NeuroReport 10, 829–835 (1999)

Tracking the time course of object categorization using event-related potentials

James Tanaka,[CA] Phan Luu,[1] Matthias Weisbrod[2] and Markus Kiefer[3]

Department of Psychology, Oberlin College, Oberlin, OH 44074; [1]Department of Psychology, University of Oregon, OR 44074, USA; [2]Department of Psychiatry, University of Heidelberg D69115 Heidelberg; [3]Department of Psychology, University of Ulm, D8970 Ulm Germany

[CA]Corresponding Author

OBJECT categorization processes were investigated by measuring event-related potentials while subjects categorized objects at the superordinate (e.g. animal), basic (e.g. dog) and subordinate (e.g. beagle) levels of abstraction. An enhanced negative deflection (N1) was found at posterior recording sites for subordinate level categorizations compared with basic level categorizations and was interpreted as a marker of increased visual analysis. In contrast, superordinate level categorizations produced a larger frontal negativity relative to basic level categorizations and were interpreted as an indicator of increased semantic processing. These results suggest a neurophysiological basis for the separate cognitive processes responsible for subordinate and superordinate object categorizations. *NeuroReport* 10:829–835 © 1999 Lippincott Williams & Wilkins.

Key words: Event-related potentials; Object recognition; Semantics; Visual perception

Introduction

In object naming, the same object can be referred to by many names. For example, the object dog can be called an animal at a general (superordinate) level of abstraction, a dog at an intermediate (basic) level, and a golden retriever at a specific (subordinate) level. While any object can be categorized at multiple levels of abstraction, it has been argued that one level, referred to as the basic level, occupies a privileged position in the object hierarchy. The basic level is defined as the most inclusive level at which category members share a common shape. Members of the same basic level category (e.g. dog) tend to share the same structural features and, therefore, bear a physical resemblance to one another. The special status of basic level categories in object recognition has been established in naming and category verification tasks where people are fastest to categorize objects at the basic level (e.g. apple, chair, dog) and slower to categorize objects at superordinate (e.g. fruit, furniture, animal) and subordinate (e.g. Macintosh apple, kitchen chair, beagle) levels of abstraction [1–3]. Based on these data, it has been suggested that the basic level is the point at which the perceived object stimulus first makes contact with an object representation stored in memory [1].

According to current object categorization theory, subsequent activation of categories superordinate and subordinate to the basic level rely on distinct cognitive processes. Relative to basic level categorizations, it has been claimed that superordinate level categorizations require abstract semantic processing. Some support for this view comes from feature listing studies in which a higher proportion of functional and behavioral properties of the object (e.g. eat, breathe, reproduce) are mentioned for superordinate level concepts (e.g. animal) [3–5]. Other studies have shown that the times required for superordinate level categorization of basic level words are highly correlated with superordinate level categorization of basic level pictures [1]. This latter result indicates that categorizations at the superordinate level are not directly bound to the format of the stimulus (i.e. word or picture), but are based on the abstract semantic association between basic and superordinate level concepts.

Subordinate level categorization, on the other hand, has been linked to perceptual processes. For example, in a feature listing task, a large number of perceptual features (e.g. small size, brown for sparrow) and modified part features (e.g. long beak for woodpecker) are mentioned for subordinate level objects, suggesting that this level of categorization requires a finer grain of perceptual analysis. Accordingly, it has been shown that in an object recognition task, short exposures of the picture stimulus (i.e. 75 ms) interfered with subordinate level categorizations but not basic or superordinate decisions [1,6]. These experiments suggest that subordinate categorizations demand an increased amount of visual analysis.

If different cognitive operations are recruited to categorize objects at different levels of abstraction then the brain areas associated with those operations should be differentially engaged in an object categorization task. Consistent with this view, prefrontal brain areas are more active when subjects categorized objects at the superordinate level than when the same objects were categorized at the basic level [7]. Presumably, prefrontal activation reflects the accessing of semantic information necessary for superordinate level judgments. In contrast, other neuroimaging studies have shown that subordinate level categorization activated fusiform and inferior temporal gyri (FIT) regions as well as the temporal poles [8], areas associated with visual object recognition. Thus, the above studies suggest that superordinate and subordinate level categorizations are dissociable with respect to their neural substrates.

Object categorization processes have not been directly examined using ERP methods. However, related electrophysioloigcal studies have indicated that semantic and visual processes can be dissociated with respect to their time course and topographic locations. As an indicator of semantic processes, it has been reported that left frontal areas are differentially activated as early as 200 ms after stimulus onset in tasks involving noun–verb associations [9]. As an index of visual analysis, an enhanced early positive component and negative component are elicited when subjects must selectively attend to a one-stimulus dimension (e.g. color) while ignoring other stimulus dimensions (e.g. shape) [10,11]. Although the exact source of the visual attentional components cannot be precisely determined from scalp recordings, both the early positive and negative components appear to originate in posterior, visual brain areas (e.g. occipital, occipital–temporal areas) [10].

In the current experiment, the temporal dynamics involved in object categorization are directly evaluated. Following the standard category verification paradigm, subjects categorized objects at superordinate (e.g. animal), basic (e.g. bird) and subordinate (e.g. robin) levels of categorization while wearing a high density electrode net that monitored electrophysiological activity at 64 scalp sites. If visual and semantic processes are localizable to different neural substrates, brain areas associated with visual processing (i.e. posterior areas) should be differentially activated when subjects categorize objects at the subordinate level of abstraction. In contrast, brain regions associated with semantic processing (i.e. frontal areas) should be differentially activated when subjects categorize objects at the superordinate level.

Materials and Methods

Subjects: Twenty-eight right-handed undergraduate students from the University of Oregon participated in this study. All subjects were native English speakers and had normal or corrected-to-normal vision. Subjects were paid for their participation.

Materials: Stimuli for the experiment consisted of 40 color pictures from four natural (animal, plant, fruit, vegetable) and six artifactual (tool, furniture, musical instrument, sports equipment, vehicle, weapon) categories. These objects were judged as typical exemplars of their category according to the Battig and Montague [12] normative ratings. Pictures were taken primarily from McMillan's Visual Dictionary [13] and digitized with a MicroTek Z Scanner. Each image was scaled to fit within a 100×100 pixel array. Left-facing and right-facing versions of each picture were generated and presented randomly in the experiment.

Procedure: Subjects were tested in an electrically-shielded booth. After subjects read a list of the 40 objects included in the study, they were seated in front of the computer monitor with their head stabilized in a chin rest. From the subjects' viewing distance of 60 cm, the picture stimuli subtended a visual angle of $\sim 0.5°$ in the horizontal and vertical dimensions. At the beginning of each trial, a fixation point (a plus sign symbol) appeared on the computer monitor for a random interval that varied between 1000 and 1500 ms. The fixation point was replaced by a category name (superordinate, basic or subordinate) for 255 ms and was then replaced by the fixation point for 570 ms. The picture stimulus was then presented for 255 ms and was replaced by the fixation point for 735 ms. At the end of the trial, the 'true/false' screen appeared prompting subjects for their response. If the picture matched the category word, subjects were instructed to press the 'true' key; otherwise they were to press the 'false' key. In the false condition, foil category labels were drawn from a contrast category that was at the same level of abstraction as the target. For example, the category labels furniture, saw, and claw hammer preceded the picture of the ball peen hammer in the superordinate, basic, and subordinate false trials, respectively. EEG was recorded for 2 s during each trial, beginning 195 ms prior to the onset of the category word and was terminated prior to subject response at the onset of the true/false screen. The inter-trial interval was randomly varied from 2 to 2.5 s. Each of the 40 picture stimuli were tested in three category conditions (superordinate, basic, subordinate) and two response conditions (true, false) yielding a total of 240 randomly presented trials.

EEG recording: EEG was recorded with the Geodesic Sensor Net [14] with an array of 57 Ag/AgCl sponge sensors placed on the scalp surface with an additional seven Ag/AgCl surface sensors placed at each infra-orbital, external canthus, and mastoid sites and at the glabella. The impedance of all electrodes was below 40 kΩ (10–40; k = 87). All recordings were referenced to the right mastoid. EEG recordings from the 64 sites were amplified and filtered with 0.01–50 Hz (3 dB attenuation) bandpass and a 60 Hz notch filter. The signals were sampled at 125 samples/s and were digitized with a 16 bit analog-to-digital converter (National Instruments NB-MIO 16). Editing of the EEG for movement, eye movement and blink artifacts was performed by a computerized artifact detection algorithm. Trials containing blinks (30 μV difference between vertical EOG channels) were omitted from analysis. Channels were also eliminated if they changed by > 50 μV between samples or exceeded ± 100 μV.

The artifact-free EEG was then averaged for each subject across all correct trials. The activity of all recorded sites was average referenced in which the activity of a given site was expressed as the difference between the site and the average of all other recording sites [15,16]. Although the underside of the brain is poorly represented by this method, analyses with the 64-channel Geodesic Sensor Net show that the average reference closely approximates the distance-weighted average reference [16]. The average reference transformation provides a reference independent estimation of scalp voltage. The averaged epochs were digitally filtered with a 30 Hz lowpass to remove any artifacts due to the 67 Hz refresh rate of the monitor. EEG was baseline-corrected for the 195 ms interval prior to the presentation of the category word and 50 ms interval prior to the picture stimulus. A final grand average was obtained by averaging across the subject averages for each experimental condition.

Results

During the word epoch (0–1020 ms), an enhanced positivity was produced by subordinate level words at the central recording sites relative to the basic and superordinate level words. However, at the end of the word epoch and prior to the picure epoch (1020–2000 ms), EEG activity returned to baseline levels and no differences were found between the superordinate, basic and subordinate level category conditions ($P < 0.10$). Because the focus on this study was on processing during the picture, no further analyses were performed on word epoch ERPs.

For analysis of the ERPs to the picture stimulus, waveforms were baseline adjusted 100 ms prior to the onset of the picture. Figure 1 shows the averaged ERP waveform for superordinate, basic and subordinate conditions collapsed across true and false responses during the picture epoch. Beginning ~100 ms after the onset of the picture stimulus, a positive deflection (P1) was demonstrated across all three category conditions. The P1 was followed by a negative deflection in the posterior channels (N1) ~140 ms after onset of the picture stimulus. While the N1 was detected in all three category conditions, subordinate level categorizations generated an enhanced N1 amplitude in left posterior channels relative to basic and superordinate level, as shown in Fig. 1. As the topographic voltage map of Fig. 2a shows, the difference between subordinate and basic level activation peaked about 160 ms after the picture was presented.

An analysis of variance performed on the latencies of the N1 component was performed for superordinate, basic and subordinate level categorizations of the posterior channels (channels 16, 20, 22, 23, 25, 26, 30, 31, 34) [17]. The ANOVA showed no differences across the three category conditions with respect to their latencies ($P > 0.10$). Using a latency window of 139–179 ms, the mean amplitudes of the N1 component were submitted to a three-way ANOVA with the factors of category level (superordinate, basic, subordinate), response (true, false) and laterality (left, right) as within-subject variables. The significant effect of category level ($F(2,26) = 15.30$, MSe = 51,080,131, $P < 0.001$) demonstrated the differences in ERP produced by superordinate, basic and subordinate level categorizations. Critically, direct comparisons revealed that the subordinate categorizations were significantly more negative than basic and superordinate categorizations ($P < 0.001$). Laterality interacted with category ($F(2,52) = 4.40$, MSe = 8,421,253, $P < 0.02$), such that subordinate categorizations produced a greater negative deflection in the left hemisphere than in the right hemisphere. No other main effects or interactions were significant.

A brain electrical source analysis (BESA) [18] was applied to further explore possible sources of the ERP effect attributable to subordinate categorization. Dipoles were fitted on the difference ERPs between the subordinate and basic level conditions within the same time window of 139–179 ms post-

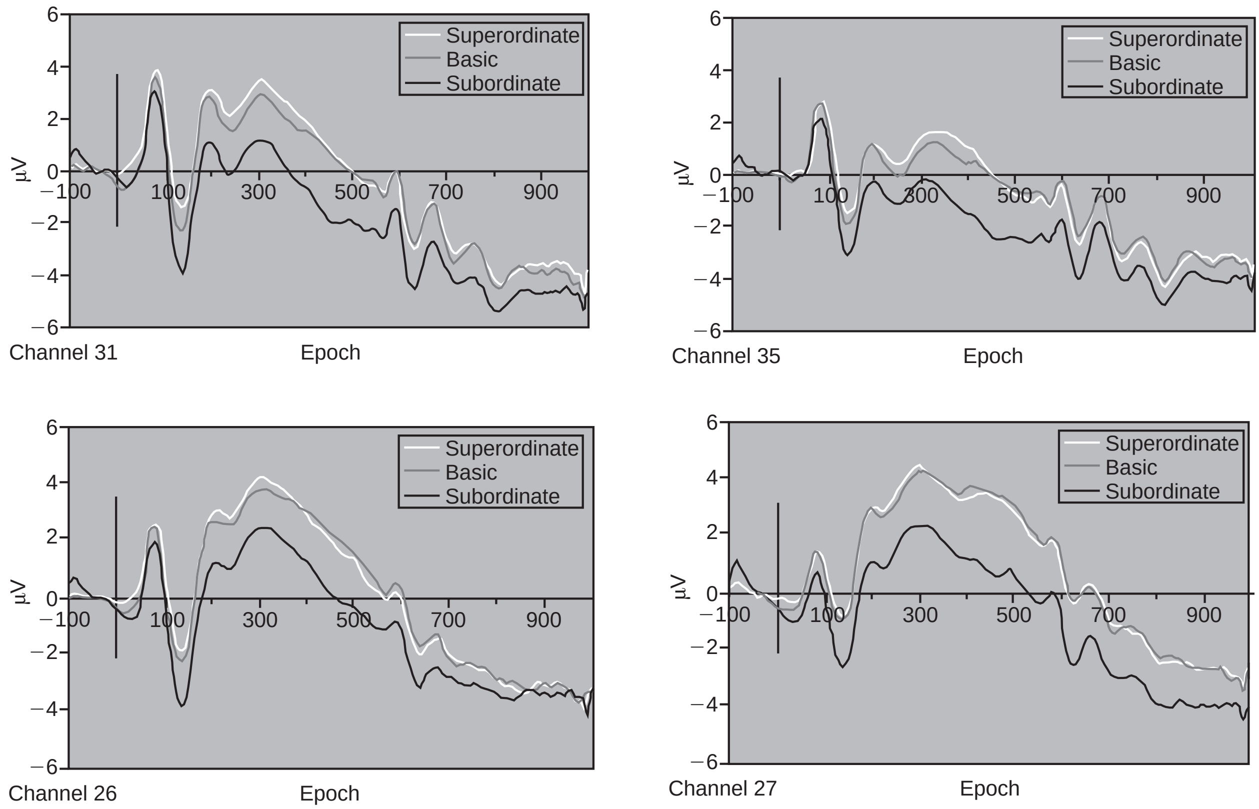

FIG. 1. Selected channels of average-referenced, grand-averaged ERP waveforms during the picture EPOCH for superordinate, basic and subordinate level category conditions. Waveforms shown are the left posterior sites, P3P (31), PzL (35), P3A (26), P3S (27). Arrow indicates the enhanced N1 negativity for subordinate level categorizations.

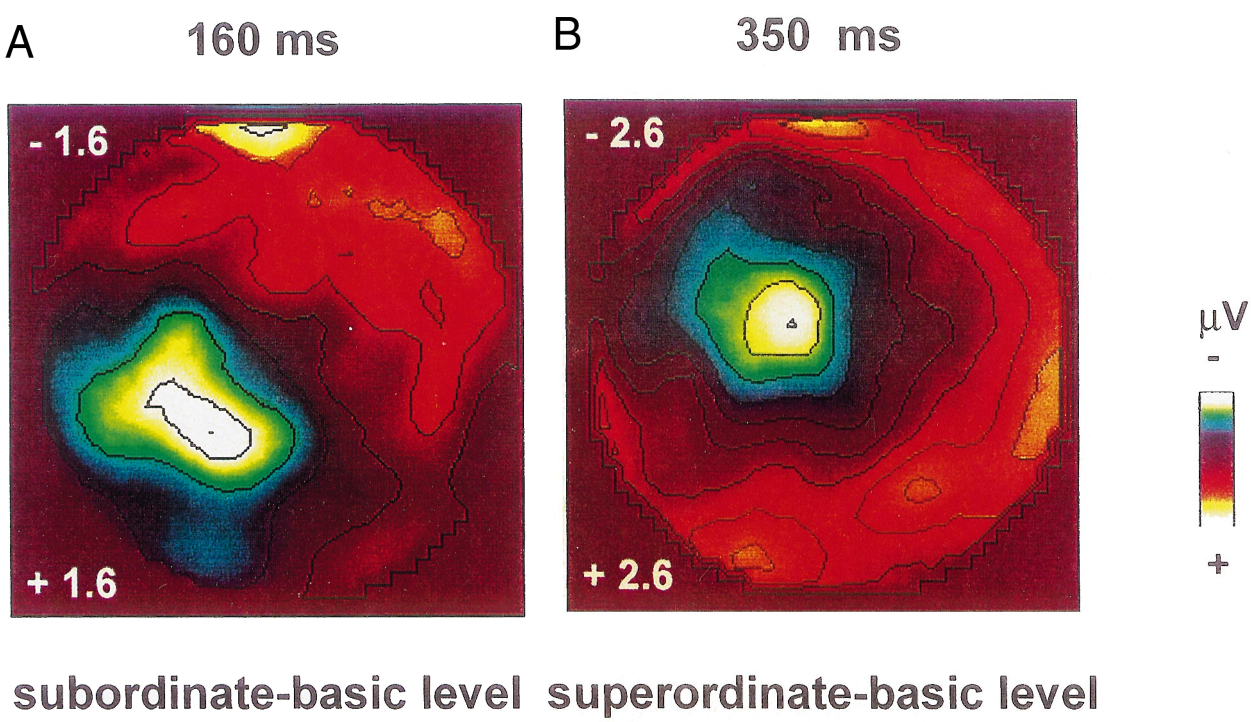

FIG. 2. Topographic maps of voltage differences between subordinate level and basic level activation and superordinate level and basic level activation. (**a**) In the subordinate level category condition, negative deflection occurred 164 ms after the onset of the picture stimulus in the posterior areas. (**b**) In the superordinate level category condition, negative activity was detected 324 ms after the onset of the picture stimulus in the frontal areas of the brain.

stimulus onset used in the statistical analysis above. As shown in Fig. 4, a solution with two dipoles in the left posterior hemisphere was obtained which accounted for about 87% of the variance. The first dipole was tangentially oriented in the left anterior ventro medial, inferior part of the temporal lobe. This dipole explained about 74% of the variance. However, a second dipole was necessary to obtain a proper fit between the dipole model and the data. BESA located this dipole more posteriorly in the superior part of the left temporo-parietal area lobe with an orientation that was radial and slightly medial. Additional dipoles did not result in a substantially better fit of the model.

Superordinate level categorizations diverged from basic level categorizations ~300 ms after the picture was presented. At this point, a negative deflection was detected in the frontal channels for superordinate level categorizations relative to basic level and subordinate level categorizations (as shown in Fig. 4). The topographic voltage map in Fig. 2b shows the difference in activation between superordinate and basic level categorizations at 324 ms post-stimulus onset.

To test for differences in frontal activation, mean amplitudes were averaged from frontal sites (channels 4, 8, 11, 14, 52, 56, 59, 55) within a latency

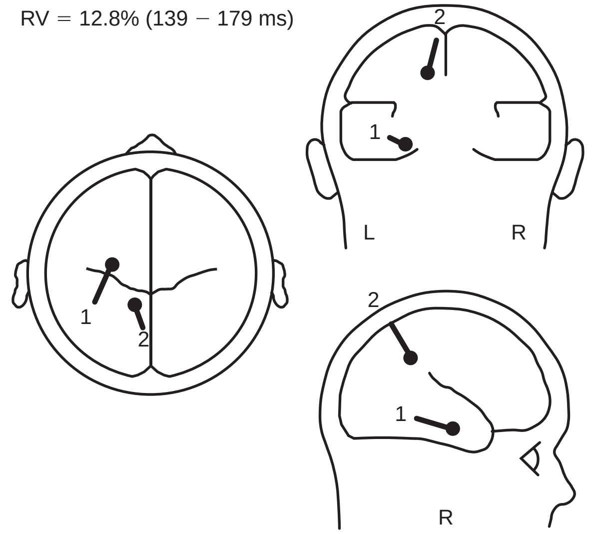

FIG. 3. A brain electrical source analysis (BESA) was performed in which the difference wave between subordinate level and basic level categorizations using a latency window of 139–179 ms post-picture stimulus onset. Two dipoles, one located in the left anterior ventro-medial part of the temporal lobe and other located in the left temporo parietal area, accounted for 87% of the difference variance.

window of 306–356 ms. Category level was significant ($F(2,52) = 38.27$, $MSe = 4.11$, $P < 0.001$), demonstrating that the waveform produced by superordinate, basic and subordinate categorizations

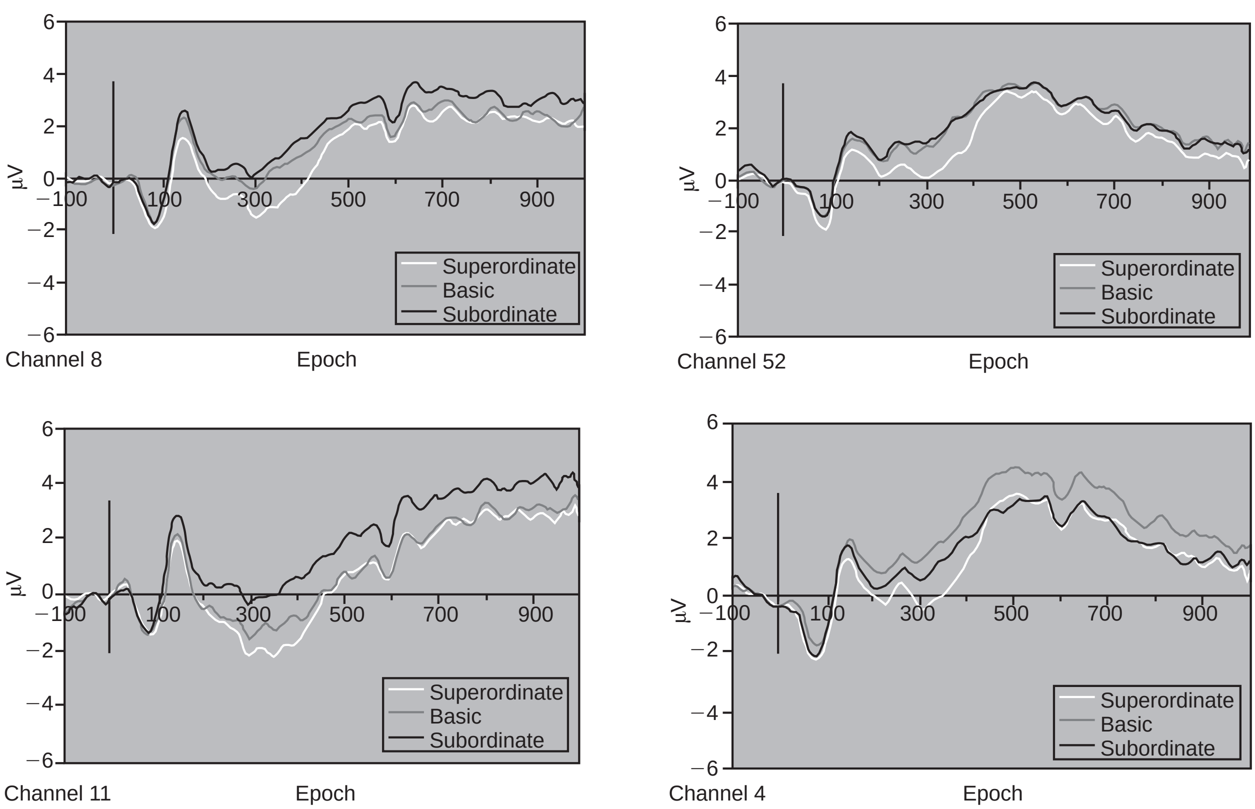

FIG. 4. Selected channels of average-referenced, grand-averaged ERP waveforms during the picture epoch for superordinate, basic and subordinate level category conditions. Waveforms shown are the left and right frontal sites, F3S(8), F4SS(52), F3I(11) and F3SS(4). Arrow indicates the enhanced negativity for superordinate level categorizations.

reliably diverged. Direct comparisons showed that superordinate categorizations were significantly more negative than basic and subordinate categorizations ($P < 0.001$). Response ($F(1,26) = 11.05$, $MSe = 1.935E7$) and laterality ($F(1.26) = 11.44$, $MSe = 5.383E7$) were significant ($P < 0.01$). Effects of response and laterality also interacted ($F(1,26) = 12.85$, $MSe = 3841817$, $P < 0.01$) such that false responses produced a greater negativity in the left hemisphere than the right hemisphere whereas true responses produced a greater positivity in the right hemisphere than the left hemisphere. No other interactions reached significant levels.

BESA was used again to further assess the ERP effect of superodinate categorization. Source analysis was applied on the difference ERPs between the superordinate and basic level conditions within the same time window as used in the parametric analysis. As shown in Fig. 5, BESA placed a dipole in the left inferior frontal lobe which accounted for about 87% of the variance. It was radially oriented and pointed slightly to the left. Additional dipoles did not result in a substantially better fit of the model.

Discussion

With respect to the time course of object categorization processes, two important results are revealed in the present ERP study. First, differences related to basic level and subordinate level categorizations appear relatively early on in processing. Specifically, subordinate level categorizations produced an enhanced negative deflection ~140 ms after the onset of the picture stimulus. The negative wave form was consistent in its approximate latency and amplitude with N1 effects found in other studies involving visual attention [10,11]. The early onset of the negative component renders it unlikely that enhanced visual processing required for subordinate categorization occurs after basic level access. Rather, these results indicate that ~140 ms after a picture is presented the visual processes required for subordinate level categorization become differentiated from the visual encoding processes necessary for basic level categorization. While the N1 is generally interpreted as indicative of enhanced visual processing conferred upon the attended object, it might also reflect non-specific arousal factors such as task difficulty or differential preparation [10,19]. However, we do not think the N1 reported in this study is a result of non-specific factors for the following reasons. First, if the magnitude of the N1 was a reflection of task difficulty, the more difficult superordinate level categorization should produce a greater N1 relative to the easier basic level judgment. However, as shown in Fig. 1, the amplitude of the superordinate level N1 was less than the amplitude of the basic level N1. Thus, the magnitudes of the N1 component for the superordinate, basic and subordinate level judgments did not mirror their relative difficulty. Second, no differences were found between superordinate, basic and subordinate level conditions in the earlier P1 component, indicating that general arousal levels were equated at the beginning of the picture onset. These findings suggest that the enhanced N1 component for subordinate level categorizations is related to the increased visual demands required for the more specific level of categorization rather than to general arousal factors.

The other important finding of this study was that wave forms associated with superordinate level categorizations did not diverge from basic level categorizations until about 340 ms after the presentation of the picture stimulus. The relatively late difference between superordinate and basic level categorization is consistent with the idea that objects are initially categorized at the basic level before abstract superordinate level information is retrieved [1]. Hence, the current findings complement previous object categorization studies [7,8] by demonstrating temporal differences between subordinate and superordinate level categorizations. Whereas subordinate level categorization differs from basic level categorization at an early stage of processing, superordinate level differences occur at a later stage of processing.

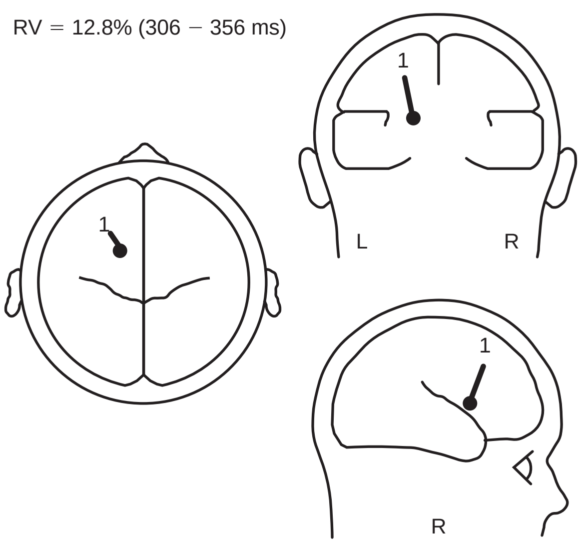

FIG. 5. BESA was performed comparing the difference waves between superordinate level and basic level categorizations using a latency window of 306–356 ms post-stimulus onset. A single dipole located in the left anterior ventro medial part of the temporal lobe and other located in the left inferior frontal lobe accounted for 87% of the difference variance.

Conclusion

The present study demonstrates that subordinate level and superordinate levels of categorization are temporally and spatially separable. About 140 ms after a picture is presented the left posterior regions of the brain are differentially engaged when an object is categorized at the subordinate level of abstraction. In contrast, about 350 ms after picture onset, frontal brain areas are activated when an object is categorized at the superordinate level. Because the picture stimuli were held constant in the superordinate and subordinate level category conditions, ERP differences must stem from differences in cognitive processes needed to categorize pictures at different levels of abstraction rather than differences in the stimuli themselves. The current results indicate that the cognitive separability of object categorization processes may be maintained and facilitated by separable neural systems.

References

1. Jolicoeur P, Gluck MA and Kosslyn SM. *Cogn Psychol* **16**, 243–275 (1984).
2. Murphy G and Brownell H. *J Exp Psychol Learn Mem Cogn* **11**, 70–84 (1985).
3. Rosch E, Mervis C, Gray W *et al. Cogn Psychol* **8**, 382–439 (1976).
4. Tanaka JW and Taylor M. *Cogn Psychol* **23**, 457–482 (1991).
5. Tversky B and Hemenway K. *J Exp Psychol Gen* **113**, 169–193 (1984).
6. Kiefer M. Poodle, dog or animal? Specificity of object names as a result of flexible concept-word activation. In: R. Meyer-Klabunde C. von Stutterheim (Hrsg.), Conceptual and semantic knowledge in language production. Proceedings of a Workshop of the Special Collaborative Research Program 245 "Language and Situation" (Vol. 2, pp. 186–204), 1996: Heidelberg: University of Heidelberg..
7. Kosslyn S, Alpert N and Thompson W. *Hum Brain Mapp* **3**, 107–132 (1995).
8. Gauthier I, Anderson A, Tarr M *et al. Curr Biol* **7**, 645–651 (1997).
9. Snyder A, Abdullaev Y, Posner M and Raichle M. *Proc Natl Acad Sci USA* **92**, 1689–1693 (1995).
10. Mangun G and Hillyard S. Mechanisms and models of selective attention. In: Coles M and Mangun G, eds. *Electrophysiology of the Mind*. Oxford: Oxford University Press, 1996: 40–85.
11. Van Voorhis S and Hillyard S. *Percept Psychophys* **22**, 54–62 (1977).
12. Battig W and Montague W. *J Exp Psychol Monogr* **80**, 1–46 (1969).
13. MacMillan Visual Dictionary. New York: MacMillan, 1992.
14. Tucker D. *Electroencephalogr Clin Neurophysiol* **87**, 154–163 (1993).
15. Bertrand O, Perrin F and Pernier J. *Electroencephalogr Clin Neurophysiol* **62**, 462–464 (1985).
16. Tucker D, Liotti M, Potts G *et al. Hum Brain Mapp* **1**, 134–152 (1994).
17. Curran T, Tucker D, Kutas M and Posner M. *Electroencephalogr Clin Neurophysiol* **88**, 188–209 (1993).
18. Scherg M. Fundamentals of dipole source potential analysis. In: F Grandori, M Hoke, and G Romani, eds. *Auditory Evoked Magnetic Fields and Electric Potentials. Advances in Audiology*. Basel: Karger, 1990: 40–69.
19. Hillyard S and Picton T. Event-related potentials and selective information processing in man. In: Desmedt J, ed. *Cogn Components in Cerebral Event-Related Potentials and Selective Attention. Progress in Clinical Neurophysiology Vol. 6*. Basel: Karger, 1979: 1–52.

ACKNOWLEDGEMENTS: We are grateful to Michael Posner and the Institute of Cognitive and Decision Sciences at the University of Oregon for their support and Tim Curran for his technical assistance. This research was supported by NIH grant R15 HD30433 and a Keck Foundation Faculty Research Grant award to J.T.

Received 22 December 1998; accepted 19 January 1999