*Chinese Science Bulletin* 2004 Vol. 49 No. 8 819—823

# Neural mechanisms of perceptual grouping in human visual cortex

MAO Lihua[1], HAN Shihui[1], GUO Chunyan[2] & JIANG Yi[1]

1. Department of Psychology, Center for Brain and Cognitive Sciences, Peking University, Beijing 100871, China;
2. Learning & Cognition Lab, Capital Normal University, Beijing 100037, China

Correspondence should be addressed to Han Shihui (e-mail: shan@pku.edu.cn)

Abstract **The current work examined neural substrates of perceptual grouping in human visual cortex using event-related potential (ERP) recording. Stimulus arrays consisted of local elements that were either evenly spaced (uniform stimuli) or grouped into columns or rows by proximity or color similarity (grouping stimuli). High-density ERPs were recorded while subjects identified orientations of perceptual groups in stimulus arrays that were presented randomly in one of the four quadrants of the visual field. Both uniform and grouping stimulus arrays elicited an early ERP component (C1), which peaked at about 70 ms after stimulus onset and changed its polarity as a function of stimulated elevations. Dipole modeling based on realistic-head boundary-element models revealed generators of the C1 component in the calcarine cortex. The C1 was modulated by perceptual grouping of local elements based on proximity, and this grouping effect was stronger in the upper than in the lower visual field. The findings provide ERP evidence for the engagement of human primary visual cortex in the early stage of perceptual grouping.**

**Keywords: calcarine cortex, event-related potential, perceptual grouping, proximity.**

**DOI: 10.1360/04wc0004**

Perceptual grouping serves to form perceptual units or objects at an early stage of visual processing for subsequent higher-level attentional processing[1,2]. Animal studies have shown that the primary visual cortex (V1) may be the earliest neural substrate underlying the grouping processes[3]. However, as stimuli used in animal studies are much smaller than those used in human behavioral studies[4,5], it is unclear to what extent the results from animal studies can be used to deduce the neural mechanisms underlying large-area perceptual grouping in humans. Recent event-related potential (ERP) studies presented subjects with stimulus arrays in which local elements were either evenly spaced (uniform stimuli) or grouped into columns or rows based on proximity or similarity[6,7]. It was found that, relative to uniform stimuli, proximity-grouping stimuli elicit a positive wave at 100—120 ms after stimulus onset over the medial occipital area, suggesting that the medial occipital cortex is engaged in grouping operations. Recent neuroimaging studies found that neural activities are stronger in the early visual areas such as V1 and V2 to global contours consisting of collinear elements than to the patterns of randomly oriented local elements[8]. However, the results could not distinguish whether the V1 activation arises from the initial grouping process when visual information first arrives at V1 or the feedback from higher visual areas because of the limitation of temporal resolution.

The current work investigated this problem by testing if an early ERP component (C1) is modulated by perceptual grouping. The C1 recorded at occipito-parietal electrodes is negative for positions above the horizontal meridian (HM) but positive for stimuli located below HM and is localized to calcarine cortex[9,10]. The C1 peaks at 50—90 ms poststimulus and reflects the initial activities of the primary visual cortex evoked by visual stimuli. Here we presented the stimulus arrays used in our previous work[6,7] randomly in one of the four quadrants of the visual field while high-density ERPs were recorded to uniform and grouping stimuli. We examined whether the C1 component was modulated by grouping of local elements in stimulus arrays to test the hypothesis that human primary visual cortex underpins the initial process of perceptual grouping defined by Gestalt laws.

## 1 Materials and methods

(ⅰ) Subjects, stimuli and procedure. Sixteen adults aged 19—29 years (mean 24, 8 male, 8 female) participated in this experiment as paid volunteers. All participants were right-handed, had normal or corrected-to-normal vision, reported no color blindness, and gave informed consent. Each stimulus consisted of a square lattice of elements (red or green disks) in an 8×8 array (see Fig. 1). The uniform stimulus consisted of alternate red and green disks distributed evenly across the lattice. The proximity-grouping stimuli consisted of alternate red and green disks arranged to form rows or columns by adjusting the distances between two adjacent rows or columns. The similarity-grouping stimuli were made by moving the red and green disks to form rows or columns with elements of identical color. The size of local elements and stimulus arrays were the same as those in our previous work[7]. A white fixation cross was continuously visible in the centre of a black background of 0.14 cd/m$^2$. The stimulus patterns had a luminance of 0.48 cd/m$^2$. The red and green disks had CIE coordinates of 0.621/0.349 and 0.308/0.587, respectively.

The stimulus arrays were randomly displayed in one of the four quadrants of the screen and lasted for 200 ms. The center of each stimulus array was 5.7° from the fixation and located on an imaginary line that went through the fixation and divided each quadrant into two equal parts. The interstimulus interval varied randomly between

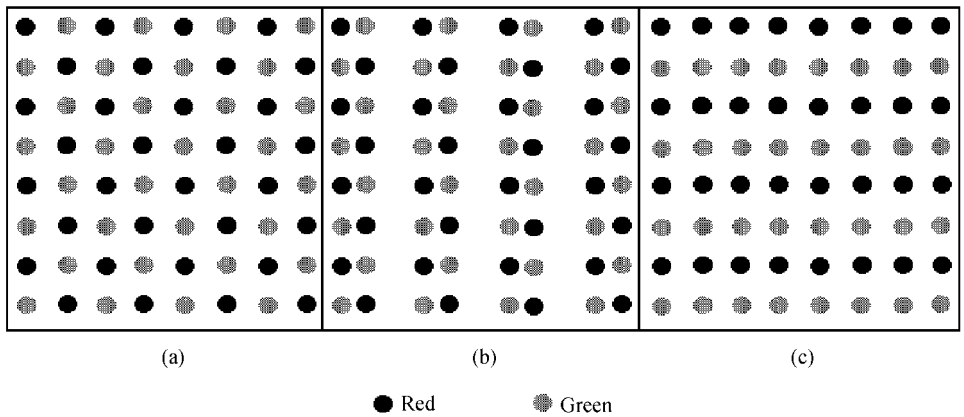


Fig. 1. Illustrations of the stimuli used in the present study. (a) The uniform stimulus; (b) the proximity-grouping stimulus in which local elements group into columns; (c) the similarity-grouping stimulus in which local elements group into rows.

1000 and 1400 ms. While keeping fixated at the fixation cross, subjects discriminated column versus row organizations of the grouping displays by button press with thumbs while ignoring the uniform stimuli. After 100 practice trials, subjects were presented with 2000 to 2600 trials in ten to thirteen blocks depending upon the amount of artifacts. The uniform, proximity-grouping, and similarity-grouping stimuli were presented randomly on 28%, 36% and 36% of the trials, respectively.

(ⅱ) Electrophysiological data recording and analysis. The electroencephalogram (EEG) was recorded from 120 scalp electrodes using an EEG/ERP system from NeuroScan Inc. The skin resistance of each electrode was made less than 5 kΩ. The recording from an electrode at the right mastoid was used as reference. Eye blinks and vertical eye movement were monitored with electrodes located below the left eye. The horizontal electro-oculogram was recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. The EEG was amplified (band pass 0.1—70 Hz) and digitized at a sampling rate of 250 Hz. The ERPs in each stimulus condition were averaged separately off-line with averaging epochs beginning 200 ms before stimulus onset and continuing for 1000 ms. Trials contaminated by eye blinks, eye movements, or muscle potentials exceeding 100 μV (peak-to-peak amplitude) at any electrode were excluded from the average.

Peak latencies were measured relative to stimulus onset. Statistical analysis was conducted at each pair of electrodes over the parietal and occipito-temporal regions and electrodes along the midline of the skull. The mean ERP amplitudes were subjected to repeated measure analysis of variance (ANOVAs) with factors being Grouping (proximity or similarity vs. uniform stimuli), Hemifield (left vs. right), Elevation (upper vs. lower), Hemisphere (left vs. right). The ANOVAs of behavioral data and ERP data at electrodes along the midline of the skull were conducted with Grouping, Hemifield, and Elevation as independent variables. Dipole modeling was used to localize the source of ERP components. Electrode positions were measured from each individual subject with a probe for sensing the 3-dimensional position of the probe tip with respect to a magnetic field source in the head support. Magnetic resonance (MR) images were obtained from 5 subjects for constructing realistic-head boundary-element models. The digitized fiducial landmarks corresponding to the electrode coordinates were coregistered with fiducial landmarks identified on whole-head MR scan so that the locations of estimated dipoles could be related to individual brain-skull anatomy. Dipoles were mapped onto the MR images of individual subjects to estimate source locations with respect to brain anatomy. The 3-dimensional coordinates of each dipole were transformed to the coordinates of Talairach and Tournoux[11] atlas by marking the anterior and posterior commissures on each subject's MR scan.

## 2 Results

(ⅰ) Behavioral data. Response accuracies were high (91.1% for proximity and 92.3% for similarity stimuli). RTs were slightly faster to proximity-than similarity-grouping stimuli (505±33.2 vs. 512±34.6 ms, $F(1,15) = 5.26$, $P < 0.04$). RTs were faster to the stimuli in the right than left visual fields (505±34.8 vs. 511±33.4 ms, $F(1,15) = 12.9$, $P < 0.003$) and to the stimuli in the lower than upper visual fields (506±34.1 vs. 511±34.1 ms, $F(1,15) = 6.43$, $P < 0.02$).

(ⅱ) ERP data. The ERPs recorded at middle parieto-occipital scalp sites showed that stimuli in the upper visual field evoked the C1 peaking at 70 ms and the P1 at 90—130 ms followed by the N1 at 100—200 ms. For stimuli displayed in the lower visual field, the C1 reversed polarity and was followed by the N1 (see Figs. 2 and 3). At lateral occipito-temporal sites, ERPs included an initial P1 and a following N1.

ANOVAs performed on mean amplitudes showed a

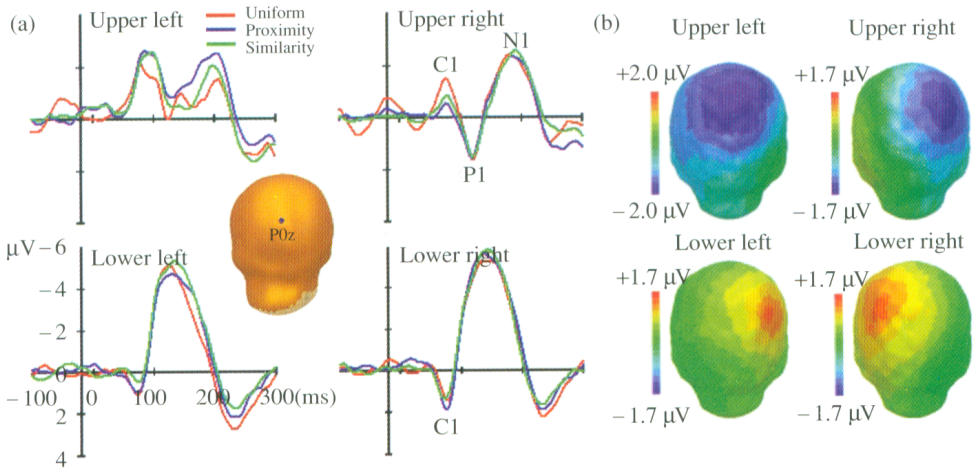


Fig. 2. (a) Grand average ERPs elicited by uniform and grouping stimuli recorded at parieto-occipital electrode (POz). (b) Voltage topographies calculated based on grand average ERPs to uniform stimuli at 60—80 ms. The foci of the positive C1 elicited by the stimuli in the lower VF are distributed slightly contralateral to the stimulated hemifields whereas this laterality effect is weaker for the negative C1 elicited by the stimuli in the upper VF.

Fig. 3. (a) Grand average ERPs elicited by uniform and grouping stimuli recorded at left lateral occipital electrode (P7). (b) Voltage topographies calculated based on grand average ERPs to uniform stimuli between 80 and 100 ms after stimulus onset. The P1 wave showed maximum amplitudes at lateral occipital areas contralateral to the stimulated hemifields.

Fig. 4. Dipole models showing the source that generated the C1 component. The best-fit dipolar source for the C1 at 60—80 ms is located to the calcarine cortex and shown in the MR images of a representative subject. The dipole orientations varied systematically as a function of stimulus positions. The positive pole of the dipole pointed inside the brain for the upper VF stimuli but outside the brain for the lower VF stimuli.

main effect of Grouping at parieto-occipital electrodes between 60 and 80 ms for the proximity-grouping condition ($F(1,15) = 5.59$, $P < 0.03$). The negative C1 was of smaller amplitude (less negative) to proximity-grouping than uniform stimuli in the upper visual field whereas the positive C1 was of larger amplitude (more positive) to proximity-grouping than uniform stimuli in the lower visual field. This C1 grouping effect was stronger for the upper than lower visual fields ($F(1,15) = 25.6$, $P < 0.001$). Moreover, the asymmetric elevation grouping effect was larger in the right than left visual fields ($F(1,15) = 32.9$, $P < 0.001$). No significant grouping effect was found in the P1 and the N1 time window. However, the P1 and N1 amplitudes were larger at electrodes contralateral than ipsilateral to the stimulated hemifields ($F(1,15) = 35.1$, $P < 0.001$). The effect of grouping was not significant in the C1 time window for similarity grouping condition ($F(1,15) = 3.05$, $P > 0.09$). However, the grouping effect at 60—80 ms was stronger when stimuli were presented in the upper than lower visual fields as indicated by the significant interaction of Grouping×Elevation ($F(1,15) = 25.5$, $P < 0.001$).

Voltage topography showed that the C1 elicited by the stimuli in the upper visual field showed negative foci over the parieto-occipital area whereas the C1 to the stimuli in the lower visual field had positive foci at the parieto-occipital region. The P1 component showed positive maximum amplitudes over the occipito-temporal regions contralateral to the stimulated hemifield regardless of stimulated elevations. The neural sources of C1 were estimated by dipole modeling based on realistic-head boundary-element models at 68—80 ms. The principle component analysis suggest that a single dipole provided the best solution for all stimulated locations during this interval. The best-fit C1 dipoles were localized to the calcarine fissure with Talairach coordinates of $x, y, z = -11, -76, -0.3$ (upper-right); 7, −72, 14 (upper-left); 23, −78, −6 (lower-left); and 7, −71, −6 (lower-right) (see Fig. 4). The goodness of fit of these dipole solutions (proportion of scalp variance accounted for) was 94% (upper-right), 86% (upper-left), 85% (lower-left), and 86% (lower-right), respectively.

## 3 Discussion

The present study showed that ERPs to either uniform or grouping stimuli were characterized by an early component at 60—90 ms post-stimulus. This component had maximum amplitudes over the parieto-occipital areas and was negative for stimuli in the upper visual field but positive for those in the lower visual field. This ERP component is identical to the C1 observed in the previous studies[9]. Our dipole modeling localized the C1 at 60—80 ms to the cortex close to the calcarine fissure, indicating that this early ERP wave possibly originates from the primary visual cortex. However, the dipole locations observed here are not exactly the same as the prediction of the cruciform model of the primary visual cortex[12], which states that stimuli presented to the upper and lower visual fields are represented by the cortex on the lower and upper banks of the contralateral calcarine fissure, respectively. The dipoles corresponding to the lower visual field stimuli were inferior to those associated with the upper visual field stimuli. It is possible that the low signal-to-noise ratio limited the precision of the dipole modeling used here. Alternatively, the incongruency between our results and the cruciform model may result from the overlap of the C1 with the early phase of the P1 in the lower visual field. As the P1 component had generators that were inferior to the C1 source (see the P1 voltage topographies), the dipoles calculated at 60—80 ms might reflect contributions of both the C1 and the early phase of the P1. The summation of the two components resulted in dipole solutions that are inferior to the area where the C1 sources are actually located.

Interestingly, the C1 was of smaller (less negative) amplitude to proximity-grouping than uniform stimuli in the upper visual field and of larger (more positive) amplitude in the lower visual field. Thus the proximity grouping generated a positive activity regardless of stimulus elevations rather than simply increased the C1 amplitude for stimuli in both the upper and lower visual fields. This is consistent with the previous ERP reports[6,7]. The current work complements the previous work by identifying the neural source of the grouping effect and provided ERP evidence that proximity grouping modulates activities of the primary visual cortex. The finding that perceptual grouping in humans has neural substrates as early as in the primary visual cortex is in agreement with the results of monkey studies that responses of neurons in V1 are modulated by grouping of stimuli inside and outside receptive fields[3]. Therefore it may be suggested that grouping operation is a common function of the primary visual cortex for both humans and monkeys. However, the significant grouping effect on the C1 was evident for proximity-grouping stimuli but not for similarity-grouping stimuli, suggesting that, relative to similarity, proximity is a factor that produces stronger grouping operations in the early visual cortex, which may contribute to the faster behavioral responses to proximity than similarity stimuli[4,5].

In addition, we found that the C1 grouping effect was stronger in the upper than lower visual fields. It has been hypothesized that processing in the lower visual field is more global and related to manipulations performed in peripersonal space whereas the upper visual field is primarily local and related to visual search and recognition mechanisms directed toward extrapersonal space[13]. If perceptual grouping reflects processing of global aspects of stimulus arrays, our observation is contrary to the

above hypothesis. However, the ecological significance of our results is still unclear.

Unlike the C1 component, the P1 with maximum amplitudes over lateral occipital areas was not influenced by grouping of local elements, nor was the following N1 component. Thus the C1 effect reflects grouping operation that is specific to the visual cortex close to the calcarine fissure. The lateral extrastriate cortex, where the P1 is generated[9,10], may not play an important role in the process of perceptual grouping.

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (Grant No. 30225026), the Ministry of Science and Technology of China (Grant No. 2002CCA01000), and the Learning & Cognition Lab, Capital Normal University.

## References

1. Lamy, D., Tsal, Y., On the status of location in visual attention, European Journal of Cognitive Psychology, 2001, 13: 305—402.
2. Vecera, P.V., Behrmann, M., Attention and unit formation: A biased competition account of object-based attention, in From Fragments to Objects——Segmentation and Grouping in Vision (eds. Shipley, T. F., Kellman, P. J.), London: Elsevier, 2001, 145—182.
3. Sugita, Y., Grouping of image fragments in primary visual cortex, Nature, 1999, 401: 269—272.[DOI]
4. Han, S., Humphreys, G. W., Interactions between perceptual organization based on Gestalt laws and those based on hierarchical processing, Perception & Psychophysics, 1999, 61: 1287—1298.
5. Han, S., Humphreys, G. W., Chen, L., Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of closure, Journal of Experimental Psychology: Human Perception Performance, 1999, 25: 1411—1432. [DOI]
6. Han, S., Song, Y., Ding, Y. et al., Neural substrates for visual perceptual grouping in human, Psychophysiology, 2001, 38: 926—935.
7. Han, S., Ding, Y., Song, Y., Neural mechanisms of perceptual grouping in humans as revealed by high density event related potentials, Neuroscience Letters, 2002, 319: 29—32. [DOI]
8. Kourtzi, Z., Tolias, A. S., Altmann, C. F. et al., Integration of local features into global shapes: Monkey and human fMRI studies, Neuron, 2003, 37: 333—346. [DOI]
9. Clark, V. P., Fan, S., Hillyard, S. A., Identification of early visual evoked potential generators by retinotopic and topographic analyses, Human Brain Mapping, 1995, 2: 170—187.
10. Noesselt, T., Hillyard, S. A., Woldorff, M. G. et al., Delayed striate cortical activation during spatial attention, Neuron, 2002, 35: 575—587. [DOI]
11. Talairach, J., Tournoux, P., Co-Planar Stereotaxic Atlas of the Human Brain, New York: Thieme, 1998.
12. Holmes, G., The organization of the visual cortex in man, Proceedings of Royal Society London [Biology], 1945, 132: 348—361.
13. Previc, F. H., Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications, Behavior and Brain Science, 1990, 13: 519—575.

(Received January 5, 2004; accepted March 1, 2004)