Review

Letter perception: from pixels to pandemonium

Jonathan Grainger, Arnaud Rey and Stéphane Dufau

CNRS and Aix-Marseille University, Laboratoire de Psychologie Cognitive, Université d'Aix-Marseille I, 3 place Victor Hugo, 13331 Marseille, France

In 1959, Oliver Selfridge proposed a model of letter perception, the Pandemonium model, in which the central hypothesis was that letters are identified via their component features. Although a consensus developed around this general approach over the years, key evidence in its favor remained lacking. Recent research has started to provide important evidence in favor of feature-based letter perception, describing the nature of the features, and the time-course of processes involved in mapping features onto abstract letter identities. There is now hope that future 'pandemonium-like' models will be able to account for the rich empirical database on letter identification that has accumulated over the past 50 years, hence solving one key component of the reading process.

What is the letter 'a'?

The cognitive scientist and philosopher Douglas Hofstadter once noted [1] that 'The central problem of Artificial Intelligence is the question: what is the letter *a*?' What Hofstadter was suggesting in his provocative statement is that understanding the mechanisms underlying invariant recognition of the arbitrary signs that compose the Roman alphabet ($a = A \neq b$) will be a major step towards understanding the essence of human intelligence. Letters represent a perfect example of the kind of symbol that humans thrive on, and letters are sufficiently limited in complexity and number to provide a highly tractable domain of investigation. Letters are also the gateway to reading [2–4], perhaps the most complex skill that humans have to master without specific genetic predisposition. Furthermore, a letter-based strategy for reading in alphabetical orthographies has probably developed because it is far more economical to solve shape invariance for 26 letters compared with tens of thousands of words [3], and understanding shape-invariant recognition is a major endeavor of current research on visual object perception [5]. In this article we review exciting new developments in this central topic of cognitive science. Recent research provides converging evidence in support of the classic account of letter perception formulated by Oliver Selfridge 50 years ago (the Pandemonium model), and the hope that a complete account of the processes involved in recognizing a letter of the alphabet is within reach.

Pandemonium in the air

The starting point of contemporary research in this field is Selfridge's [6,7] seminal work, laying the foundations for a cognitive theory of letter perception. In Pandemonium, letter identification is achieved by hierarchically organized layers of feature and letter detectors. Support for such a hierarchical organization was provided at that time by neurophysiological studies of the cat visual cortex [8]. In spite of this evidence, an alternative theoretical approach, template-matching, has also been favored by cognitive scientists. Template-matching models assume that several shape-exemplars of a given letter are stored in memory, and that recognition consists in finding the best match between a target item and one of these memory traces. In this approach, a new template is learned and stored each time a new target stimulus differs notably from existing templates. Simple versions of template-matching compare descriptions of the stimulus as a set of pixel intensities with corresponding representations in long-term memory, but the distinction between feature-based and template-matching approaches can be blurred by incorporating template matchers as feature detectors in Pandemonium [9]. Nevertheless, the major drawback with template-matching models, as already noted by Neisser [10], is that the matching procedure requires prior normalization of the stimulus (adjusting the stimulus to a prototypical position, size and orientation), and proposals for such a process in most cases lack psychological and neurophysiological plausibility. Furthermore, a general consensus has developed over the years in favor of feature-based approaches. What is the key evidence for this, and what are the features?

Show me the features

The confusion matrix is the traditional method used to hunt for features. In a typical experiment used to generate a confusion matrix, isolated letters are presented in data-limited conditions (brief exposures and/or low luminance and/or masking) and erroneous letter reports are noted. Error rate (e.g. reporting F when E was presented) is hypothesized to reflect visual similarity driven by shared features. An analysis of the pattern of letter confusions was therefore expected to reveal the set of features used to identify letters. There are >70 published studies on letter confusability, and some have formed the basis of concrete proposals of lists of features for letters of the Roman alphabet, mainly consisting of lines of different orientation and curvature [11–13].

One major drawback of standard letter confusion data is that the method used to degrade stimuli (to generate confusion errors) influences the nature of the confusions

Corresponding author: Grainger, J. (jonathan.grainger@univ-provence.fr).

1364-6613/$ – see front matter doi:10.1016/j.tics.2008.06.006 Available online 27 August 2008

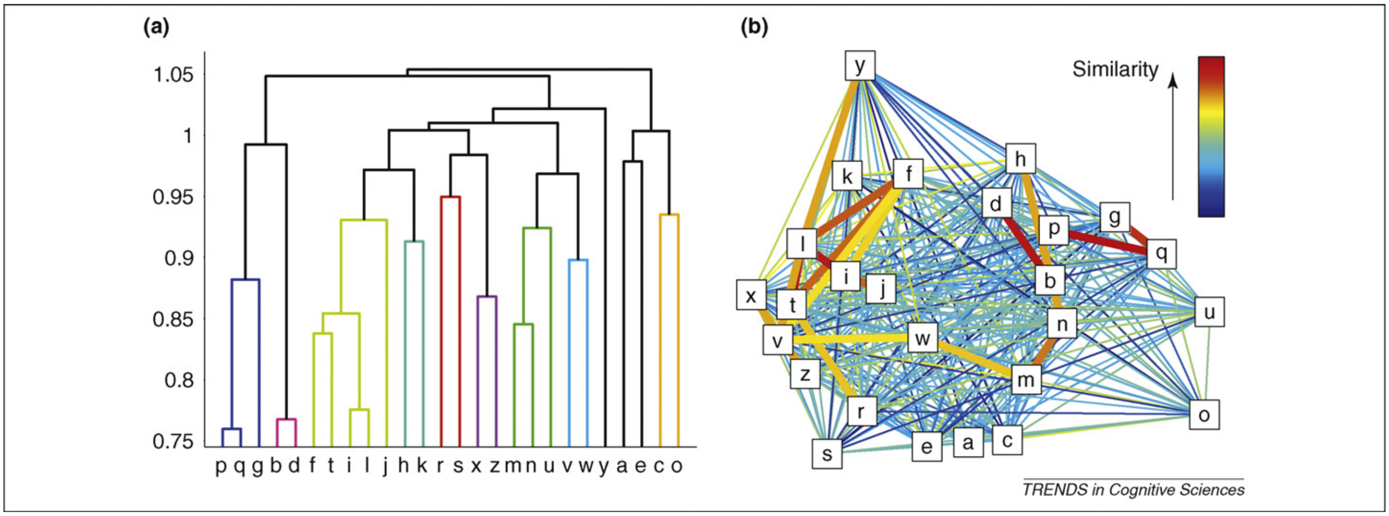

Figure 1. Illustration of inter-letter similarities (Euclidian distances) revealed by a same–different matching experiment [16]. **(a)** A dendogram plot in which the height of connecting bars reflects Euclidian distance (smaller = closer). Inter-letter similarities form clusters shown in different colors. **(b)** A projection of the same Euclidian distances on a plane [44]. The original distances in high-dimensional Euclidian space are coded by the color of connecting lines. In both panels it can be seen that letters with specific combinations of simple features group together, such as 'vertical line' plus 'circle' (b, d, g, p, q). This corresponds to the main similarity class obtained by principal component analysis. The second similarity class, corresponding to small curvilinear shapes (a, c, e, o, s), is also visible, in addition to combinations of similarity classes such as 'vertical line' plus one small feature (f, t, i, j, l).

and furthermore confounds perceptual confusions with post-perceptual guessing. In one of the first studies to overcome this drawback, Podgorny and Garner [14] used a same–different matching task and showed that the resulting discrimination-time matrix for letters correlated well with judgments of perceptual similarity (see also Ref. [15] for confusion matrices expressed as saccade latencies). Furthermore, having response time (RT) as the dependent measure avoids the problem of empty cells and provides a ratio scale that enables the application of more powerful metric analyses. More recently, Courrieu, Farioli and Grainger [16] used a Go–NoGo variant of the same–different matching task, with participants responding only when the two letters were different. The discrimination times were transformed into Euclidean distances by using a 'monotonic embedding' technique [17]. A principal component analysis revealed 25 dimensions, many of which were clearly interpretable as elementary visual features (Figure 1).

A new look at features

Two recent articles open up an exciting new perspective for research investigating feature-based letter perception. Pelli *et al.* [18] measured contrast thresholds to Roman letters presented in different fonts, in addition to letters and characters from other languages. The authors expected letters to be identified optimally, like a single feature (i.e. a spatial frequency channel or filter in a grating experiment). This is what would be expected from a template-matching approach. To test this, Pelli *et al.* [18] measured efficiency of letter identification (Box 1) under varying viewing conditions and found that efficiency was independent of stimulus duration, eccentricity and size, but did vary across different alphabets and fonts. The sub-optimal performance

Box 1. The psychophysics of letter perception

The visual world can be described in terms of variations in spatial frequency, that is, changes in luminance across space [45]. Letters have a broad spatial frequency spectrum, so what part of the spectrum is used to identify letters? Two main techniques have been used to answer this question – measuring identification thresholds for bandpass filtered letters, and measuring variations in identification thresholds as a function of the bandpass characteristics of a masking stimulus (critical-band masking, Figure 2 in main text). The key to both approaches involves comparison of human performance with that of an ideal observer. Letters are better identified with high-pass filters than with low-pass filters, but this is because more information is available in the high-pass stimuli. Ideal observer analysis enables information availability to be equated, hence enabling biases in information utilization to emerge [46]. The measure of human performance relative to the ideal observer is called efficiency.

On the one hand, several studies have reported that observers identify octave-band wide filtered letters with almost as much efficiency as unfiltered letters for all but extreme (very high and very low frequency) bandpass filters [46–48]. On the other hand, critical-band masking experiments indicate that only a single channel is used to identify broadband (i.e. unfiltered) letters, with the centre frequency of the channel varying as a function of letter size, font and alphabet [21,49]. This could be because the presence of noise in the critical-band masking procedure makes it harder to learn to use different spatial frequency channels in the course of an experiment [50] or because filtered stimuli force participants to use information that they would otherwise ignore.

This line of research was extended further in the recent work of Fiset *et al.* [20] applying the 'bubbles' technique of Gosselin and Schyns to letter identification. The bubbles technique aims to uncover the parts of the image that are diagnostic of the observer's performance (i.e. correct letter identification). The technique, therefore, not only manipulates spatial frequency filtering, but also explores the significance of different parts of the image. As illustrated in Figure 2 (In the main text), the image is first band pass filtered at different channel frequencies and a random set of 'bubbles' is extracted from the filtered letters. The sampled images are summed across frequency channels to generate a 'bubbelized' image presented to observers. Variation in performance is then traced back to information available in the image (coordinate of bubble center) at each frequency band using image classification procedures (least-squared multiple linear regression on spatial co-ordinates and performance).

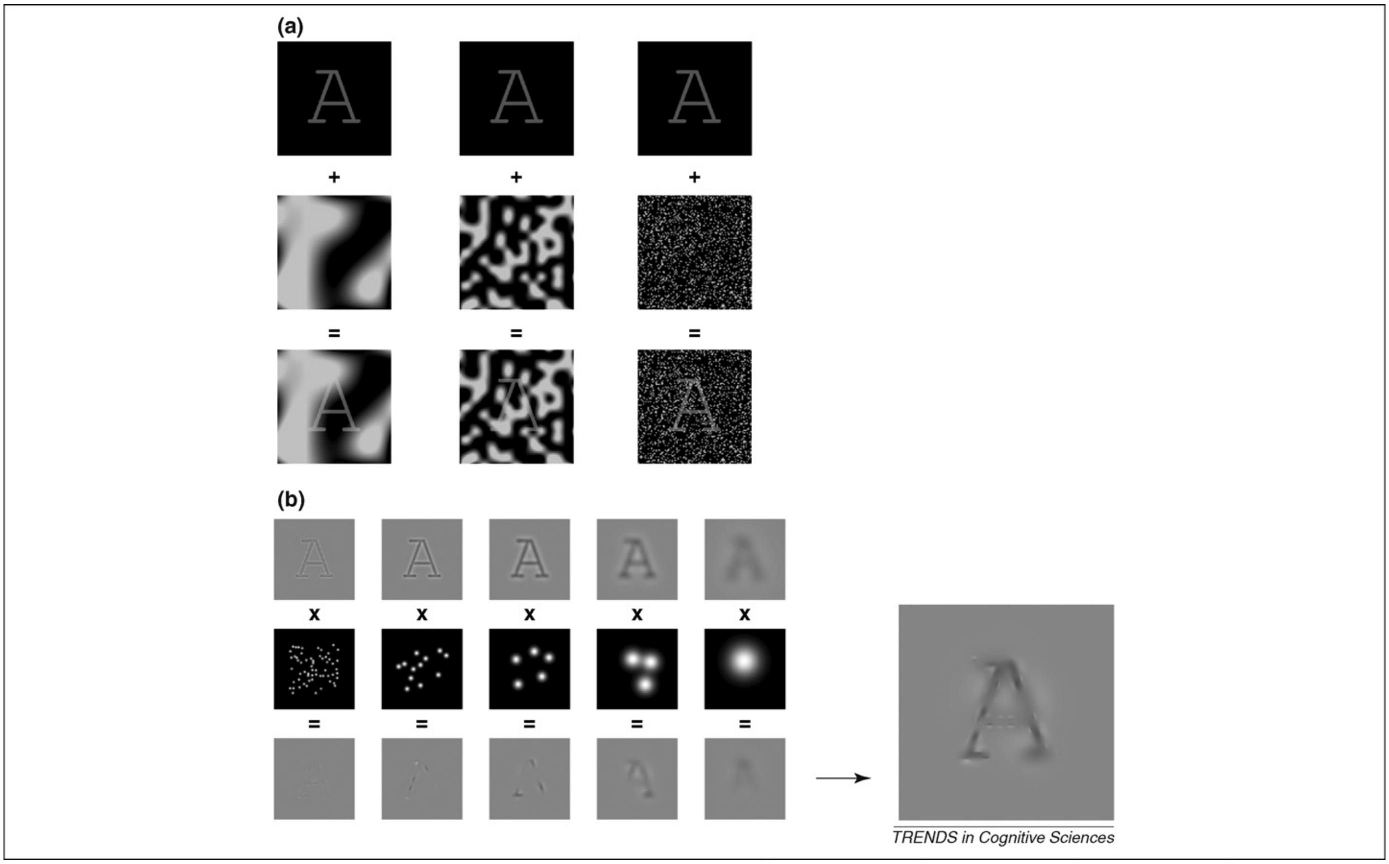

Figure 2. **(a)** Illustration of critical-band masking. Broadband letters (top row) are combined with bandpassed noise (middle row) to form a masked letter stimulus presented for identification. Three cycles per letter noise (middle) acts as a better mask than masks formed by either lower (left) or higher (right) spatial frequencies. **(b)** Illustration of the 'bubbles' technique applied to letters. Letters are first bandpass filtered with a series of five octave-wide spatial frequency channels (top row). Each filtered image is then randomly sampled (randomly selected *x* and *y* coordinates), and each sample smoothed by a Gaussian kernel with standard deviation proportional to the channel frequency. This generates a set of 'bubbles' representing parts of the image (middle row), the number and size of which is determined by the channel frequency of the filtered letter in the top row. The bubbles are then used as a spatial window that determines what is extracted from the filtered images (top row) to form the parts of the image shown in the bottom row. The sum of these image parts forms the final bubbelized image of the letter A (right image). Adapted, with permission, from Ref. [20].

of human observers was therefore taken to reflect feature-based letter identification, in which identification of the whole is affected by the identification of each component feature. What is it that changes across alphabets and fonts that might be driving these changes in efficiency? Pelli *et al.* [18] found one particular measure that correlated highly with letter identification efficiency. That was perimetric complexity – the square of the length of inside and outside perimeter, divided by ink area (for size invariance). In the absence of independent evidence concerning the nature of the features subtending letter identification, perimetric complexity provided a measure of visual complexity thought to be proportional to the number of features.

The second breakthrough has come from recent research applying Gosselin and Schyns' [19] 'bubbles' technique (Box 1 and Figure 2) to explore the nature of the critical features for letter perception. The classification images obtained by Fiset *et al.* [20] for 26 lowercase and 26 uppercase Roman letters in Arial font revealed several important pieces of evidence. First, on average only 32% of the printed area of uppe and 24% of lowercase letters was used by observers to identify letters, and the greatest proportion of useful information was apparent in the 2–4 cycles per letter frequency band, in line with estimates from critical-band masking studies [21]. Second, the analysis revealed that terminations were by far the most diagnostic piece of information for letter identification, with intersections and horizontal lines providing further significant sources of information for uppercase letters. For example, the letter W was mainly distinguished from other letters by the presence of two terminations, one in the upper left corner and the other in the upper right corner.

The time-course of letter perception

Standard behavioral measures of letter identification, such as percentage of correct responses in data-limited conditions and RTs to non-degraded stimuli, all represent the final result of an accumulation of component processes. However, a complete understanding of letter perception must incorporate knowledge about how the component processes develop through time. The masked priming paradigm [22] has proven its utility as a tool for examining the earliest phases of visual word recognition (see Ref. [3] for review), and has been usefully applied with letter stimuli to separate out the role of visual factors from phonological and articulatory factors in letter identification and letter naming (Box 2). Masked priming can also be combined with measures of brain activity that provide a moment-to-moment reflection of on-going target processing. In a recent study this combination was used to investigate the time-course of component processes in letter identification [23]. This work revealed a cascade of

Box 2. Masked priming with single letters

In the masked priming paradigm [22], prime stimuli are presented briefly enough to prevent the use of conscious predictive strategies that often contaminate standard priming effects. This paradigm is ideally suited for investigating early perceptual processing of familiar visual objects such as letters, without having to degrade target stimuli, and therefore complements the psychophysical procedures described in Box 1. In the majority of masked priming studies of letter identification performed up to now, primes were complete letters that varied in terms of their visual similarity with the target (e.g. c-C versus a-A) and whether or not they were nominally identical to the target (e.g. a-A versus c-A). The two main tasks used in these studies were found to be differentially sensitive to these two priming manipulations. The alphabetic decision task (speeded classification of letters versus non-letters) was found to be more sensitive to visual overlap than nominal overlap [51–54], whereas the letter naming task was much more sensitive to nominal overlap than to visual overlap [51,54,55] (Figure I). One methodological conclusion from this research is that the letter naming task might be overly sensitive to phonological-articulatory factors, which are interesting in their own right, but render the task relatively insensitive to visual factors. However, phonological priming effects in letter naming only occur with complete phonological overlap across primes and targets. Similar sounding letters (P-B) do not facilitate letter naming relative to different sounding letters (F-B) [51,55], whereas homophones of letter targets (e.g. sea-C) facilitate letter-naming responses to the same extent as same letter primes (c-C) [51]). The absence of priming from similar sounding letters might be because the overlap is practically always on the second phoneme, with mismatching information on the initial phoneme. Masked priming studies of word and object naming have shown that initial phoneme overlap is the key factor driving priming effects [56–58].

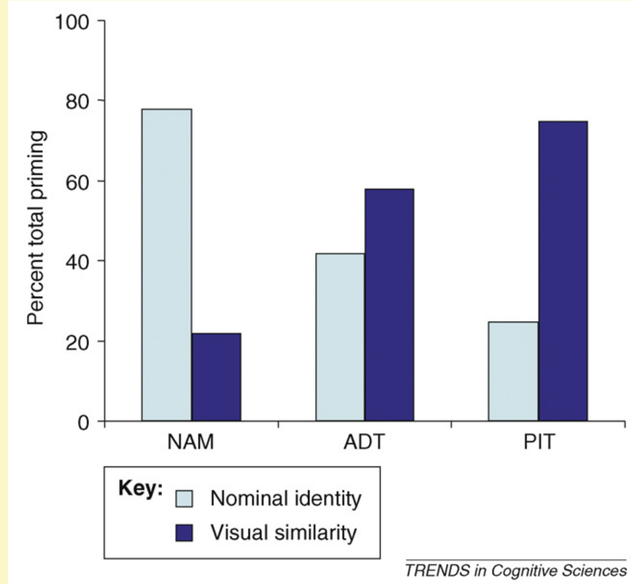

Figure I. Effects of nominal identity and visual similarity in three different tasks combined with masked priming: letter naming (NAM), alphabetic decision (ADT) and perceptual identification (PIT). Effects of nominal identity are calculated by comparing visually dissimilar same letter primes (a-A) and different letter primes (b-A). Effects of visual similarity are calculated by subtracting the priming effect of a visually similar same letter prime (c-C versus b-C) from the effect of a visually dissimilar prime (a-A versus b-A). Net priming effects are averaged over relevant conditions in two studies testing these conditions [51,54], and expressed as a percentage of the combined priming effect size for each task.

effects in the event-related potential (ERP) signal as a function of (i) prime-target visual overlap (peaking at 150 ms post-target onset); (ii) whether or not primes were the same letter as targets in the same case (180 ms post-target) and (iii) whether or not primes were the same letter as targets independently of case (200 ms post-target). This constitutes important evidence in favor of a generic hierarchical model of letter identification in which visual features are mapped onto abstract letter identities via a series of increasingly invariant representations.

A comparison of ERP waveforms generated by letter and pseudo-letter stimuli can also provide useful information about the time-course of letter identification. One study found that the amplitude of the N170 component is larger for letters than pseudo-letter (false font) stimuli [24]. Another study using pseudo-letters that were matched to the letter stimuli in terms of component features, found that ERP waveforms diverged as early as 145 ms post-stimulus onset [25] (Figure 3). In this study, amplitude of the P2 ERP component differed across the 14 letters that were tested, but some information about individual letter identities was already available in the waveform before the peak of the P2. This was indicated by the fact that item-level voltage values in this time window were found to correlate significantly with predicted letter identification latencies derived from different versions of a generic interactive-activation model of letter perception, comprising feature and letter detector layers [4]. The best model was one with excitatory feedforward and feedback connections between layers, and within-level inhibition across letters, but no between-level (feature-letter) inhibition.

The likely neural source of this early letter-specific brain activity is left occipital-temporal cortex [26–32]. This fits well with research locating orthographic processing associated with printed words in a small strip of left fusiform gyrus called the visual word form area (VWFA, Ref. [33]). Single letters dissociated from objects, faces, digits and letter strings in an area situated anterior to the peak of the VWFA [27,28,32] (Figure 3). Given their more anterior location, neurons in this area possibly instantiate abstract location-invariant letter detectors to be distinguished from more location-specific letter detectors involved in processing letter strings [34,35]. Furthermore, the fact that the letter-specific region and the VWFA fall within a more general object processing region [29] is in line with the neuronal recycling hypothesis of Dehaene *et al.* [34].

Putting the parts together

Recent evidence in favor of a generic feature-based hierarchical approach to letter perception therefore provides hope that a complete account of the processes involved in identifying isolated letters is within our grasp. More sophisticated models, such as the one shown in Figure 4, will probably include multiple layers of simple and complex features converging on case-specific and possibly font-specific letter detectors, which map in turn onto more abstract shape-invariant letter representations [23,34, 36–38].

There are several key questions that need to be addressed in future developments of hierarchical feature-based models. One concerns how the spatial relations between the different features are coded, if at

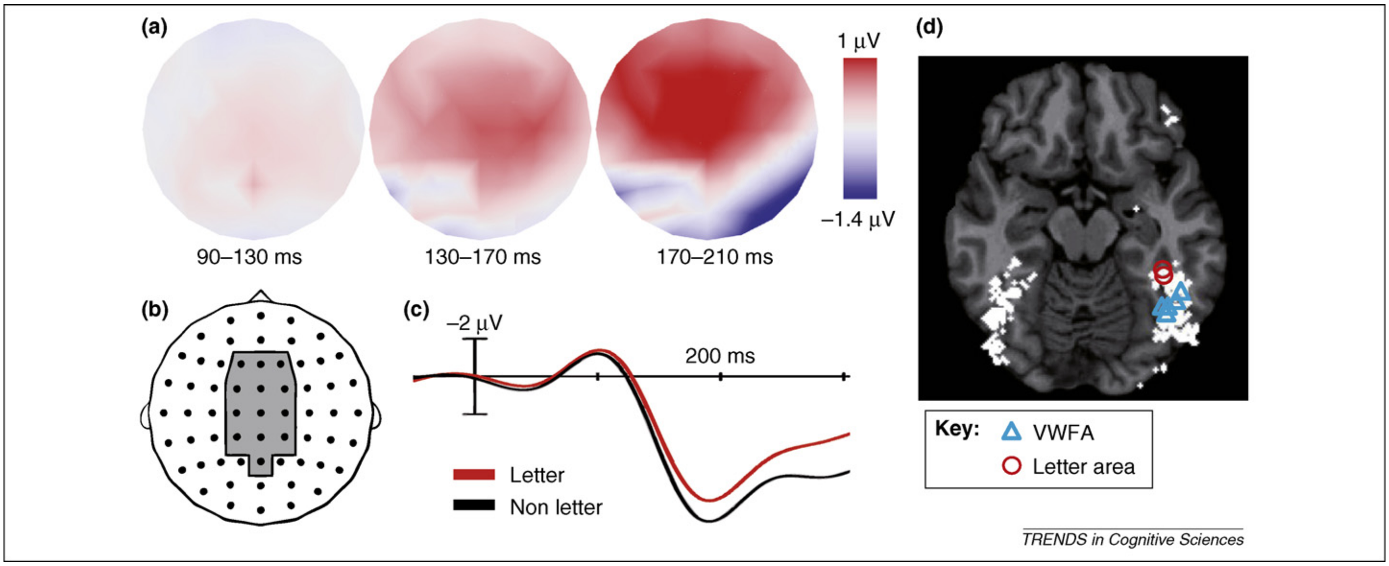

Figure 3. Examples of letter-specific brain activity. ERP results from Rey *et al.* [25] shown as **(a)** scalp maps of voltage differences obtained by subtracting ERP amplitudes generated by letter and pseudo-letter stimuli in three different time windows, and **(c)** grand average waveforms for letters and pseudo-letters over electrodes shown in **(b)**. **(d)** Regions of brain activation measured by fMRI, with the single letter region [28,32] indicated by red circles (left anterior fusiform), and for comparison the visual word form area (VWFA, Ref. [33]) indicated by blue triangles. The region in white is the conjoined activation of objects and letters found by Joseph *et al.* [29] (left hemisphere on the right). Reproduced, with permission, from Ref. [29].

all (because letters could be coded as an unstructured list of features). One influential theory of object recognition posits a key role for structural representations – that is a description of the object in terms of its component parts and their positional relations (such as ‘x is left of y, y is above z’, Ref. [39]). A promising alternative is to code for the position of object parts by using object-centered coordinates, and there is recent neurophysiological evidence in favor of such object-centered coordinate systems [40]. This approach can be easily transposed to the domain of letter identification, using the features derived from empirical investigations of letter perception, augmented with position-in-letter information. Thus, for example, the letter ‘W’ could be coded as: termination upper left, termination upper right, intersection lower left, intersection upper centre, intersection lower right. It is this level of sophistication that might enable feature-based models to better account for patterns in the empirical data (such as letter confusions). Finally, one might also consider that such spatial relations are implicitly coded in hier-

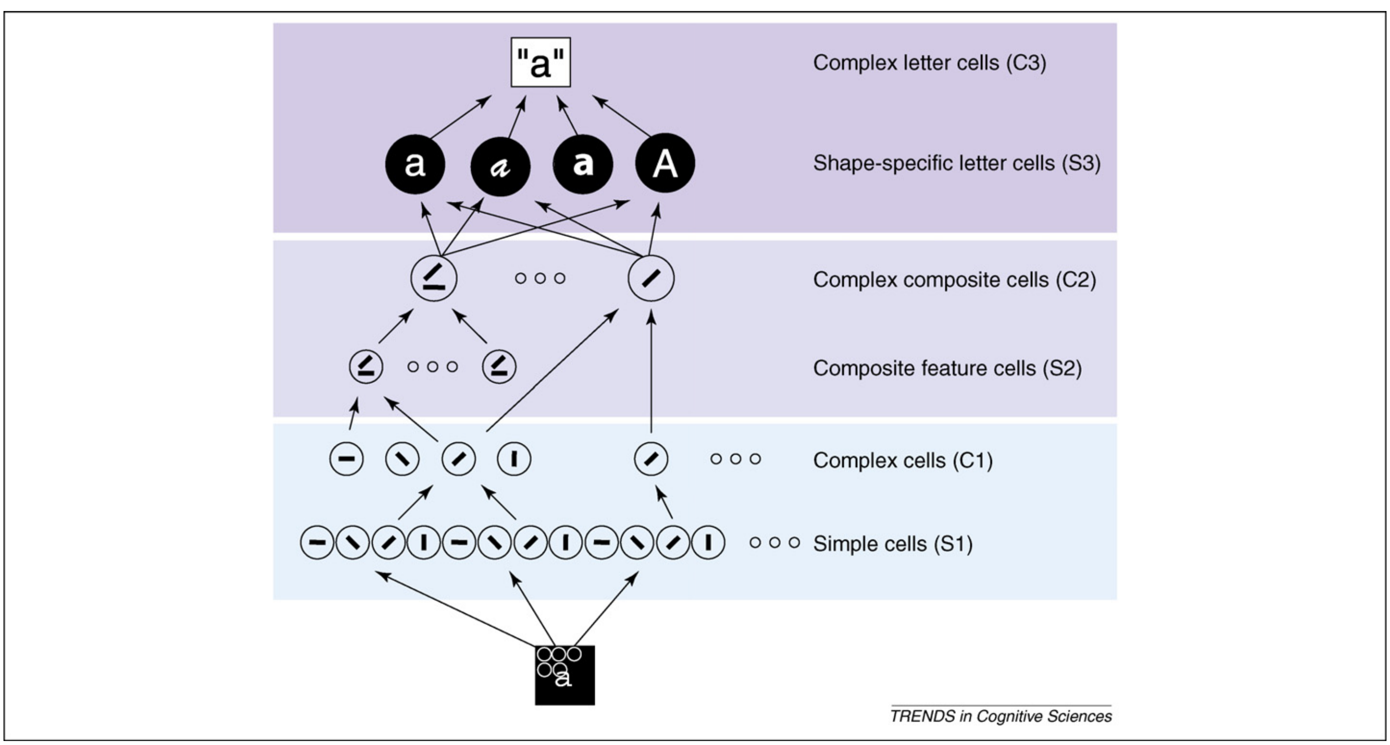

Figure 4. A hierarchical model of letter perception. Shape and location invariance are gradually achieved via a hierarchy of increasingly complex neural processors. Adapted, with permission, from Ref. [38].

archical multi-level networks such as the one depicted in Figure 4.

Another key question is: how the kind of hierarchical structure illustrated in Figure 4 could be learned by a biological system? Could the arbitrary association between lowercase 'a' and uppercase 'A' be learned without supervision (explicit tuition)? Polk and Farah [41] proposed a solution to this problem based on the influence of common contexts on Hebbian learning. The basic idea is that we learn to associate case-specific forms of the same letter by the fact that they commonly occur in the context of case invariant letters (such as in 'map' and 'MAP', in which case-specific 'a' is surrounded by case-invariant 'm' and 'p'). Alternatively, or perhaps in conjunction with this unsupervised learning, children could learn to link 'a' and 'A' on hearing the same letter name associated with each form. Finally, Hinton [42,43] has recently demonstrated that a multi-layered network implementing increasingly complex features can be trained to recognize handwritten digits with high levels of accuracy.

Conclusions and future directions

What is the letter 'a'? This review article has shown that a rich empirical database has accumulated over the years, and there is hope that this will provide the necessary constraints for finding the answer to Hofstadter's [1] question. A major step towards achieving this goal has been made in recent years. Research applying psychophysical techniques has provided convincing evidence that letters are indeed identified via their component features, and there is important new evidence concerning the precise nature of these features. At the same time, research using electrophysiological recordings has started to provide valuable information about the time-course of component processes in letter identification, and functional magnetic resonance imaging (fMRI) studies have begun to isolate the brain regions involved in letter perception (as opposed to letter strings and other visual objects – Figure 3).

The challenge now is to develop a computational model of letter identification that can successfully predict empirical data obtained from the wide spectrum of techniques that have been used to investigate letter perception up to now. Future developments of these models need to be constrained by behavioral, electrophysiological and brain imaging results, and should be articulated with concurrent developments in visual object recognition and printed word perception. It is the application of such multiple constraints that will guarantee success in putting the right parts in the right place at the right time in future computational models of letter identification.

References

1 Hofstadter, D. (1985) *Metamagical Themas: Questing for the Essence of Mind and Pattern,* Basic Books
2 Pelli, D.G. *et al.* (2003) The remarkable inefficiency of word recognition. *Nature* 423, 752–756
3 Grainger, J. (2008) Cracking the orthographic code: an introduction. *Lang. Cogn. Process.* 23, 1–35
4 McClelland, J.L. and Rumelhart, D.E. (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychol. Rev.* 88, 375–407
5 DiCarlo, J.J. and Cox, D.D. (2007) Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–341
6 Selfridge, O.G. (1959) Pandemonium: a paradigm for learning. In *Proceedings of the Symposium on Mechanisation of Thought Processes* (Blake, D.V. and Uttley, A.M., eds), pp. 511–529, H.M. Stationary Office
7 Selfridge, O.G. and Neisser, U. (1960) Pattern recognition by machine. *Sci. Am.* 20, 60–68
8 Hubel, D.H. and Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154
9 Larsen, A. and Bundesen, C. (1996) A template-matching pandemonium recognizes unconstrained written characters with high accuracy. *Mem. Cognit.* 24, 136–143
10 Neisser, U. (1966) *Cognitive Psychology,* Appleton-Century-Crofts
11 Gibson, E.J. (1969) *Principles of Perceptual Learning and Development,* Appleton-Century-Crofts
12 Geyer, L.H. and DeWald, C.G. (1973) Feature lists and confusion matrices. *Percept. Psychophys.* 14, 471–482
13 Keren, G. and Baggen, S. (1981) Recognition models of alphanumeric characters. *Percept. Psychophys.* 29, 452–466
14 Podgorny, P. and Garner, W. (1979) Reaction time as a measure of inter- and intraobject visual similarity: Letters of the alphabet. *Percept. Psychophys.* 26, 37–52
15 Jacobs, A.M. *et al.* (1989) Letter perception in peripheral vision: a temporal discrimination matrix using eye movements. *Percept. Psychophys.* 46, 95–102
16 Courrieu, P. *et al.* (2004) Inverse discrimination time as a perceptual distance for alphabetic characters. *Vis. Cogn.* 11, 901–919
17 Courrieu, P. (2002) Straight monotonic embedding of data sets in Euclidean spaces. *Neural Netw.* 15, 1185–1196
18 Pelli, D.G. *et al.* (2006) Feature detection and letter identification. *Vision Res.* 46, 4646–4674
19 Gosselin, F. and Schyns, P.G. (2001) Bubbles: a technique to reveal the use of information in recognition. *Vision Res.* 41, 2261–2271
20 Fiset, D. *et al.* Features for uppercase and lowercase letter identification. *Psychol. Sci.* (in press)
21 Solomon, J.A. and Pelli, D.G. (1994) The visual filter mediating letter identification. *Nature* 369, 395–397
22 Forster, K.I. and Davis, C. (1984) Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698
23 Petit, J.P. *et al.* (2006) On the time-course of letter perception: a masked priming ERP investigation. *Psychon. Bull. Rev.* 13, 674–681
24 Wong, A.C.N. *et al.* (2005) An early electrophysiological response associated with expertise in letter perception. *Cognitive, Affective, &. Behav. Neurosci.* 5, 306–318
25 Rey, A. *et al.* (2008) Testing computational models of letter perception with item-level ERPs. *Cogn. Neuropsychol.* DOI: 10.1080/09541440802176300
26 Flowers, D.L. *et al.* (2004) Attention to single letters activates left extrastriate cortex. *Neuroimage* 21, 829–839
27 James, K.H. and Gauthier, I. (2006) Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia* 44, 2937–2949
28 James, K.H. *et al.* (2005) Letter processing in the visual system: different activation patterns for single letters and strings. *Cognitive, Affective, &. Behav. Neurosci.* 5, 306–318
29 Joseph, J.E. *et al.* (2006) fMRI correlates of cortical specialization and generalization for letter processing. *Neuroimage* 32, 806–820
30 Longcamp, M. *et al.* (2003) Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage* 19, 1492–1500
31 Pernet, C. *et al.* (2005) Selective response to letter categorization within the left fusiform gyrus. *Neuroimage* 28, 738–744
32 Polk, T.A. *et al.* (2002) Neural specialization for letter recognition. *J. Cogn. Neurosci.* 14, 145–159
33 Cohen, L. *et al.* (2002) Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 125, 1054–1069
34 Dehaene, S. *et al.* (2005) The neural code for written words: a proposal. *Trends Cogn. Sci.* 9, 335–341
35 Grainger, J. *et al.* (2006) Letter position information and printed word perception: the relative-position priming constraint. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 865–884

36 Cadieu, C. *et al.* (2007) A model of V4 shape selectivity and invariance. *J. Neurophysiol.* 98, 1733–1750
37 Fukushima, K. (2003) Neocognitron for handwritten digit recognition. *Neurocomputing* 51, 161–180
38 Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025
39 Biederman, I. (1987) Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147
40 Connor, C.E. *et al.* (2007) Transformation of shape information in the ventral pathway. *Curr. Opin. Neurobiol.* 17, 140–147
41 Polk, T.A. and Farah, M.A. (1997) Simple common contexts explanation for the development of abstract letter identities. *Neural Comput.* 9, 1275–1287
42 Hinton, G.E. (2007) Learning multiple layers of representation. *Trends Cogn. Sci.* 11, 428–434
43 Hinton, G.E. and Salakhutdinov, R.R. (2006) Reducing the dimensionality of data with neural networks. *Science* 313, 504–507
44 Sammon, J. (1969) A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.* 18, 401–409
45 Campbell, F.W. and Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551–566
46 Parish, D.H. and Sperling, G. (1991) Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Res.* 31, 1399–1415
47 Alexander, K.R. *et al.* (1994) Spatial-frequency characteristics of letter identification. *J. Opt. Soc. Am. A* 11, 2375–2382
48 Gold, J. *et al.* (1999) Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Res.* 39, 3537–3560
49 Majaj, N.J. *et al.* (2002) The role of spatial frequency channels in letter identification. *Vision Res.* 42, 1165–1184
50 Sowden, P.T. and Schyns, P.G. (2006) Channel surfing in the visual brain. *Trends Cogn. Sci.* 10, 538–545
51 Bowers, J.S. *et al.* (1998) Orthographic, phonological, and articulatory contributions to masked letter and word priming. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1705–1719
52 Jacobs, A.M. and Grainger, J. (1991) Automatic letter priming in an alphabetic decision task. *Percept. Psychophys.* 49, 43–52
53 Jacobs, A.M. *et al.* (1995) The incremental priming technique: a method for determining within-condition priming effects. *Percept. Psychophys.* 57, 1101–1110
54 Ziegler, J. *et al.* (2000) Visual and phonological codes in letter and word recognition: evidence from incremental priming. *Q. J. Exp. Psychol.* 53A, 671–692
55 Arguin, M. and Bub, D. (1995) Priming and response selection processes in letter classification and identification tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1199–1219
56 Grainger, J. and Ferrand, L. (1996) Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *J. Mem. Lang.* 35, 623–647
57 Kinoshita, S. and Woollams, A. (2002) The masked onset priming effect in naming: computation of phonology or speech planning. *Mem. Cognit.* 30, 237–245
58 Schiller, N.O. (2008) The masked onset priming effect in picture naming. *Cognition* 106, 952–962