaning. In the experiment described above, the subject is primed on the first six words of a sentence, and the seventh word is either semantically related to the prime or not. It was later discovered that the N400 effect not only occurs in sentences, but it is shown that a whole range of stimuli can be used. The underlying strategy is to first show a *prime* stimulus and a short time after, show a *probe* stimulus, where the prime and probe can for example be word–word, image–word, image–image or even sound–word pairs [2][10].

There are two competing theories as to the cause of the N400 potential [12]. The *integration* view states that the N400 is caused by a difference in difficulty integrating the symbol in a context. This theory is obviously in line with the results of the experiments of Kutas et al. where the last word had to be integrated with the rest of the sentence. The results with word pairs can be explained by regarding the first word as a context where the second word has to be integrated into. The *lexical* view states that the N400 is caused by a difference in the difficulty of long term memory access. According to the spreading activation model, the activation of a symbol from memory causes nearby symbols to preactivate, making subsequent access of these symbols easier. Findings in [8], in which fMRI is used, and [7], in which MEG is used to localize the N400 effect, suggest that the effect is primarily due to facilitated memory access, which states a case for the lexical view.

1.3 Goals of the Present Research

Like the P300, the N400 potential can give insight into high level brain processes by using probes. In this paper, the possibility is explored of using the N400 potential to determine which one out of several possible objects the user is thinking of. For example, to differentiate between the user thinking about a coffee mug or a tomato. The advantage of using the N400 over the P300 is that the N400 effect will also occur on stimuli that do not correspond completely with the object the subject is thinking of, but also on stimuli that are closely related. This could in the future allow for a system to deploy a binary search algorithm in order to find the target object, allowing for a much larger choice for the subject. For instance, the system can first try to determine whether the object is a living organism or not, before descending down the search tree, playing a BCI version of 20 questions[1] with the user.

When showing a probe word, the N400 effect can be used to detect whether the subject was primed on a stimulus related to the probe word or not. Current N400 research does not leave any choice to the subject as to which stimulus he will be primed, so the prime will always be known in advance. If however the subject is allowed to choose his prime and this choice can be detected, the N400 potential will be a useful feature to use in a BCI. This choice presents a problem when showing the priming stimulus: how can a system know which stimulus the subject wants to be primed on? Showing the stimuli corresponding to all the choices will most likely prime the subject for all of them, disallowing any choice. In this paper, a method is investigated in which the subject is not presented with a priming stimulus, but must achieve the proper priming by actively thinking about a physical object. When a user has primed himself on an object, showing probe words corresponding with all the possibilities and examining the N400 potential elicited by them may enable an automated system to determine which object the user was thinking about.

The ambiguous term 'thinking about' is now defined as 'being primed on'. Since the priming effect occurs for many different types of stimuli, such as words, images and sounds, the hypothesis that a subject can prime himself by being told to think about an object is assumed to be possible and is evaluated with an experiment. The goal of the experiment is to determine whether a subject can prime himself on an object without being shown a priming stimulus.

2 Method

Probe words have been prepared that correspond to one of two possible primes (e.g. a book or a bottle). The problem is how to convey the choices of prime to the subject. Telling him the choices may cause him to be primed on all of them. In the experiment, the subject was therefore not given a choice. He was given a physical object, such as a book, mug or tomato, to hold. Physical objects were chosen to not limit the subject's mind to a visual, auditory or lingual stimulus, but allow him to choose for himself how to go about thinking about the object. In order to promote the latter, two auditory beeps are played before showing the probe word. The subject is instructed to close his eyes on the first (low pitched) beep and concentrate on the object and to open his eyes on the second (high pitched) beep and look at the screen, where the probe word appears.

2.1 Participants

The experiment was performed on three participants aged between 23 and 28, all of which were males and native speakers of Dutch. All of them were right handed. They were placed in a comfortable chair in front of a desk with a computer screen and did not leave the chair for the duration of the experiment.

[1] A game in which the player is allowed 20 yes-or-no questions to determine what the opponent is thinking of. See also `http://www.20q.net`

2.2 Design

Each participant completed one session with consisted of three blocks. The procedure during a block is as follows:

1. An object is given to the subject, allowed to be held and placed before him on the table.
2. Instructor leaves the room.
3. 100 Trials were performed. 50 trials matching the object shown to the subject, 50 matching a different object, not shown to the subject.
4. Instructor enters the room.
5. 5-10 minute break.

Fig. 2 summarizes the way each trial was presented to the subject. Prior to each word, a low beep was heard. The subject was instructed to close his eyes when hearing this beep and think about the shown object. Two seconds later, a high beep followed. The subject was instructed to open his eyes when hearing this beep. A fixation cross appears and the subjects eyes are drawn to the center of the screen. This closing and opening of the eyes will produce a large EOG[2] artifact, which is dealt with in the signal processing step later on. After 2 seconds to prevent overlap with the artifact, a probe word replaces the fixation cross for 200ms. The next beep would sound 1800ms after that, prompting the subject to close his eyes again for the next trial.

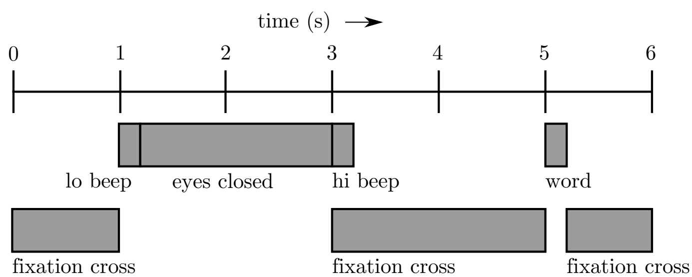


Fig. 2. Timeline of a single trial. Two beeps sound, one indicating the subject to close his eyes and think about the object, one indicating him to open his eyes.

2.3 Procedure

1. Subject is seated in a comfortable chair in front of a computer screen.
2. Subject is told about the goal of the experiment and given instructions on the procedure.
3. Subject is fitted with electrodes during the explanation of the experiment.
4. 10 Test trials were presented to acquaint the subject with the procedure.
5. Three blocks were performed.
6. End of experiment.

[2] Electro-OculoGraphy, in this context meaning the electric current produced by eye movements that shows up in the EEG recording.

2.4 Stimuli

The experiment consists of 3 blocks. In each block, the subject is given a physical object and 10 words are shown that match the given object and 10 words are shown that match a different object. Each word was included in the randomized sequence 5 times, in order to average the recordings of the 5 repetitions later on.

The words that were shown to the subject have to be closely related to one object, but not at all related to the other objects. To archive this, the Dutch Wordnet [13] was used. This wordnet is a graph with Dutch words as nodes, with semantic relations drawn between them as edges, such as synonyms, hyponyms, opposites etc.

The 6 physical objects are chosen to be exemplars which can be described with one word (e.g. a book, a mug, a tomato, etc.), hereafter called the object name. The goal is to associate 10 Dutch words with each object name. Candidate words are generated by traversing the Dutch Wordnet, starting at the object word o_1 (for instance 'book') and spreading outwards in a breadth-first fashion. For each word that is encountered in this fashion, the distance to each of the 6 object names is calculated and a score is calculated:

$$s = \left(\sum_{i=2}^{6} d(w, o_i) \right) - 5d(w, o_1) \tag{1}$$

Where s is the score, w is the word under consideration, o_i is one of the object names, with o_1 being the object name from which the search was started. $d(x, y)$ is the distance function between two words: the number of edges in the shortest path between them, which is used as a measure of relatedness of the words. Words with a high score will be close to the object we are searching from, but distant from any of the other objects. A search like this will generate many words that are either uncommon or not necessarily associatively related to the object. For each object name, a list of 30 words is created by sorting all the generated words by score, highest to lowest, and manually taking the first 30 words that were judged to be strongly associatively related to the object by the instructor. This subjective selection improves the effectiveness of the dataset, because the distance function used takes only purely semantic relatedness into account, whereas the N400 effect is also attributed to associative relatedness between words.

The total list of 180 words (30 words times 2 objects times 3 blocks) has been presented to each subject at least a week before the experiment. The subject was asked to score each word in relation to a photograph of the corresponding object. A 5 point scale was used, 1 being not related at all, 5 being practically the same thing. For each object, the 10 words that the subject scored highest where chosen to be used in the experiment. When choosing between words with the same score, the scores by the other subjects was taken into account and the words with a high score assigned by all subjects were favored. For example, fig. 3 lists the words that the first subject scored highest in relation to an object. The full list of words used during the experiment is included in appendix A.

Dutch original:

woordenboek	bibliotheek
bijbel	bladzijde
hoofdstuk	paragraaf
uitgever	verhaal
auteur	kaft

English translation:

dictionary	library
bible	page
chapter	section
publisher	story
author	cover

Fig. 3. left: a photograph of a sample object that was shown to the subjects. **right:** the 10 words marked as most related to the object in the photograph by the first subject.

2.5 Method of Analysis

A schematic overview of the data analysis method is presented in fig. 4. Numbers in the text below correspond to the blocks in the diagram.

The recordings were made with a 32 channel EEG cap and 3 external electrodes placed in the middle of the forehead and below each eye (1). All data was recorded with a samplerate of 256Hz, average referenced (2) and bandpass filtered between 0.3Hz and 30Hz (3). Trials were extracted on the interval $t = -3$s

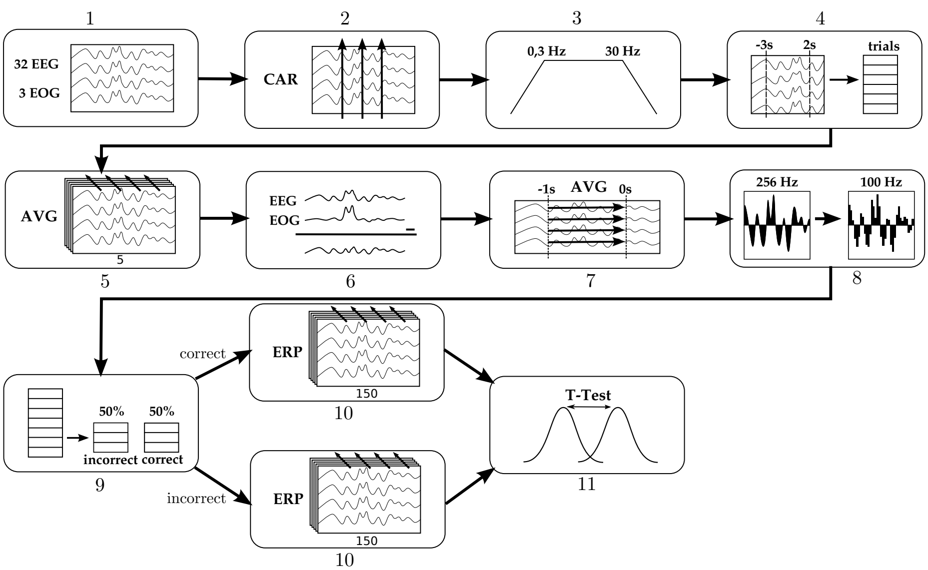


Fig. 4. Diagram summarizing the data analysis process. Each block has a number which correspond to the numbers in the text.

to $t = 2$s, relative to the onset of the probe word at $t = 0$s (4). This includes the moment the subject opens his eyes until the moment any N400 effects should no longer be visible. Each probe word occurred 5 times in the presentation sequence. The corresponding trials were averaged (5) to form the final trials. These averaged trials were filtered with an automated method for reducing EOG effects [11] which involves calculating a mixing matrix between the recorded EOG signal and the recorded EEG (6). Using this matrix, the EOG can be subtracted from the EEG, reducing the effect of eye movements in the data, which are severe but predictable in this experiment, since the subject was instructed to close and open his eyes before being shown a probe word. Application to the average of 5 trials instead of unaveraged data increases the effectiveness of this filtering [3]. Each trial was baselined on the interval $t = -1$s to $t = 0$s (7) and resampled to 100Hz (8) to reduce the number of data points. For each class, an Event Related Potential (ERP) plot was created (10). Student's t-tests were performed on each 10ms segment (11) between both classes to determine the statistical significance between any differences between the two classes.

3 Results

The resulting ERP plots are presented in fig. 5. It can be seen that starting around 400ms, the waveforms diverge between the classes for all subjects. From the topo plots can be seen that the location of the N400 effect differs for each subject. This could be explained by the fact that the subjects employed different strategies for concentrating on an object, ranging from visualizing it to thinking about related symbols. The ERP plots show that there is a dipole effect: the N400 is a positive deflection in relation to the baseline at frontal/right positions and a negative deflection in occipital/left positions. Variation of the N400 amplitude and timing is to be expected between subjects, as it is also the case in other studies (see for an example [4], figure 16). There are also differences in the duration of the effect, in the recordings for subject 3, the effect is measurable for more than a second, up until the subject closes his eyes again, while subject 1 only shows the effect for a few hundred milliseconds.

All data preprocessing steps are performed on the dataset as a whole, except for the baselining. It is possible that the separate calculation of the baseline for each class creates an artificial difference between the ERPs. This could for instance be the case when the baseline is calculated on a unstable portion of the signal containing lots of artifacts. Such periods exists in the recordings, where the subject opens and closes his eyes causing an EOG artifact. Calculating the baseline on the wrong portion of the signal will effectively generate a random baseline value for each class.

In order to rule this possibility out, the exact same data analysis was performed again, but the trials were assigned random class labels by shuffling them and assigning the first half the label *correct* and the last half the label *incorrect*. The result was that any differences between the classes that could be seen were not statistically significant and randomly distributed over channels and time. This bestows confidence in the method of analysis.

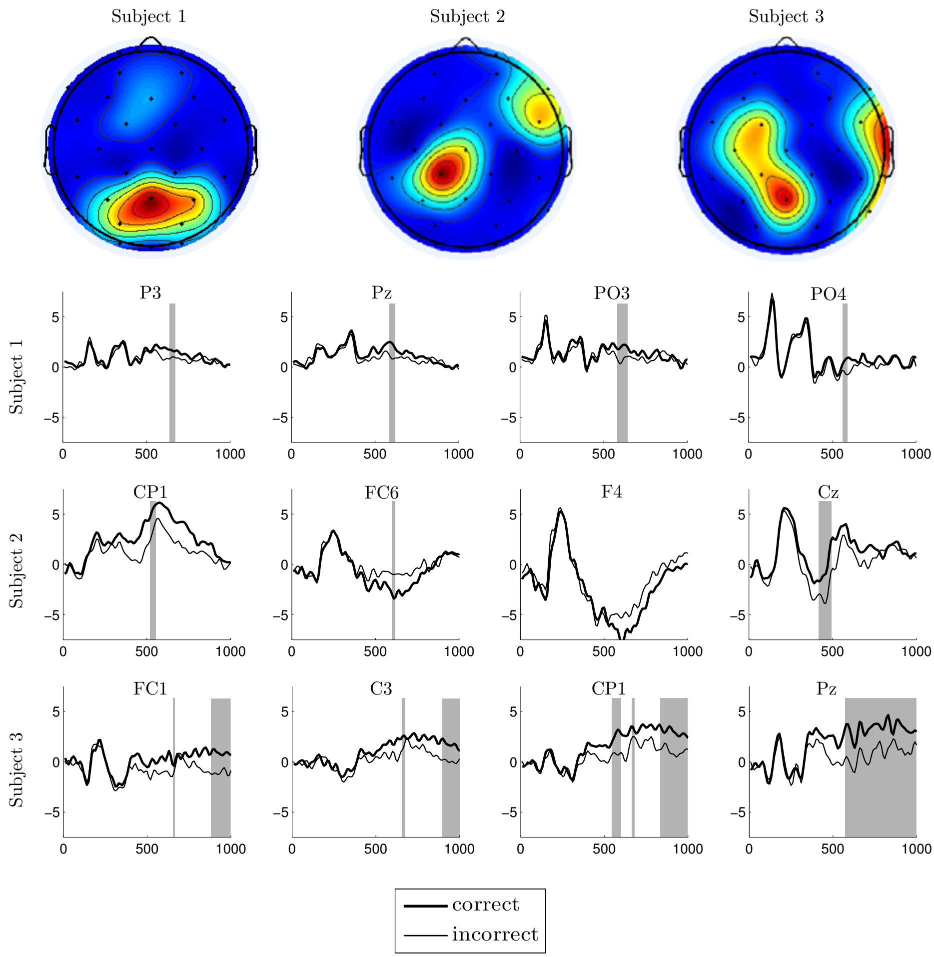


Fig. 5. Top: For each subject, the topo plot shows the mean significance for each channel during the time interval 500–1000ms. The values are given as 1/p-value, so a higher value means the difference between the classes is more significant. **Bottom:** ERP plots for all three subjects. Each row corresponds to a subject and the 4 most significant channels. Shaded area's indicates a statistically significant ($p \leq 0.01$) difference between the waveforms.

4 Conclusions

The purpose of the experiment was to explore whether a subject can be primed without being shown an explicit stimulus, such as a word, image or sound. The subject was instead asked to prime himself by actively thinking about an object when hearing a beep. Single probe words were used to trigger a N400 response when the word matched the previously shown object.

The recordings indicate that the N400 effect is indeed elicited using this strategy and so, the experiment gives support for the hypothesis that a subject can prime himself by thinking about an object in such a way that the N400 effect occurs when shown probe words. Evidence is given that priming can be achieved without using explicit stimuli, leaving choice for the subject to prime himself. Using this pilot experiment as a basis for further research, a BCI could be constructed which can guess the object that the user is primed on and by extension what it is the user is 'thinking about'.

5 Future Work

Many questions remain to be answered before the N400 signal can be reliably used in a BCI context as envisioned in this paper.

The decision to ask the subject to close his eyes, made signal processing considerably harder, because of the generated EOG artifacts. It was included to make it easier for the subject to concentrate and not be distracted by outside stimuli. In retrospect, this might have done more harm than good, so a future experiment can be performed to compare results without the closing of the eyes.

A first attempt has been made on trying to automatically classify the trials. A linear support vector machine was trained on the data segment corresponding to the 4 channels with the lowest average t-test scores (the most significance) and the time interval 400ms–600ms, resulting in 80 datapoints per trial. However, naive classification of this kind proved to be insufficient as the performance was around chance level.

After the experiment, the subjects were familiar with the words used, which can have an influence on their ability to trigger an N400 effect if there same dataset would be used again. Research can be done to determine the impact of these repetition effects. When constructing a BCI, care must perhaps be taken to present different probe words every time it is used by the same user. Improvements due to user training can also be explored.

Acknowledgements

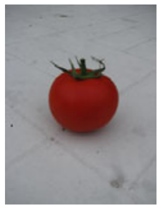
The authors gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture and Science. This work is also made possible by the academic-assistant project of the Royal Dutch Academy of Science, financed by the Dutch Ministry of Education, Culture and Science. We would also like to thank the test subjects for their efforts.

References

1. Abootalebi, V., Moradi, M.H., Khalilzadeh, M.A.: A new approach for EEG feature extraction in P300-based lie detection. Computer methods and programs in biomedicine 94(1), 48–57 (2009)

2. Bajo, M.T.: Semantic facilitation with pictures and words. Journal of Experimental Psychology: Learning, Memory, and Cognition 14(4), 579–589 (1988)
3. Croft, R.J., Barry, R.J.: EOG correction: a new aligned-artifact average solution. Electroencephalography and clinical neurophysiology 107(6), 395–401 (1998)
4. Hagoort, P., Brown, C.M., Swaab, T.Y.: Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. Brain: a journal of neurology 119, 627–649 (1996)
5. Kutas, M., Hillyard, S.A.: Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. Advancement of Science 207(4427), 203–205 (1980)
6. Kutas, M., Hillyard, S.A.: Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163 (1984)
7. Lau, E., et al.: A lexical basis for N400 context effects: evidence from MEG. Brain and language 111(3), 161–172 (2009)
8. Lau, E.F., Phillips, C., Poeppel, D.: A cortical network for semantics (de)constructing the N400. Nature reviews. Neuroscience 9(12), 920–933 (2008)
9. Meegan, D.V.: Neuroimaging techniques for memory detection: scientific, ethical, and legal issues. The American journal of bioethics: AJOB 8(1), 9–20 (2008)
10. Orgs, G., Lange, K., Dombrowski, J.H., Heil, M.: Conceptual priming for environmental sounds and words: an ERP study. Brain and cognition 62(3), 267–272 (2006)
11. Schlögl, A., et al.: A fully automated correction method of EOG artifacts in EEG recordings. Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology 118(1), 98–104 (2007)
12. Thompson-Schill, S.L., Kurtz, K.J., Gabrieli, J.D.E.: Effects of Semantic and Associative Relatedness on Automatic Priming. Journal of Memory and Language 38(4), 440–458 (1998)
13. Vossen, P., Bloksma, L., Boersma, P.: The Dutch Wordnet (1999)

A Words Used in Experiment 4

	zin	sentence		karton	cardboard
	bibliotheek	library		verhuizen	move
	trilogie	trilogy		bewaren	keep
	letter	letter		verpakken	package
	paragraaf	section		etiket	label
	hoofdstuk	chapter		inpakken	pack
	inleiding	introduction		tillen	lift
	voorwoord	foreword		magazijn	storehouse
	nawoord	epilogue		plakband	duct tape
	uitgever	publisher		stapelen	stack
	kaft	cover		opbergen	store
	woordenboek	dictionary		schoenendoos	shoe box
	bladzijde	page		opslag	storage
	verhaal	story		zolder	attic
	auteur	author		dragen	carry
	pint	pint		fruit	fruit
	glas	glass		groente	vegetable
	hals	neck		tortilla	tortilla
	kroonkurk	bottle cap		paprika	paprika
	alcohol	alcohol		peterselie	parsley
	biertap	beer		voeden	feed
	pul	pint		ovenschotel	oven dish
	doorzichtig	transparent		tros	bunch
	gezellig	merry		plant	plant
	fust	cask		voedsel	food
	bier	beer		lekker	delicious
	wijn	wine		ketchup	ketchup
	statiegeld	deposit		saus	sauce
	krat	crate		pizza	pizza
	flessenopener	bottle opener		groentenboer	greengrocer
				plukken	pick
				gerecht	dish
				maaltijd	meal
				salade	salad
	snuiten	blow		beker	cup
	servet	napkin		thee	tea
	snuiven	sniff		chocolademelk	chocolate
	wegdoen	put away		theelepel	teaspoon
	hoesten	cough		slurpen	slurp
	doekje	handkerchief		schenken	pour out
	niezen	sneeze		gieten	pour
	afvegen	wipe		bord	plate
	broekzak	pocket		oortje	ear
	neus	nose		onderzetter	coaster
	snotteren	snivel		koffie	coffee
	weggooien	throw away		breken	break
	niesbui	sneezing fit		keukenkast	cupboard
	opvouwen	fold		drinken	drink
				drank	beverage