

How Order of Label Presentation Impacts Semantic Processing: an ERP Study

Jelena Batinić (jelenabatinic1@gmail.com)

Laboratory for Neurocognition and Applied Cognition,
Department of Psychology, Faculty of Philosophy, University of Belgrade,
Čika Ljubina 18-20, 11000 Belgrade, Serbia

Andrej Savić (andrej_savic@etf.rs)

Tecnalia Serbia Ltd.
University of Belgrade, School of Electrical Engineering,
Bulevar Kralja Aleksandra 73, 11000 Belgrade, Serbia

Vanja Ković (vanja.kovic@f.bg.ac.rs)

Laboratory for Neurocognition and Applied Cognition,
Department of Psychology, Faculty of Philosophy, University of Belgrade,
Čika Ljubina 18-20, 11000 Belgrade, Serbia

Abstract

In this study, we wanted to investigate whether the processing of semantic information is easier when mapping names to pictures or is it the other way around. In order to test this hypothesis, we ran a behavioural and an ERP (Event Related Potential) study, with specific interest in the N400 component as an indicator of semantic processing. We compared three groups of participants who did a match/mismatch task with the only difference being that the labels would appear before, after or simultaneously with the pictures. Not surprisingly, the hardest condition was the one where the two information were presented simultaneously. The amplitude of the N400 was more prominent in the condition where labels were presented after the pictures in comparison to the condition where labels preceded picture presentation, suggesting that this second experimental situation led to smaller violation of expectation for our participants (word to picture condition) in comparison to mapping pictures to words.

Keywords: semantic processing; Event Related Potentials; N400; mental representations; word processing; picture processing

Introduction

We are in constant interaction with novel and familiar objects on a daily basis. When learning about an object for the first time, we examine its visual characteristics and associate them with its name.

In this study, we wanted to investigate whether the processing of semantic information is easier when mapping names (as more abstract representations) to pictures (as more specific representations) or is it the other way around. Given that the most informative component which is well known to be sensitive to semantic processing/integration is the N400 component it will be of our primary interest to test whether these mappings elicit differences in the N400 amplitude.

The discovery of the N400 component came in the now classical study of Kutas and Hillyard (1980) in which participants were presented with sentences (one word at a time), that ended with either congruent or

incongruent words. There were two types of incongruent endings: possible but improbable (She drinks tea with salt) or completely semantically unrelated to the previous context (She drinks tea with house). Incongruent words elicited a negative response at around 400 ms from the stimulus onset. Authors concluded that the N400 is sensitive to context and semantic anomalies (Kutas & Hillyard, 1980).

Since that study, different authors have reported finding the N400 in a variety of experimental tasks which required semantic processing such as match/mismatch task, semantic priming, word or picture recognition (Anderson & Holcomb, 1995; Boutonnet, & Lupyan, 2015; Ganis, Kutas, & Sereno, 1996; Holcomb & Anderson, 1993). In general, N400 is most prominent in the central and parietal regions of the scalp (Anderson & Holcomb, 1995; Kutas & Federmeier, 2011), but the topography changes depending on the experimental condition. For example, anterior regions are particularly active when processing pictures (Anderson & Holcomb, 1995). The latency of the component is usually in the time window of 200-600 ms from stimulus onset (Kutas & Federmeier, 2011). The most interesting characteristic of the N400 is its amplitude, given that it is most responsive to experimental manipulations, whereby a more negative amplitude is elicited by unexpected stimuli which are in turn harder to process (Kutas & Federmeier, 2009).

Semantic processing and the N400

Anderson and Holcomb used a semantic priming task in order to investigate the differences in processing of auditory and visually presented words (Anderson & Holcomb, 1995). Word pairs (prim and target) were presented in the same modality (visual or auditory) using different stimulus onset asynchronies SOA= 0 ms, 200 ms, 800 ms. N400 component was found in all of the experimental conditions, but lasted longer when word pairs were presented simultaneously (SOA=0 ms), suggesting that the processing of the two stimuli was

parallel. Apart from that, the highest error rates were in this experimental condition, supporting the hypothesis that it is harder to process two pieces of information at the same time.

Following a different line of research, Boutonnet and Lupyan (2015) were investigating whether visual processing of objects would be easier when they were primed with names or with nonverbal cues. In a match/mismatch task pictures of familiar animals and artifacts were preceded by their names or equally informative nonverbal cues (sound of dog barking preceding a picture of a dog). Participants were more successful when they were cued with words, and the authors suggest that this is because words denote categories and are better at evoking mental representations which facilitates responding both to match and mismatch trials.

If in fact words evoke more general and abstract mental representations, it would be interesting to see in what way do pictures, that always represent a specific exemplar, can influence the processing of an object's name.

As previously mentioned, words can be treated as more abstract representations and refer to entire categories of objects, while a picture is always representing a single instance of an object and therefore evokes a more narrow and specific mental representation (Ković, Plunkett, & Westermann, 2009; Ković, Plunkett, & Westermann, 2010). Given that, to our knowledge, there are no studies directly comparing word to picture versus picture to word processing it remains unclear whether these processes differ, and if they do, which one is easier.

In order to investigate this, we constructed an experiment in which we manipulated the order of label presentation, thereby contrasting three experimental conditions: words preceding pictures, pictures preceding words, and words and pictures presented together. This allowed us to compare the processes of mapping abstract (word) to specific (pictures) representations and specific to abstract representations by examining the amplitude of the N400 across conditions. Our hypothesis is that the hardest condition for our participants would be simultaneous presentation of words and pictures, given that they have to process two pieces of information at the same time (Anderson & Holcomb, 1995). Furthermore, we expect that the easiest condition, which would elicit the smallest negative response, would be the pictures to words condition. Since names evoke broad mental representations (Boutonnet & Lupyan, 2015), any picture shown after the label, no matter how typical of an exemplar it is, would most likely be somewhat different from our evoked mental representation which makes the task harder for the participant to respond. On the other hand, a picture can evoke only one name for a given object which makes the name easier to process when displayed after the picture.

Method

Participants

We tested sixty participants, twenty per experimental condition. Participants were psychology students at the University of Belgrade, all native Serbian speakers. They gave informed consent and received course credit for their participation. All participants reported normal or corrected-to-normal vision.

Stimuli

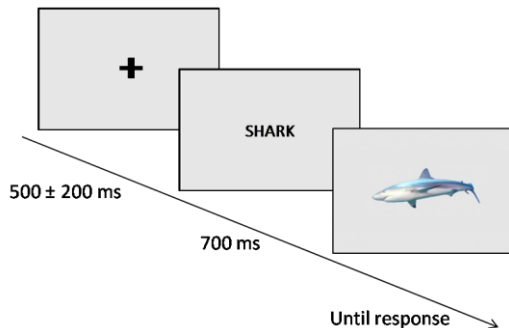
The study consisted of 120 familiar, everyday objects from different categories such as: mammals, fruits, furniture, tools, clothes, etc. These objects were represented by pictures (original stimuli list taken from Kovic et al., 2009) and their corresponding labels. Labels were presented visually in order to control the duration of stimuli presentation, which wouldn't be possible in the case of auditory presentation. All stimuli were pretested and qualified as highly typical and highly familiar objects. We also conducted a naming task in which 8 participants were asked to name the objects presented in the pictures in order to ensure that there was only one appropriate name for a given picture. Hence, only pictures that were named in the same way by every participant, were included in the study.

Experimental Design and Procedure

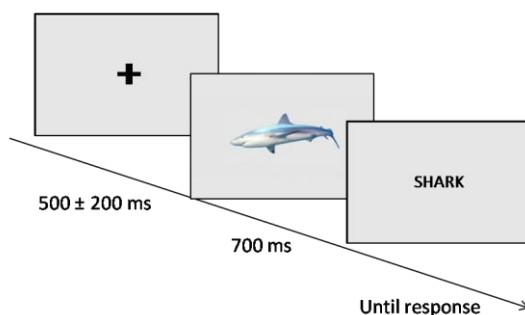
Participants completed 240 trials of a simple match/mismatch task. They were instructed to judge if the picture and the label represented the same object, and indicate their response by pressing one of two keys (*C* or *N*) on a keyboard (which were counterbalanced across participants). The number of match and mismatch trials was equal and the order of trials was randomized across participants. Depending on the experimental condition, participants were responding to *pictures* – when they were preceded by words (WP condition); *words* – when they were preceded by pictures (PW condition) or *words and pictures* when they were presented together (TO condition). The labels and pictures in the mismatch trials were from different categories and paired in a way to avoid phonological similarities and phonological onset competition (cat-cow); rhyme (dog-frog) as well as semantic association (cat-dog). Trials would start with a fixation cross, followed by a 700ms presentation of word, picture, or word and picture together (depending on the experimental condition, with the difference being that in TO condition the stimuli would last until response, not only 700ms) after which they would see a picture or a word on which they had to respond to. The time sequence of a single trial for each experimental group is presented in Figure 1. In order to avoid preparatory movement potentials during the task a jitter of ± 200 ms for the fixation cross was introduced (Luck, 2005). According to Luck (2005) expecting a stimulus that requires a response can cause preparatory movement

potentials that are known to appear as contingent negative variations (CNV), a low frequency negative wave preceding an expected stimulus.

A Word to Picture Condition



B Picture to Word Condition



C Together Condition

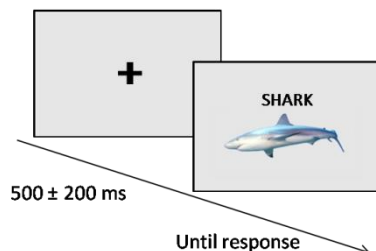


Figure 1: Time sequence of individual trials for all three experimental conditions.

The experiment was conducted in a Faraday Cage. The participants were sitting in front of a computer, at approximately one meter distance from the screen. The stimuli were presented on a grey background at the center of the screen at eye level. Participants were instructed to avoid frequent blinking and reduce muscle movement as much as possible.

ERP recordings

EEG signals were recorded continuously throughout the experiment. The signals were recorded from 15 electrodes placed at: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, PC5, PC6, T5, T6, O1 and O2 sites according to the international 10–20 standard. Two electrodes were placed on the earlobes as a reference, and the ground electrode was positioned on the participant's forehead.

PSYLAB EEG8 biological amplifier in combination with PSYLAB SAM unit (Contact Precision Instruments, London, UK) were used for EEG measurements. Skin-electrode contact impedance was below 5 k Ω at the beginning of the trials. EEG signal amplification was 20 k and hardware band-pass filtering over the range 0.03–40 Hz. Signals were sampled at 500 Hz using NI USB- 6212 (National Instruments, Austin TX) card for analog to digital signal conversion. For EEG signal acquisition and online display a custom software with graphical user interface developed in LabVIEW 2010 was used (National Instruments, Austin, TX, USA) (Savic, Malešević, & Popovic, 2013). For determining the exact moment of stimulus onset upon which we time-lock the ERPs a sensor for detecting changes in brightness was placed in the upper-left corner of the screen. The stimuli had a black square in the sensor area which was not visible to the participants. This allowed a precision of 1 ms for determining stimulus onset.

ERP processing

Offline EEG processing was conducted using custom routines in MATLAB (version 2010a, The Mathworks, Natick, MA, U.S.A.). EEG signals from all channels were filtered using a zero-phase 4th order Butterworth bandpass filter with 0.1–25 Hz cut-off frequencies. The high pass component of the filter removes near-DC drift and the low pass component filters out muscle artifacts and 50 Hz noise, along with related harmonics. Data were then segmented into epochs including 100 ms baseline prior to stimulus onset, 900 ms following stimulus onset. The baseline was corrected in all EEG channels by subtracting from each epoch the mean of a 100 ms interval prior to the stimuli onset. Epochs contaminated with ocular-movements and/or other artefacts were rejected from further analysis if absolute value of the signal from any of the channels exceeded a threshold manually determined for each subject within a range of 40–60 μ V (mean value: $48 \pm 6.4 \mu$ V). The individual event related potential was calculated for each electrode site in each of the three experimental conditions. In the case where an individual electrode contained substantial noise compared to the average signal for the participant, only that individual electrode was removed, resulting in a small number of exclusions. Only three participants (one in each experimental condition) were excluded from the study on the basis of poor EEG signal.

Results

Behavioral Data

Accuracy rates for all three experimental conditions were extremely high, on average participants were correct 97% percent of the time. Mixed ANOVA analysis showed no difference in accuracy across conditions, but revealed a significant effect of match/mismatch ($F(1,59) = 133.33$; $p < .01$, $\eta^2 = 0.72$), with participants making more errors in the match (5%) trials in comparison to mismatch trials (1%). Regarding

RTs, we found a main effect of experimental condition ($F(2,59) = 478.86$; $p < .01$, $\eta^2 = 0.95$). Post hoc tests revealed that only the TO condition differed significantly from both WP and PW condition (See Figure 3.).

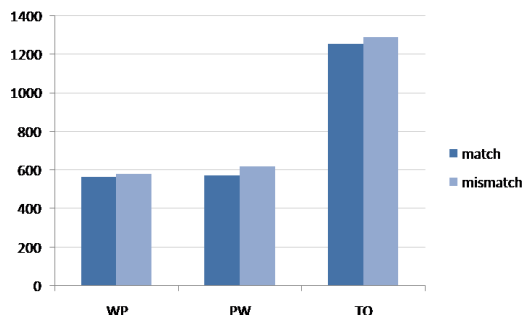


Figure 3. RTs for match and mismatch trials across three experimental conditions (WP-word to picture; PW-picture to word; TO-together condition)

ERP Data

The model used for analysis consisted of 9 electrode sites: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, divided into three bands of coronal orientation (frontal-central-parietal), and three lateral regions (left-central-right). All analysis, including the determination of time windows of interest were done with difference waves. Namely, we subtracted ERP wave forms of match trials from the ERP wave forms of mismatch trials in order to isolate the N400 component more accurately. For determining time windows of interest we adopted an exploratory approach to data analysis. Following the analysis of Kovic et al. (2010), mean amplitude measurements were extracted from the continuous EEG signal into 20 ms bins for each participant across all experimental conditions. Successive ANOVAs were conducted on each time bin. Windows of interest were defined if at least 3 consecutive 20 ms bins were significant ($p < .05$). After identification of windows, mean amplitudes across the window were computed for each experimental condition, and further analysis conducted. Two windows of interest (260ms - 440 ms; 440ms - 680ms) were analysed with a 3x3x2 repeated measures ANOVA with within-subjects factors of Frontality (Frontal, Central, Parietal) and Laterality (Left, Midline, Right), and between-subjects factor Time-condition (WP condition, PW condition). Given that the latency of the component of interest (namely N400) and pattern of responding was completely different between TO in comparison to WP and PW conditions, we decided to exclude the TO condition from the amplitude analysis.

Time window 260-440ms We found that WP and PW condition differed in the N400 amplitude, given that there was a significant effect of Time-condition in the first time window ($F(1,38) = 6.01$, $p < .01$, $\eta^2 = 0.14$). PW condition elicited a more negative response than the WP condition (See Figure 4.) . Apart from that, we also

found a main effect of Laterality $F(2,30) = 79.24$; $p < .01$, $\eta^2 = 0.58$ and an interaction Laterality x Frontality $F(4,30) = 7.86$; $p < .01$, $\eta^2 = 0.18$). The same effect reported here can be easily recognised in Figure 5. whereby the dark blue colour indicates the more prominent N400 effect across scalp distribution.

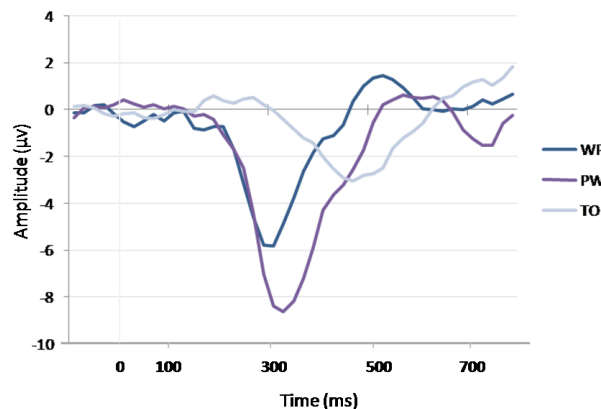


Figure 4. Difference waves showing the N400 and P600 effects across experimental conditions on the Cz electrode (WP-word to picture; PW-picture to word; TO-together condition)

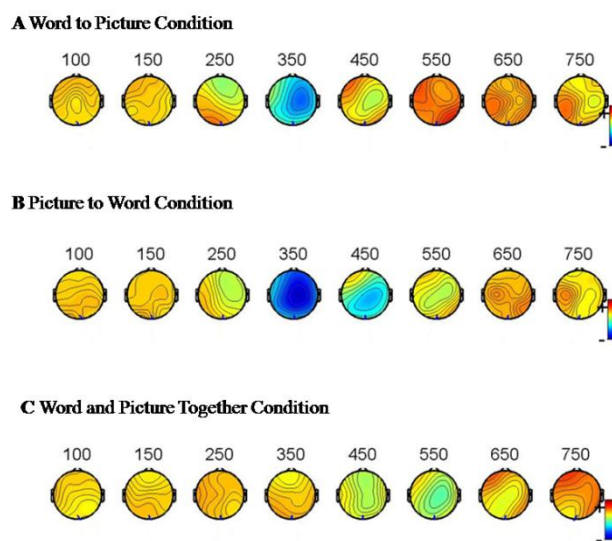


Figure 5. Heat maps showing time course of the distribution of the N400 and P600 effects across the scalp in all three experimental conditions. The dark blue color indicates the more negative amplitudes

Time window 440-680ms A repeated measures ANOVA was conducted in order to analyze the difference of the P600 amplitude between the WP condition and PW condition. Analysis revealed a main effect of Time-condition $F(1,38) = 4.99$; $p < .01$, $\eta^2 = 0.12$) with WP condition eliciting a more positive response (See Figure 4.). Similarly to the first, earlier time window, we found a main effect of Laterality

($F(2,30) = 4.34$; $p < .05$, $\eta^2 = 0.11$) and a Laterality x Frontality interaction ($F(4,30) = 8.95$; $p < .01$, $\eta^2 = 0.19$). Figure 5. shows the distribution of activity through time and across scalp. Orange colour indicates the more positive responses in amplitude.

Discussion

In this study we tested if and how the order of stimuli presentation impacts semantic processing. In particular, we tested the hypothesis that the mapping from picture to name would be easier for processing in comparison to name-to-picture mapping given that there are many instances of pictures and thus mapping a single name to multiple potential objects seemed as a harder experimental condition in comparison to mapping from a particular picture to the name (which we pretested to select the most adequate for the given object).

The results we obtained demonstrate that the hardest condition for the semantic processing is the one in which the labels and pictures were presented at the same time. This finding was in accordance with our expectations because participants needed to process both information (word and picture) in parallel, which opens possibilities of interference as well as competitive processes making the task harder. Additionally, in the together condition, there was no priming in a strict sense as in the other two conditions, which prevented participants from forming any expectations which would help them with the task. It is noteworthy to say that in this study latency of ERP response in TO condition corresponds to a time window commonly associated with the P600 component. However, unlike the P600 component, here its polarity is negative. This is why we believe it is in fact a late N400 effect, delayed because of the difficulty of the task, which was also represented through longer RTs. All of this could account for the different morphology of the N400 in the TO condition.

Regarding the other two conditions, we observed a larger N400 amplitudes in PW condition in comparison to WP condition. Thus, in accordance with our expectations we found that picture to label mapping (PW) was easier for participants to process given that the N400 amplitude was more prominent in this condition in comparison to label to picture (WP) condition.

A more parsimonious way of interpreting this data would be in terms of violation of expectation reflected through the amplitude of the N400 component. This is consistent with a hypothesis that one can predict a word from a picture with more precision than the opposite (which leads to a larger violation of expectations).

The observed pattern of results in picture to word mapping would potentially be different in the case of less typical or atypical pictures. Violation of expectation in that case would certainly be higher than observed in the current study. Similarly, when mapping from word to picture, we would also expect greater violation of expectation than the one reported in this study.

Another interesting stream of research would be to contrast WP and PW mapping in the situation of novel object formation, that is – during the process of category formation. Here, we would have better control of the variability of the objects used in the study, given that with familiar objects the variability is much higher for the pictures (then for words).

Another component that turned out to be sensitive to semantic processing in this study, was the P600 component which in relevant literature is commonly related to syntactic processing (Kotz, Frisch, von Cramon & Friederici, 2003; Osterhout & Holcomb, 1992). However, there are a few studies which also reported P600 to be sensitive to semantic processing and interpreted as additional processing of meaning (Frisch, Schlesewsky, Saddy & Alpermann, 2002; Martín-Loeches, Nigbur, Casado, Hohlfeld & Sommer, 2006).

In our study, WP condition elicited a more positive P600 response in comparison to PW condition, which would suggest the information in this condition required additional processing in the later stages. However, given that there is an ongoing debate over the meaning of P600 in semantic processing, and since this component wasn't of main interest in this study, we would reserve from making firm claims when interpreting these results.

Practical implications of this research would be that in a classical priming experiments the best way to design the experiment would be to consistently map from pictures to words (that is, from more specific to more general representations), at least in the situation when typicality, familiarity and frequency of the selected pictures are high.

Acknowledgments

We would like to express our gratitude to prof. Suzy Styles from Nanyang Technological University in Singapore for the valuable comments she made on the study, as well as to anonymous reviewers whose suggestions were incorporated in the final version of this paper.

This research was partially supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, Belgrade, Project No. 179033 and Project No. 179006.

References

- Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, 32(2), 177-190.
- Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *The Journal of Neuroscience*, 35(25), 9329-9335.
- Frisch, S., Schlesewsky, M., Saddy, D. & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, 85, B83-B92.
- Ganis, G., Kutas, M., & Sereno, M. E. (1996). The search for "common sense": An electrophysiological

- study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience* 8(2), 89-106.
- Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, 8(4), 379-411.
- Kotz, S. A., Frisch, S., von Cramon D.Y. & Friederici, A. D. (2003). Syntactic language processing: ERP lesion data on the role of the basal ganglia. *Journal of the International Neuropsychological Society*, 9, 1053–1060.
- Ković, V., Plunkett, K., & Westermann, G. (2009). Shared and/or separate representations of animate/inanimate categories: An ERP study. *Psihologija*, 42(1), 5-26.
- Ković, V., Plunkett, K., & Westermann, G. (2010). A unitary account of conceptual representations of animate/inanimate categories. *Psihologija*, 43(2), 155-165.
- Kutas M, & Federmeier K. D. (2009). N400. *Scholarpedia*. 4(10):7790.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology*, 62, 621.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. MIT press.
- Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun-adjective agreement in Spanish. *Brain Research*, 1093, 178-189.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785-806.
- Savić, A., Malešević, N., & Popović, M. (2013). Motor imagery driven BCI with cuebased selection of FES induced grasps. In J. L. Pons, D. Torricelli, & M. Pajaro (Eds.). *Converging clinical and engineering research on neurorehabilitation* (Vol. 1, pp. 513–516). Berlin, Heidelberg: Springer.