Perception, 2005, volume 34, pages 237 – 240

DOI:10.1068/p5412

# Last but not least

**Abstract.** When two three-letter words are flashed up in sequence, observers cannot tell whether the top halves of the words are the same or different. It follows that words, like faces, are processed holistically, not as a set of separate features.

## Holistic word processing

Should we think of the brain as a general-purpose computer that can be programmed and re-programmed to do an almost infinite variety of tasks? Or is it more like a collection of pre-set modules, each specialized to perform only one particular task, but to do it extremely well? The task of the visual system is to ask questions of the environment, and since some questions are probably asked more frequently than others, it might make sense to evolve 'fast, dumb mechanisms' or modules that are specialized to answer such FAQs. Fodor (1983) proposed that visual modules do exist, and are characterized by hidden internal workings but obligatory outputs. Examples, as discussed by Nakayama (2001), include color vision and depth perception. The workings of color vision are hidden from consciousness, since we are unaware of the outputs of our retinal cones or of our opponent R/G and B/Y channels. The outputs from the color module are obligatory, in that we are unable to see a colored scene as drained of its colors, but perforce see it in full color. Similarly, in stereo vision we cannot access the view seen by each eye, but the output is obligatory in that we cannot avoid seeing the depth in (say) a random-dot stereogram.

More to our purpose, there is ample evidence of such modules for (at least) two tasks, namely face recognition and the use of language. Face recognition is handled largely by the temporal lobes, as evidenced by single-unit studies (reviewed by Gross and Sergent 1992). Face-selective neurons appear to be members of ensembles for coding faces rather than individual face detectors or grandmother cells. In humans, damage to different areas can produce a variety of face-processing impairments that reflect interference at different levels of processing of the facial image. The most remarkable such deficit is known as prosopagnosia, or an inability to recognize previously known familiar faces. Prosopagnosia is caused by bilateral lesions of the mesial occipitotemporal regions (Damasio 1985).

Language is handled by the left hemisphere in the vast majority of people, with Broca's area responsible for speech production and the more posterior Wernicke's area responsible for language comprehension. The left hemisphere also handles reading. This is evidenced by alexia, or the inability to read. Lesions to interhemispheric fibers in the left paraventricular region (beneath and beside the occipital horn of the left lateral ventricle) can cause alexia without agraphia, sometimes accompanied by color anomia (an inability to name colors), whereas lesions to the left inferior angular gyrus can cause alexia together with agraphia (Damasio 1983).

Psychophysical evidence for face modules in the normal brain comes from holistic processing of faces. It is widely agreed that faces can be perceived in two ways: either by separate inspection of individual features, or else by holistic or configural perception, defined as an obligatory processing of all parts (Diamond and Carey 1986; Tanaka and Farah 1993). Feature-based recognition relies upon information from local regions of the face such as nose shape, eye color, or hairline. For instance, Sinha and Poggio (1996) presented an illusion in which Bill Clinton's face, when shown with Al Gore's hair, produced an initial percept of Gore. Here, the hairline feature drove the recognition process. Configural processing relies upon information integrated across the entire internal face region, which is matched to a stored representation coding the expected form of an upright face (McKone et al 2001). Evidence for configural processing comes from the detrimental effect of disrupting facial layout, even when this leaves the individual features intact. Thus, face recognition is considerably impaired by scrambling the position of the features (Tanaka and Farah 1993) or by 'exploding' them into separated face parts (Farah et al 1995).

Nakayama (2001) reviews some interesting neurological evidence for face-specific and word-specific modules in the brain. For instance, Caramazza and Hillis (1990) studied a brain-damaged patient with a left parietal lesion that led to unilateral neglect. The patient could read the left but not the right half of words. This was still true even when the words were displayed horizontally, vertically, or mirror-reversed, so the damage was not merely to the perceived spatial layout of the word on the page, but to a more subtle abstract internal representation of the word. This very specific pattern of loss suggests the existence of underlying word-processing modules.

Coltheart et al (1993) found that in some cases of acquired dyslexia patients could read whole words and understand their meaning, yet they could not read the individual letters of the words. This indicates the existence of a 'whole word' module that makes available the identity of the word but not its constituents. This is a remarkably close analogy to a case of face recognition in a brain-damaged patient (CK) studied by Moscovitch et al (1997). This patient had lost the ability to recognize objects, but had retained the ability to recognize faces. CK was shown an ambiguous trompe-l'oeil portrait by Giuseppe Arcimbaldo (1530–1593), comprising a face that was composed of vegetables. The patient could perceive the face but not the vegetables! This suggests the existence of a module specific for a 'whole face' but not for its constituent features. Conversely, a typical prosopagnosic patient would perceive the vegetables but not the face.

Configural processing is already present in 4-year and 5-year olds (Pellicano and Rhodes 2003), although it develops more slowly than feature perception in the first ten years of life (Mondloch et al 2002). In a striking demonstration of configural processing, two faces can be presented in quick succession, where the top half (including the eyes) remains the same but the bottom half (including the mouth) changes. When observers are asked to report whether the top half is the same or different in the two faces, observers often mistakenly report that they are different. Thus, observers are unable to perceptually isolate the features in the upper half of the face, but are obliged to process the face as a whole (Farah et al 1998).

In this paper, I use the same technique to show that configural processing is not confined to faces, but can also apply to the perception of words. (I did not use non-word letter strings.) Two three-letter words were presented in quick succession, and observers were asked to mark on a printed form whether the *upper halves* were the same or different in the two words. The duration of each stimulus word was 40 ms, with an interstimulus interval (ISI) of 520 ms. Patterns were drawn in black lines on a white ground, and were controlled by a Macromedia Director program running on a Macintosh computer. The display was projected on a large screen in front of a class of 140 undergraduate students, who viewed the projection screen from a wide range of distances and angles.

Four pairs of three-letter words were presented (figure 1). These were hand-drawn so that their top halves were identical and only the bottom halves were different. (In most commercial fonts, both top and bottom parts tend to be made as different as possible in the interests of legibility). The word pairs were hoe–bag, ban–hop, nag–pea, and pen–nap. Although the top halves within each word pair were identical, 54% of the students incorrectly identified them as being different. These errors were not simply due to the general difficulty of the task. For comparison, a control word pair (kit–kit) that was identical top and bottom was correctly identified as 'the same' by 94.3% of the

| Stimulus | Error rate/% | Stimulus | Error rate/% |
|---|---|---|---|
| hoe<br>bag | 61 | nag<br>pea | 55 |
| ban<br>hop | 50 | pen<br>nap | 51 |

**Figure 1.** Word pairs whose top halves are identical with percentage of observers who erroneously reported that the top halves were different. Results show that holistic word processing prevented observers from perceptually isolating the top halves of the words.

observers, and another pair of control words that had different tops and bottoms (nap – kit) was correctly identified as 'different' by 97.8%. I conclude that the errors were made because students had difficulty in perceptually isolating the top halves from the bottom halves of the words. This means that students were processing the words not as separable parts, but holistically as perceptual units that could not be perceptually split apart. These results show that in normal circumstances, the visual system cannot, or does not, divide words into upper and lower halves. Admittedly, while being perceived, the words could still be decomposed in other ways: into individual letters, for example. Therefore the results of this experiment are limited in their implication for the important issue whether or not there are two separable, modular systems for visual recognition, namely, holistic versus part-based, the impairment of which underlies prosopagnosia and alexia, respectively, with visual object agnosia being a mixture of the two.

In deciding whether modules really exist, everybody can agree that the perception of faces and the perception of non-face objects share common early visual processing stages. At this point, opinions divide in favor of one or other of two possibilities: Either the brain eventually processes faces separately from other objects, within a domain-specific module dedicated to face perception, or this apparent specialization for faces could result from people's expertise with this category of stimuli. Gauthier et al (2003) measured interference between face processing and object processing. If the expert processing of faces and cars depended on common mechanisms related to holistic perception (obligatory processing of all parts), then for human subjects who are presumed to be face experts, car perception should interfere with concurrent face perception. Furthermore, such interference should increase with greater expertise in car identification, and indeed this is what Gauthier et al found. Event-related potentials (ERPs) suggested that this interference arose from perceptual processes contributing to the holistic processing of both objects of expertise and faces. Gauthier et al conclude that face perception is not modular. However, this conclusion may be premature, since the learned skills of car perception might conceivably be piggy-backing on the more basic skills of face perception.

In summary, there are two possible criteria to define what a module is. Either the modularity of the face-recognition system can be defined by the kinds of stimuli being processed. On this argument the face-recognition system is not modular because it can process other objects such as cars (Gauthier et al 2003). This argument suffers from the fact that, in the remarkably plastic human brain, highly learned responses can overlay and obscure the operation of prior modules. However, if we define modularity in terms of the mode of operation of the system, namely holistic versus part-based, then the weight of evidence, as briefly reviewed above, favors modularity for both face- and word-recognition systems and the present results add a mite to this evidence.

The errors I found can be regarded as a new class of Stroop effect. Stroop tasks (reviewed by Melara and Algom 2003) present stimuli having two dimensions, and participants respond to one dimension whilst ignoring the other. Thus Stroop (1935) presented color-words printed in colored inks, for example the word GREEN printed in red ink, and asked observers to name the ink. Observers were slow to say "red" (the correct response), because of interference from the overlearned, incorrect response "green". Similarly, in this experiment the same overlearned response of reading the letters in words interfered with the unusual task of responding only to parts of the letters.

It would be interesting to repeat such experiments with languages that are unknown to the observers, such as Chinese characters shown to non-speakers of Chinese. It is known that the color Stroop effect does not work for color words printed in an unfamiliar language. Possibly, Western observers would find it easier to isolate the top halves of unfamiliar Chinese characters than they do for the overlearned characters of the Roman/English alphabet.

In conclusion, the physiological evidence for specialized modules to process faces and language, briefly reviewed above, is very strong. The finding that words, like faces, can be processed holistically, provides additional psychophysical evidence for the existence of these modules.

Stuart Anstis

Department of Psychology, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, USA; e-mail: sanstis@ucsd.edu

**Acknowledgments.** Supported by grants from the Academic Senate and Department of Psychology, University of California at San Diego. I thank an anonymous referee for helpful comments. Thanks to Professors Oliver Braddick and Brian Rogers, Department of Psychology, University of Oxford, for their hospitality, and to the Master and Fellows of Pembroke College, Oxford, for electing me to a Visiting Fellowship, during the summer sabbatical when this work was completed.

**References**

Caramazza A, Hillis A E, 1990 "Spatial representation of words in the brain implied by studies of a unilateral neglect patient" *Nature* **346** 267 – 269

Coltheart M, Curtis B, Atkins P, Heller M, 1993 "Models of reading aloud: Dual-route and parallel-distributed-processing approaches" *Psychological Review* **100** 589 – 608

Damasio A R, 1983 "Pure alexia" *Trends in Neurosciences* **6** 93 – 96

Damasio A R, 1985 "Prosopagnosia" *Trends in Neurosciences* **8** 132 – 135

Diamond R, Carey S, 1986 "Why faces are and are not special: An effect of expertise" *Journal of Experimental Psychology: General* **115** 107 – 117

Farah M J, Tanaka J N, Drain M, 1995 "What causes the face inversion effect?" *Journal of Experimental Psychology: Human Perception and Performance* **21** 628 – 634

Farah M J, Wilson K D, Drain M, Tanaka J N, 1998 "What is 'special' about face perception?" *Psychological Review* **105** 482 – 498

Fodor J, 1983 *Modularity and Mind: An Essay on Faculty Psychology* (Cambridge, MA: MIT Press)

Gauthier I, Curran T, Curby K M, Collins D, 2003 "Perceptual interference supports a non-modular account of face processing" *Nature Neuroscience* **6** 428 – 432

Gross C G, Sergent J, 1992 "Face recognition" *Current Opinion in Neurobiology* **2** 156 – 161

McKone E, Martini P, Nakayama K J, 2001 "Categorical perception of face identity in noise isolates configural processing" *Journal of Experimental Psychology: Human Perception and Performance* **27** 573 – 599

Melara R D, Algom D, 2003 "Driven by information: a tectonic theory of Stroop effects" *Psychological Review* **110** 422 – 471

Mondloch C J, Le Grand R, Maurer D, 2002 "Configural face processing develops more slowly than featural face processing" *Perception* **31** 553 – 566

Moscovitch M, Winocur G, Berhmann M, 1997 "What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition" *Journal of Cognitive Neuroscience* **9** 555 – 564

Nakayama K, 2001 "Modularity in perception: its relation to cognition and knowledge", in *Blackwell Handbook of Perception* Ed. E B Goldstein (Oxford, UK: Malden, MA: Blackwell) pp 737 – 759

Pellicano E, Rhodes G, 2003 "Holistic processing of faces in preschool children and adults" *Psychological Science* **14** 618 – 622

Sinha P, Poggio T, 1996 "I think I know that face" *Nature* **384** 404

Stroop J R, 1935 "Studies of interference in serial verbal reactions" *Journal of Experimental Psychology* **18** 643 – 662

Tanaka J N, Farah M J, 1993 "Parts and wholes in face recognition" *Quarterly Journal of Experimental Psychology: Human Experimental Psychology A* **46** 225 – 245

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 34 2005

www.perceptionweb.com