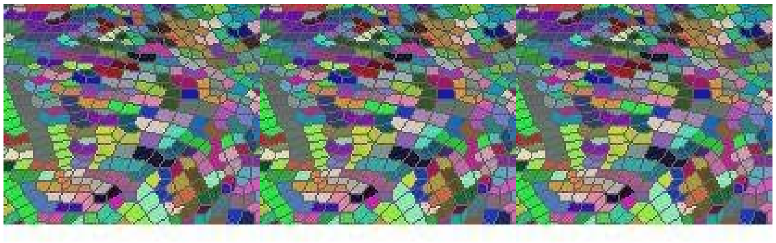

Cognitive Science Online

Vol.1, Issue 1, Winter 2003
http://cogsci-online.ucsd.edu

Letters

Articles

Discussion

Editors: Christopher Lovett, Ayşe Pınar Saygın, Hsin-Hao Yu
Department of Cognitive Science, University of California San Diego
9500 Gilman Drive, La Jolla, CA 92093-0515
cogsci-online@cogsci.ucsd.edu

Information

Cognitive Science Online is an online journal published by the UCSD Cognitive Science Department and seeks to provide a medium for the cognitive science community in which to exchange ideas, theories, information, advice and current research findings. This online publication is a peer-reviewed and highly interdisciplinary academic journal seeking contributions from all disciplines and methodologies investigating the mind, cognition and their manifestation in living, and possibly artificial, systems. For more information about this journal, submissions, back issues, please visit our website at http://cogsci-online.ucsd.edu

Contact Information

Department of Cognitive Science
University of California San Diego
9500 Gilman Drive
La Jolla, CA 92093-0515
cogsci-online@cogsci.ucsd.edu

Editors

Christopher Lovett
Ayşe Pınar Saygın
Hsin-Hao Yu

Advisory Editorial Board

Letter from the Editors

Welcome to the first issue of Cognitive Science Online. In creating this journal we have tried to provide our readers with an insightful and sometimes entertaining glimpse into the world of cognitive science, and whether you are part of the department or interdisciplinary program here at UCSD, a member of the far-reaching, global cognitive science community, or are just curious as to what strange and esoteric research we cognitive scientists have been up to, we're sure you'll be pleased with the results. In creating a journal of this kind we felt it particularly crucial to represent the diversity of ideas floating around in our highly variegated field of cognitive science, as too often the lines that have traditionally partitioned its sub-disciplines begin to form impenetrable barriers around isolated laboratories, and the integrative perspective can begin to fade if left unchecked. As a medium to keep the interdisciplinary spirit of cognitive science alive and flourishing, one of this journal's main aims is to provide a convenient and highly visible forum for communicating information, knowledge and ideas between various researchers and theoreticians who are devoted to studying and ultimately understanding cognition in all its instantiations. Hopefully it will prove to be a valuable resource to those of you wishing to keep abreast of the current research, methodologies, ideas and opinions making up the science of the mind, as well as fostering the incorporation of this information with your own ideas and activities.

As graduate students, we are perhaps in the best position to draw and integrate information from various laboratories, as well as having the freedom to push methodological limits in creating truly novel and creative research designs. In this journal we are particularly interested in publishing scholarly papers written by graduate students in cognitive science or a related field, not only to provide these students with exposure to the outside world, but also to provide examples of the type of cutting-edge work being done in the spirit of a truly unified cognitive science. In addition, this journal provides a forum within which to discuss current opinions and issues, exchange information, facilitate solidarity and cohesion within the department, as well as between various departments within and outside of the UCSD cognitive science community. We would also like to increase the visibility of the field, the people comprised by it, their ideas and their achievements, bringing a greater sense of what cognitive science is and why it is so important to the world at large. Hopefully we will be successful. Enjoy!

Christopher Lovett
Ayşe Pınar Saygın
Hsin-Hao Yu

Letter from *Edwin Hutchins*, Chair of the Cognitive Science Department

It is a pleasure to contribute a note to the inaugural edition of Cognitive Science Online. This peer-reviewed electronic journal edited and produced by the graduate students is a great idea for many reasons. Three aspects of the project are especially appealing.

First, the way the journal project captures and focuses the energy of the department. Cognitive Science is an exciting and rapidly changing field. Our community is a unique group of talented people. One of the best elements of the department chair's job is that it brings one into contact with the activities of the entire department. The chair sees the full scope of the work of the department as reflected in teaching, research proposals, publications, honors, and awards. I can tell you that an enormous amount of groundbreaking work is going on in the department. However, our current institutional practices leave much of that work invisible to the department as a whole. Our second-year project and third-year thesis prospectus presentations are valuable in part because they bring the diverse work of our graduate students into the public eye. Cognitive Science Online provides a forum for communicating to the entire community not just in the month of June and not just about core projects. The vision of the journal is to encourage the exchange of ideas in this very interdisciplinary community. This seems to me to be exactly right. The emphasis on graduate student contributions is also appropriate. Historically, the department's graduate students have supplied, through their research with multiple mentors, the integration that individual faculty members could not accomplish. I welcome Cognitive Science Online as a context in which we can show each other what we do.

Second, a journal run by the students fits perfectly with the wider mission of the department. As a department, we have a stewardship relationship to a body of knowledge. We are responsible for developing the science of the mind through research, passing that knowledge along to another generation of scholars through teaching, applying that knowledge where it can do good in the world, and defending the knowledge against corruption. The interdisciplinary nature of cognitive science makes an in-house journal especially appropriate. To build a department of cognitive science (singular) rather than cognitive sciences (plural), we must continue to foster communication across traditional disciplinary boundaries. The extent to which our department has achieved and maintained integration across a wide range of domains and methods is truly remarkable. This is perhaps the single attribute that most clearly distinguishes this community from other similar efforts. An in-house forum for the open exchange of ideas is thus perfectly suited to our mission. We are truly fortunate to have students with motivation required to make this project go.

Finally, publication is an essential function in our profession. Cognitive Science Online provides graduate students, and others, a convenient early step in a process that is central to our lives as academics. I hope that everyone will contribute. The project also provides the editors a context for learning essential skills in editing and management of a journal.

In retrospect, the journal seems like the obvious thing for the world's best graduate program in cognitive science to do. The founding board of editors, Christopher Lovett, Ayşe Pınar Saygın, and Hsin-Hao Yu, deserve our collective thanks for developing the idea and providing the vision and documentation required to get the project started. Initiative like theirs makes this a department we are happy to come to work in and proud to call our own.

Edwin Hutchins

Chair, UCSD Cognitive Science Department

Cognitive Science Online, Vol.1, pp.1–7, 2003 http://cogsci-online.ucsd.edu

Nonlinear reverse-correlation with synthesized naturalistic noise

Hsin-Hao Yu
Department of Cognitive Science
University of California San Diego
La Jolla, CA 92093
`hhyu@cogsci.ucsd.edu`

Virginia R. de Sa
Department of Cognitive Science
University of California San Diego
La Jolla, CA 92093
`desa@cogsci.ucsd.edu`

Abstract

Reverse-correlation is the most widely used method for mapping receptive fields of early visual neurons. Wiener kernels of the neurons are calculated by cross-correlating the neuronal responses with a Gaussian white noise stimulus. However, Gaussian white noise is an inefficient stimulus for driving higher-level visual neurons. We show that if the stimulus is synthesized by a linear generative model such that its statistics approximate that of natural images, a simple solution for the kernels can be derived.

1 Introduction

Reverse-correlation (also known as white-noise analysis) is a system analysis technique for quantitatively characterizing the behavior of neurons. The mathematical basis of reverse correlation is based on the Volterra/Wiener expansion of functionals: If a neuron is modeled as the functional $y(t) = f(x(t))$, where $x(t)$ is the (one dimensional) stimulus to the neuron, any nonlinear f can be expanded by a series of functionals of increasing complexity, just like real-valued functions can be expanded by the Taylor expansion. The parameters in the terms of the expansion, called *kernels*, can be calculated by cross-correlating the neuronal responses to the stimulus, provided that the stimulus is Gaussian and white (Wiener, 1958; Lee & Schetzen, 1965; Marmarelis & Marmarelis, 1978).

Reverse correlation and its variants are widely used to study the receptive field (RF) structures of the sensory systems. In vision, the circular RF's of LGN neurons and the gabor-like RF's of simple cells in the primary visual cortex are revealed by calculating the first-order (linear) kernels. Neurons with more nonlinearity, such as complex cells, can also be studied by the second-order kernels (Szulborski & Palmer, 1990). However, reverse correlation is rarely applied to extrastriate visual areas, such as V2. One of the many factors that limit reverse correlation to the study of the early visual system is that Gaussian white noise is an inefficient stimulus for driving higher order neurons, since visual features that are known to activate these areas (Gallant et al., 1996; Hegdé & Van Essen, 2000) appear very rarely in Gaussian white noise.

The goal of this paper is to show that if we generate more "interesting" stimuli by training a linear generative model from natural images, solutions to the kernels can be obtained easily. We will proceed by first formulating the Volterra/Wiener series, describe the linear generative model of stimulus synthesis, derive the kernels, and then compare this scheme to other reverse-correlation methods using natural stimuli. The design of physiological experiments using this stimulus is in progress.

2 The Wiener series and reverse correlation

For simplicity, we will only consider systems of two inputs: $y(t) = f(x_1(t), x_2(t))$. Systems of more than two inputs (that is, driven by a stimulus of more than two pixels) follow the same mathematical form.

The Volterra series of f is given by:

$$
\begin{aligned}
y(t) &= f(x_1(t), x_2(t)) \\
&= V_0 + V_1 + V_2 + \dots \\
V_0 &= k_1 + k_2 \\
V_1 &= \int k_1(\tau)x_1(t-\tau)d\tau + \int k_2(\tau)x_2(t-\tau)d\tau \\
V_2 &= \iint k_{11}(\tau_1,\tau_2)x_1(t-\tau_1)x_1(t-\tau_2)d\tau_1\tau_2 \\
&+ \iint k_{22}(\tau_1,\tau_2)x_2(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2 \\
&+ \iint k_{12}(\tau_1,\tau_2)x_1(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2
\end{aligned}
$$

V_0 is the constant term. V_1 describes the linear behavior of the system. The kernels $k_1(\tau)$ and $k_2(\tau)$ are called the *first-order kernels*. V_2 describes the nonlinearity involving interactions between the two inputs. The kernels in V_2 are called the *second-order kernels*. There is a second-order kernel for each pair of inputs. $k_{11}(\tau_1,\tau_2)$ and $k_{22}(\tau_1,\tau_2)$ are called the *self kernels* and $k_{12}(\tau_1,\tau_2)$ is called the *cross kernel*.

In order to solve for the kernels, Wiener re-arranged the Volterra series such that the terms are orthogonal (uncorrelated) to each other, with respect to Gaussian white inputs (Wiener, 1958; Marmarelis & Naka, 1974; Marmarelis & Marmarelis, 1978).

$$
\begin{aligned}
y(t) &= f(x_1(t), x_2(t)) \\
&= G_0 + G_1 + G_2 + \dots \\
G_0 &= h_1 + h_2 \\
G_1 &= \int h_1(\tau)x_1(t-\tau)d\tau + \int h_2(\tau)x_2(t-\tau)d\tau \\
G_2 &= \iint h_{11}(\tau_1,\tau_2)x_1(t-\tau_1)x_1(t-\tau_2)d\tau_1\tau_2 - P\int h_{11}(\tau,\tau)d\tau \\
&+ \iint h_{22}(\tau_1,\tau_2)x_2(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2 - P\int h_{22}(\tau,\tau)d\tau \\
&+ \iint h_{12}(\tau_1,\tau_2)x_1(t-\tau_1)x_2(t-\tau_2)d\tau_1\tau_2
\end{aligned}
$$

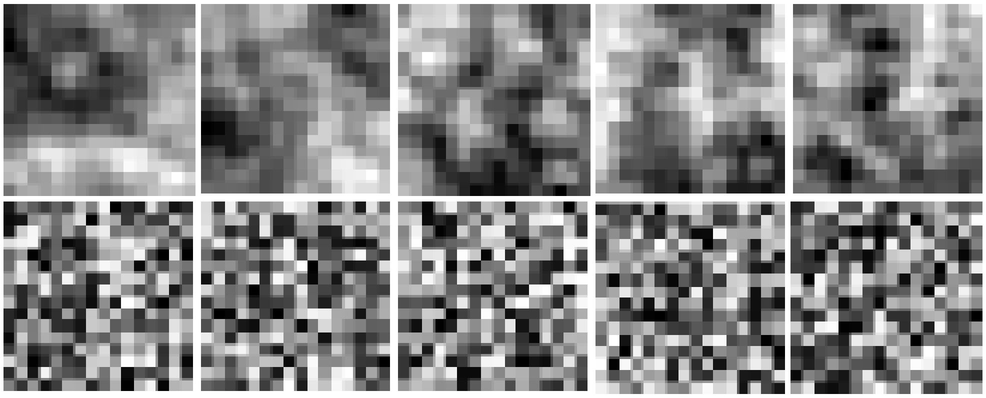

Figure 1: The stimuli (vector x, upper row) are synthesized by linearly transforming a white noise cause (vector s, lower row) via a linear generative model: $x = A\ s$. Matrix A is learned from samples of natural images.

where $x_1(t)$ and $x_2(t)$ are independent Gaussian white inputs, with equal power (or variance) P. The kernels are called the *Wiener kernels*.

Lee and Schetzen (Lee & Schetzen, 1965) showed that the Wiener kernels can be calculated by cross-correlating the neuronal response $y(t)$ with the inputs. For example, the first-order kernel $h_1(\tau)$ can be calculated from $\langle y(t)x_1(t-\tau)\rangle$, self-kernel $h_{11}(\tau_1, \tau_2)$ from $\langle y(t)x_1(t-\tau_1)x_1(t-\tau_2)\rangle$, and the cross-kernel $h_{12}(\tau_1, \tau_2)$ from $\langle y(t)t_1(y-\tau_1)x_2(t-\tau_2)\rangle$[1]. See (Marmarelis & Naka, 1974; Marmarelis & Marmarelis, 1978) for details.

3 Synthesis of naturalistic noise and kernel calculation

3.1 The synthesis model

Instead of using Gaussian white noise for reverse correlation, we can linearly transform white noise such that the the statistics of the transformed images approximate those of natural images. This should produce a better stimulus for higher-order visual neurons since it contains more features found in nature.

More specifically, let the stimulus $x(t) = (x_1(t) \ldots x_n(t))^T$ be synthesized by:

$$x(t) = A\ s(t)$$

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{bmatrix}$$

where $s(t) = (s_1(t) \ldots s_n(t))^T$ is white. The vector $s(t)$ is called the *cause* of the stimulus $x(t)$. The constant matrix A can be learned from patches of natural images by various algorithms, for example, Infomax Independent Component Analysis (Infomax ICA) (Bell & Sejnowski, 1995, 1996). In this case, the causes $s_1(t) \ldots s_n(t)$ are required to be Laplacian distributed.

---

[1] $\langle\ \rangle$ denotes expectation over t

Examples of the synthesized stimuli are illustrated in Figure 1. Visual features that occur very rarely in white noise, such as localized edges, corners, curves, and sometimes closed contours, are much more common after the A transformation.

Using linear generative models to synthesize stimulis for physiological experiments was also suggested in (Olshausen, 2001).

3.2 Kernel calculation

To calculate the kernels, one can follow Wiener and orthogonalize the Volterra series with respect to the distribution of the new stimulus, instead of Gaussian white noise. Here we provide a much simpler solution, using a trick that is similar to the treatment of non-white inputs in (Lee & Schetzen, 1965).

The derivation is illustrated in Figure 2. Instead of directly solving for the kernels of system f, we consider system f', which is formed by combining system f with the linear generative model: $f' = f \circ A$ (Figure 1b). The kernels of system f' can be calculated by the standard cross-correlation method, because its input $s(t)$ is white[2]. After f' is identified, we consider a new system f'', formed by combining f' with the inverse of the generative model: $f'' = f' \circ A^{-1}$ (Figure 1c). The kernels of system f'' can be easily obtained by plugging $s(t) = A^{-1}x(t)$ into the kernels of f', and expressing the kernels as functions of $x(t)$ instead of $s(t)$. But since $f'' = f' \circ A^{-1} = f \circ A \circ A^{-1} = f$, system f'' is equivalent to f. We therefore calculate kernels of f by transforming the kernels of f'.

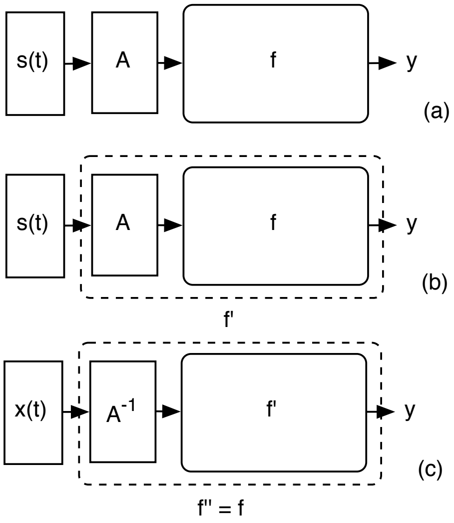

Figure 2: The derivation of formulas for kernels. (a) In order to calculate the kernels of system f, we form the system f' as in (b). Kernels of system f' can be obtained by the standard cross-correlation method because the input s is white. After the kernels of f' are identified, we construct system f'' as in (c). The kernels of system f'' can be obtained by transforming the kernels of f'. But since f'' is equivalent to f, this yields the kernels that we wanted in the first place.

---

[2] Note that $s(t)$ is Laplacian distributed, instead of Gaussian distributed. Kernels higher than the first order need to be calculated according to (Klein & Yasui, 1979; Klein, 1987).

Let $\phi_1(\tau)\ldots\phi_n(\tau)$ be the first-order kernels of f', obtained by cross-correlating system response with white noise $s(t)$. The first-order kernels of the original system f, $h_1(\tau)$ $\ldots h_2(\tau)$, are simply

$$\begin{bmatrix} h_1(\tau) \\ \vdots \\ h_n(\tau) \end{bmatrix} = A^{-t} \begin{bmatrix} \phi_1(\tau) \\ \vdots \\ \phi_n(\tau) \end{bmatrix}$$

The second-order kernels of system f,

$$h_{ij}(\tau_1,\tau_2), \qquad i,j = 1\ldots n, \quad h_{ij}(\tau_1,\tau_2) = h_{ji}(\tau_1,\tau_2)$$

can be calculated from $\phi_{ij}(\tau_1,\tau_2)$, kernels of system f', by the following equation:

$$\begin{bmatrix} c_{11}h_{11}(\tau_1,\tau_2) & \ldots & c_{1n}h_{1n}(\tau_1,\tau_2) \\ \vdots & & \vdots \\ c_{n1}h_{n1}(\tau_1,\tau_2) & \ldots & c_{nn}h_{nn}(\tau_1,\tau_2) \end{bmatrix} = A^{-t} \begin{bmatrix} c_{11}\phi_{11}(\tau_1,\tau_2) & \ldots & c_{1n}\phi_{1n}(\tau_1,\tau_2) \\ \vdots & & \vdots \\ c_{n1}\phi_{n1}(\tau_1,\tau_2) & \ldots & c_{nn}\phi_{nn}(\tau_1,\tau_2) \end{bmatrix} A^{-1}$$

where $c_{ij} = 1$ if $i = j$, and $c_{ij} = \frac{1}{2}$ if $i \neq j$. Higher order kernels can also be derived.

3.3 Notes on implementation

First, since training ICA on natural images usually produces a matrix whose row vectors resemble gabor functions(Bell & Sejnowski, 1996), we can construct matrix A directly as rows of gabor patches. This is similar to the synthesis model in (Field, 1994), and has the advantage of not being biased by the particular set of images used for training. From this point of view, the synthesized stimulus is a random mixture of edges.

Second, the synthesis method described so far generates each frame independently. If ICA is trained on movies, we can synthesize image sequences with realistic motion (van Hateren & Ruderman, 1998; Olshausen, 2001). The frames in the sequences are correlated, but described by independent coefficients. The spatiotemporal kernels of neurons with respect to synthesized movies can also be derived by the same procedure.

4 Comparison to related work

To overcome the limitations of using Gaussian white noise for reverse correlation, researchers have recently started to use natural stimuli (Theunissen et al. (2000) in the auditory domain, and Ringach et al. (2002) in vision). They found RF features that were not revealed by white noise. The analysis strategy of these methods is to model receptive fields as linears filter with zero memory, and solve for the mean square error solution by regression (DiCarlo et al., 1998) or the recursive least square algorithm (Ringach et al., 2002). This involves estimating and inverting the spatial autocorrelation matrix of the stimulus.

The advantages of our approach using synthesized stimulus are:

- Dealing with natural images usually requires a large amount of memory and storage. In our method, unlimited number of frames can be generated on demand, once the synthesis matrix A is learned. Kernel calculation is also easier.
- In our method, all the statistics about the stimulus is contained in the matrix A, allowing us to formulate reverse correlation in terms of the Wiener series and derive formulas for higher order kernels, which can be important for studying the

non-linear behavior of neurons (Szulborski & Palmer, 1990). Higher order kernels for natural images are much more difficult to derive, due to their complicated (and largely unknown) statistical structure. The existing regression methods for natural image reverse correlation assume linearity and do not allow the calculation of higher order kernels.

- The synthesis model is motivated by the redundancy reduction theory of the early visual code (Barlow, 1961; Field, 1994; Olshausen & Field, 1996; Bell & Sejnowski, 1996), which states that the goal of early visual code is to transform the retinal representations of natural images to an independent, sparse code. If this theory is to be taken literally, the computation of the early visual system is essentially A^{-1}, and the synthesized stimulus $x(t)$ is represented as $s(t)$ by the first-order system (the primary visual cortex). Under this assumption, second-order neurons are receiving (Laplacian distributed) white noise stimuli. The kernels ϕ's can therefore be interpreted as the kernels of higher-order systems with respect to cortical codes, instead of retinal codes. This can be useful for interpreting the non-linear behavior of neurons(Hyvärinen & Hoyer, 2000; Hoyer & Hyvärinen, 2002)

5 Discussion

We have shown how to easily derive kernels for a specific form of naturalistic noise. As this stimulus has more of the features of natural stimulation, it should more strongly activate visual neurons and allow us to more efficiently explore receptive fields.

We are currently designing physiological experiments to test this procedure on simple and complex cells in the primary visual cortex of squirrels. Specifically,

- We will calculate first-order kernels using white noise, synthesized naturalistic noise, and natural images, and compare the quality of the receptive field maps.
- Examine if second-order kernels can be reliably calculated, and see if they help to predict the behavior of neurons.
- Analyze the relationship between h's (kernels with respect to retinal code) and ϕ's (kernels with respect to cortical code, under the whitening hypothesis), and examine if the coding hypothesis helps us understand the structure of the complex cells.

References

Barlow, H. B. (1961). The coding of sensory messages. In *Current problems in animal behavior.* Cambridge: Cambridge University Press.

Bell, A. J., & Sejnowski, T. J. (1995). An information maximisation approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159.

Bell, A. J., & Sejnowski, T. J. (1996). Edges are the “independent components” of natural scenes. *Advances in Neural Information Processing Systems*, *9*.

DiCarlo, J. J., Johnson, K. O., & Hsiao, S. S. (1998). Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. *Journal of Neuroscience*, *18*, 2626–2645.

Field, D. J. (1994). What is the goal of sensory coding. *Neural Computation*, *6*(4), 559–601.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. E., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, *76*, 2718–2739.

Hegdé, J., & Van Essen, D. (2000). Selectivity for complex shapes in primate visual area v2. *The Journal of Neuroscience*, *20*(RGB 1 of 8).

Hoyer, P. O., & Hyvärinen, A. (2002). A multi-layer sparse coding network learns contour coding from natural images. *Vision Research*, *42*(12), 1593-1605.

Hyvärinen, A., & Hoyer, P. O. (2000). Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces. *Neural Computation*, *12*(7), 1705-1720.

Klein, S. A. (1987). Relationship between kernels measured with different stimuli. In P. Z. Marmarelis (Ed.), *Advanced methods of physiological system modeling* (Vol. I, pp. 278–288). Plenum Press.

Klein, S. A., & Yasui, S. (1979). Nonlinear systems analysis with non-gaussian white stimuli: general basis functionals and kernels. *IEEE Transactions on Information Theory*, *IT-25*(4).

Lee, Y. W., & Schetzen, M. (1965). Measurement of the wiener kernels of a non-linear system by cross-correlation. *International Journal of Control*(2), 234–254.

Marmarelis, P. Z., & Marmarelis, V. Z. (1978). *Analysis of physiological systems: the white noise approach.* Plenum Press.

Marmarelis, P. Z., & Naka, K. (1974). Identification of multi-input biological systems. *IEEE Transactions on Biomedical Engineering*, *BME–21*(2).

Olshausen, B. A. (2001). Sparse coding of time-varying natural images. In *Society for neuroscience abstracts* (Vol. 27). Society for Neuroscience.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.

Ringach, D. L., Hawken, M. J., & Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision*, *2*(1), 12-24.

Szulborski, R. G., & Palmer, L. A. (1990). The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Research*, *30*(2), 249–254.

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of Neuroscience*, *20*(6), 2315-2331.

van Hateren, J. H., & Ruderman, D. L. (1998). Independent component analysis of image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc.R.Soc.Lond. B*(265), 2315–2320.

Wiener, N. (1958). *Nonlinear problems in random theory.* The M.I.T. Press.

Cognitive Science Online, Vol.1, pp.8-21, 2003 http://cogsci-online.ucsd.edu

Tapping into the continuum of linguistic performance: Implications for the assessment of deficits in individuals with aphasia

Suzanne Moineau

Joint Doctoral Program in Language and Communicative Disorders
University of California San Diego and
San Diego State University
`smoineau@crl.ucsd.edu`

Abstract

The fundamental goal of every speech and language clinician is to provide services that will enhance the functional communicative abilities of the patients they treat. The cornerstone of developing a successful intervention program is careful patient assessment. Historically, clinicians have relied on traditional standardized language and neuropsychological assessment tools to determine performance baselines from which to plan the treatment course. Although informative in many ways, the batteries that are used can also be limiting. Most often they force clinicians and researchers into forming categorical diagnostic groups, which may result in the loss of critical information essential for the planning of therapeutic interventions. The purpose of the current paper is to review some empirical evidence that suggests we should strongly consider redefining classic syndromes, redesigning standard assessment tools, and utilizing new technologies to map out the symptom space in individuals with brain injury.

Introduction

In 1861, Paul Broca published an historically influential paper that aimed to systematically map behavioral symptoms to particular brain regions. Specifically, Broca claimed that the third convolution of the left frontal lobe was the seat of articulate speech, and that damage to this area would result in a defect in the motor realization of language (Goodglass, 1993). Soon after, Broca's aphasia became widely accepted as an impairment in the production of language resulting in a patient having non-fluent speech output but intact auditory comprehension. The patient thus exhibits an apparent ability to fully understand directives, questions and even simple conversation despite speech production that is telegraphic, primarily consisting of content words, and noticeably labored. Although the last 40 years have brought about minor revisions in this classic definition (i.e., auditory comprehension deficits can be seen, but only with complex syntax and grammar (Grodzinsky, 1995, 2000)), the core of the classification remains unchanged. Likewise, the cognitive and behavioral deficits associated with left temporal lobe damage, as outlined by Carl Wernicke, have undergone very little, if any, revision since 1874. Damage to this area of the brain typically results in deficits in comprehension of spoken language, however, as non-speech sensory images are purportedly intact, the Wernicke's aphasic demonstrates fluent, albeit paraphasic speech output. A patient with this classic profile typically suffers from an inability to comprehend even the simplest of linguistic stimuli (e.g., 'Is your name Bob?' or 'Touch your nose'). Also, despite having the natural flow and contours of normal speech production, the Wernicke's aphasic frequently produces non-words or misuses words in a given context. Though current diagnostic categories are grossly sufficient in describing the prototypical syndrome characteristics, a vast number of individuals with aphasia do not 'fit' these prototypes. This early observation fueled numerous debates, which continue to this day, about the nature of brain organization for language production and comprehension.

Despite strong evidence from the start against the theory that the brain comprises discontinuous sensorimotor centers and connections, this was the dominant view up until the latter half of the twentieth century (summarized in Kean, 1985)[1]. This view was then replaced (at least in some scientific circles) by another theory that had equally strong ties to the claim that brain areas are discontinuous by nature. However, the new account shifted from holding that these centers are separated along distinct sensorimotor lines to the notion that they are differentiated along content lines (i.e. grammar vs. lexicon, Caramazza, Berndt, Basili & Koller, 1981). Either way, the predominant view over the last 150 years has been one of discontinuous centers and mental organs which are localizable and domain specific. This view also holds that the deficits resulting from damage to such centers are easily distinguishable from one another and are able to be classified into distinct categories. It is out of these theories about brain organization for language, the autonomy of linguistic structures from one another and from all other cognitive functions, and specific competence and performance patterns following neurologic insult that aphasia subtypes and classifications were born. It is also with these same theories in mind that we design diagnostic tools with which to assess linguistic function and build our treatment plans.

Today, although most clinicians in the trenches would readily agree that the symptoms observed in aphasia are more accurately defined as continuous rather than discrete, as evidenced by the lack of 'pure' cases and the high proportion of unclassifiable patients, our standard assessment tools do not adequately reflect our understanding of this complex terrain of linguistic deficits. Instead, we continue to utilize and rely on tools that give composite scores and provide classifications as a way to conveniently define the linguistic behaviors that we are observing. By fitting all outliers into one or another syndrome type we fail to refine our diagnostic tools and criteria so that they are sensitive to and reflect the actual nature of language processing, which through strong evidence appears to be continuous and dynamic.

We do, however, have the means by which to more accurately measure, and thus define, the unique symptom space that every individual with aphasia occupies. These

---

[1] A more 'holistic' or distributed alternative to the centers-and-connections view has been offered throughout the modern history of aphasia -- from Freud, Hughlings Jackson, Pierre Marie, Arnold Pick, Kurt Goldstein, Henry Head, and in more recent times, by researchers like Hermann Kolk, Claus Heeschen and (as cited) Sheila Blumstein. It is just the case, however, that the *predominant* view, at least in Anglo-American circles, has been the modular one.

methods include investigating the nature of real-time lexical processing via on-line priming studies, processing in ‘noisy’ environments, and implicit processing as measured by event-related potentials (ERPs) and eye movements. These techniques can assist us in more precisely outlining profiles of processing strengths and weaknesses which can help us in designing effective intervention plans, and they are also more ecologically valid than traditional standardized tests. Though most of these techniques are still in the beginning stages of implementation with clinical populations, and have found their primary utilization in the research community, their systematic clinical application is essential if we are to do the greatest service to our patients. They enable us to uncover vulnerabilities in the system that may have otherwise gone undetected and as well, capture preserved implicit processing in individuals with severe deficits who, due to linguistic or physical limitations, may not be able to comprehend and/or perform even the simplest of explicit tasks. For example, both lexical priming studies and experiments investigating the effects of acoustic degradation on speech processing have led to the discovery that lexical activation and single word comprehension are not fully intact in Broca’s aphasia, as once thought. On the flip side, recent work in neuroimaging has provoked some researchers to conclude that both Wernicke’s and global aphasics, who may be reported to have no explicit comprehension of linguistic stimuli, may have preserved implicit semantic priming. Below is a summary of some of the work being done in these fields to better define the processing landscape of individuals with aphasia.

Lexical Processing via Priming Studies

A classic view of lexical-semantic impairments in individuals with aphasia would predict that Broca’s aphasics are largely unaffected in this domain, while Wernicke’s aphasics demonstrate severe deficits in lexical-semantic processing (Zurif, Caramazza, Myerson, and Galvin, 1974; Goodglass & Baker, 1976). In other words, Broca’s aphasics have been considered to have an intact ability to activate and integrate lexical items, as evidenced by ‘spared’ comprehension of content words, while Wernicke’s aphasics have been claimed to have deficiencies in this domain as noted by their failure to comprehend even simple, common lexical items. These views arose, however, mainly from measurements that required subjects to perform explicit judgment tasks, such as those presented in standardized tests. It was not until the 1980’s that the accepted overall picture of intact lexical processing in Broca’s aphasics began to be challenged. Initially, neuropsychological studies used a priming paradigm to investigate the integrity of the lexical-semantic system in individuals with aphasia, which provided evidence that neurologically intact individuals are faster and more accurate to perform a lexical decision task (i.e. to decide if a presented utterance is a real word or a non-word) following a primed word that is related to the target than to an unrelated word (Milberg and Blumstein, 1981; Blumstein, Milberg & Schreir, 1982; Milberg, Blumstein & Dworetzky, 1987, 1988). Contrary to what might be predicted of Broca’s aphasics, however, the majority of the evidence revealed abnormal priming patterns in these patients. A few examples of the nature of the differences follows:

In their work, Milberg and colleagues (1988) found a group of Broca’s aphasics to have reduced activation of lexical targets following the presentation of a phonologically altered prime (i.e., “gat” - “dog” vs. “cat” - “dog”). In other words, when the prime was a good production (having accurate place, manner and voiceonset time) Broca's aphasics demonstrated normal priming; however, if the prime was a poor exemplar (altered along one of these dimensions) they showed reduced priming as compared to normal control subjects. More recently, Utman, Blumstein and Sullivan (2001) found a similar effect in a new group of individuals with Broca's aphasia. Like the results obtained by Milberg, *et al.*, these Broca's aphasics demonstrated a larger and longer-lasting reduction in semantic priming in response to word-initial acoustic manipulations when compared to normal controls.

Swinney, Zurif & Nicol (1989) demonstrated that Broca's aphasics fail to show exhaustive access of secondary meanings of ambiguous words. Instead, the subjects they tested only accessed the most frequently occurring meaning, independent of contextual bias: a finding not seen in young adults, the elderly or fluent aphasics, who automatically primed all meanings of ambiguous words. An earlier study conducted with normal college students revealed, however, faster activation times for the primary meaning of an ambiguous word than for its secondary meanings, independent of context (Simpson, 1984). Given this finding, Swinney, *et al.*, speculated that Broca's aphasics may just have a slower-than-normal time course of meaning activation, with a corresponding failure to activate meanings beyond the most frequent one.

This *speed* of activation account of the deficits seen in lexical activation in Broca's aphasics has found further support through work conducted by Penny Prather and colleagues (Prather, Zurif, Stern & Rosen, 1992; Prather, 1994; Prather, Zurif, Love & Brownell, 1997). Prather, *et al.* demonstrated that Broca's aphasics do in fact show automatic priming, with normal decay, however the time course is protracted. In contrast to normal elderly controls who prime at relatively short interstimulus intervals (ISIs) beginning at 500 ms, Broca's aphasics show reliable automatic priming only at ISIs of 1500ms.

From the evidence accumulated from on-line priming studies it is becoming clear that Broca's aphasics do show deficits in lexical access as compared to normal control subjects, a finding that is rarely apparent in single word comprehension tasks where reduced speeds and incomplete activation may not be very sensitive in revealing underlying deficits. These studies did more than just uncover deficits in the lexical-semantic processing of Broca's aphasics, however. They uncovered some priming effects of lexical items in Wernicke's aphasics that more closely resembled normal controls than the priming seen in Broca's aphasics. In particular, Blumstein et al (1982) and Milberg (1987) found that unlike Broca's aphasics, Wernicke's aphasics prime poor exemplars, while Prather *et al.* (1997) found that their subject with Wernicke's aphasia showed a normal, rapid initial activation. While all studies also revealed differences between the priming of Wernicke's aphasics and normal controls (i.e., Wernicke's tend to hyper-prime distant exemplars and have abnormally long ranges of priming), classic taxonomies would suggest that on a lexical task, Broca's aphasics would show more normal patterns of processing than Wernicke's aphasics. This however does not appear to be the case.

Processing In 'Noisy' Environments

When first initiated, much of the work being done studying language deficits focused on the differential effects of acoustic degradation (distortions of the speech signal) on spoken language processing in aging and hearing impaired populations, as compared to normal, healthy young adults (Dirks, Morgan & Dubno, 1982; Helfer & Wilber, 1990; Gordon-Salant & Fitzgibbons, 1993, 1995, 1997). Soon after, however, this experimental paradigm began to be used to test a very different type of theory from the ones initially investigated.

As evidence began to mount against a theory of autonomous, domain-specific linguistic modules that can be independently impaired in individuals following stroke, accounts predicting reductions in processing resources and verbal working memory as the cause of processing breakdowns in aphasia began to be proposed (Just & Carpenter, 1992; Caplan & Waters, 1999).[2] Theories such as these suggested that any individual facing a reduction in resources may be susceptible to breakdowns in processing. In 1991, Kerry Kilborn set out to test this very hypothesis. He predicted that subjecting normal, healthy individuals to noise, and thereby reducing their general processing resources, could create isolated deficits in specific grammatical features. The deficits he induced in normal controls mirrored the performance of a group of German Broca's aphasics who were tested in a similar sentence interpretation task by Bates, Friederici and Wulfeck (1987). In short, the acoustic manipulation caused a selective breakdown in the processing of grammatical morphology, and an increased reliance on word order.

Since the time of this initial research by Kilborn, a flurry of similar studies have been conducted to investigate the vulnerability of different grammatical, syntactic and lexical structures to 'noise' (Miyake, Carpenter, & Just, 1994; Blackwell & Bates, 1995; Utman & Bates, 1998; Dick, Bates, Wulfeck, Utman, Dronkers, Gernsbacher, 2001). In these experiments, subjects were required to process linguistic stimuli (presented either via auditory or visual modality) under any one or more of the following experimental conditions: digit load manipulations, speeded presentation, or filtering of the speech signal. The findings of such studies converge along similar lines: the vulnerability seen in individuals with aphasia can be reproduced in neurologically intact individuals under a range of compromising processing conditions. These studies have, however, tested the effects of acoustic degradation on more complex structures than just the single lexical item, and have focused on deficits seen in normal controls. However, this experimental paradigm appeared to be ideal for investigating the vulnerability of single word comprehension in Broca's aphasics, who have historically been considered to be intact in this domain (Berndt & Caramazza, 1999). Though evidence from the lexical priming literature clearly reflects deficits in the lexical access system of Broca's aphasics, priming is a measure that may quite possibly be more indirect than a simple, direct single word lexical comprehension task.

In 2001, an experiment was conducted in our laboratory to assess the effects of acoustic degradation on single word comprehension in a group of elderly controls, right-hemisphere-damaged (RHD) individuals and individuals with aphasia (Moineau & Bates, 2001). The results indicated the following: when presented with unaltered speech, Broca's aphasics are as accurate as neurologically intact, age-matched adults at identifying whether or not a single spoken word correctly matches a visually presented picture. That is, under optimal listening conditions, such as those present in a standardized testing room, Broca's aphasics appear to have intact single word

---

[2] While Caplan & Waters do propose that there are different resources for syntax and all other language-related resources, they support the view that the deficits seen in individuals with aphasia are due to processing limitations rather than damage to a specific linguistic structure.

comprehension skills. Furthermore, as predicted by the traditional classification models, the group of Wernicke's aphasics tested in this experiment was the only population to demonstrate a significant comprehension deficit in the unaltered condition. When the stimuli were altered, however, the group of Broca's aphasics demonstrated significant decreases in accuracy as compared to elderly controls and RHD individuals. Their performance did not, however, differ significantly from that of anomic aphasics or Wernicke's aphasics. This pattern revealed a gradient of diminished performance that was based on group severity: Wernicke < Broca < Anomic < RHD < Elderly.

In summary, this experiment demonstrated that the range of deficits seen in aphasic patients when performing even a simple, single word comprehension task is continuous and fragile. Given the fact that we are most often processing speech with some type of environmental noise in the background, the conditions presented in this study appear to be more consistent with natural processing settings than the conditions under which we engage in standardized testing. This study reveals a gradient of performance that can be easily uncovered when we subject individuals to exogenous stressors that mimic common environmental conditions, and may also serve to simulate endogenous alterations in the processing climate following stroke.

To summarize thus far, evidence has accumulated to suggest that discrete classifications of symptoms as 'present' or 'absent' should be abandoned in favor of diagnostic criteria that respect the true nature of language impairments as they vary along a continuum from mild to severe. Patients who do not "have" the impairment in question under traditional classification schemes will show vulnerabilities when their processing abilities are pushed to the limit. Indeed, this gradient approach to processing deficits extends to normal, neurologically intact adults, whose performance resembles that of brain-injured patients when the processing climate is altered through exogenous stressors that mimic conditions of brain damage. Let us turn now to the other side of the coin: studies demonstrating spared (residual) abilities in patients who (by classic criteria) are believed to have lost the ability in question.

Implicit Processing: Event Related Potentials

To date, there is a substantial body of research that has utilized event related potentials (ERPs) to investigate the neural bases of cognition and language in both healthy, normal adults and those with neurologic damage (Neville, 1980; Kutas & Van Petten, 1994; Rugg, 1995; Hagoort, Brown & Swaab, 1996; Revonsuo & Laine, 1996; Swaab, Brown & Hagoort, 1997; Friederici, Hahne, von Cramon, 1998; Swaab, 1998; Swaab, Brown & Hagoort, 1998; Connolly, Mate-Kole & Joyce, 1999; Connolly, Major, Allen, & D'Arcy, 1999; Friederici & Jacobsen, 1999; Friederici, vonCramon & Kotz, 1999; Connolly & D'Arcy, 2000). Of particular interest in the investigation of language processing is the N400 effect. This effect was first discovered in 1980 by Kutas and Hillyard. In their study, Kutas and Hillyard found that a negative deflection in the ERP waveform was evident between 380-440 milliseconds following the presentation of a semantically anomalous word in a sentence context. The difference between the amplitude of the N400 to the semantically congruous versus the incongruous word is considered the N400 effect. This methodology is particularly useful in the assessment of individuals with aphasia as it does not require overt responses, and is capable of measuring brain activity during the normal continuous stream of speech, without interruption.

One of the first studies to investigate ERP effects in individuals with aphasia was carried out by Revonsuo & Laine (1996). The subject in their case study was noted to have a lesion that involved the entire region of Broca's area, part of Wernicke's area, and included the insula and underlying white matter. The authors began testing the subject 2.5 months post onset. He was reported to be globally aphasic, as measured by the standardized Finnish version of the Boston Diagnostic Aphasia Examination (Laine, Goodglass, Niemi *et al.*, 1993) at one week post onset. Prior to examining the N400 effect, a sentence categorization task revealed chance level performance for explicit comprehension. The authors presented the subject with 400 test items: 200 congruent and 200 anomalous sentences. Results indicated a pattern of performance consistent with normal controls. The subject in this study demonstrated a significantly greater negativity, occurring at approximately 400 ms after the presentation of the target word, for anomalous words as compared to congruent words.

In another study of implicit processing in a global aphasic, Connolly, *et al.* (1999) found similar findings to that of Revonsuo & Laine. The subject in their study was reported to be severely compromised by his injuries and was not indicated for rehabilitation. Formal, traditional assessment could not be conducted as the patient was not capable of performing the explicit tasks required. The authors tested the N400 effect to 320 sentences: 160 presented in the visual modality and 160 presented in the auditory modality. They found that their subject exhibited brain response patterns indicative of intact implicit processing. Comparisons to a control group revealed that the grand average for the aphasic subject did not differ significantly from the group of control subjects. This finding led to a reinstatement of individualized rehabilitative intervention, with a successful outcome.

In 1996, Hagoort and colleagues (Hagoort, Brown & Swaab, 1996) conducted a larger scale study, looking at the N400 effect in a group of 20 aphasic patients. Their goal was to see if there was a difference in the N400 effect : 1) based on aphasia syndrome (Broca's vs. Wernicke's); 2) based on severity (as measured by scores on comprehension subtests independent of syndrome classification); and 3) between aphasics and normals. The results indicated the following: 1) There was no significant difference in the N400 effect based on aphasia syndrome. The Wernicke's aphasics did show a larger reduction in the size of the N400 effect, as compared to the Broca's aphasics, however, this effect was not significant; 2) There was a significant difference noted in the N400 effect based on severity of symptoms. Individuals that were rated as low comprehenders, based on comprehension scores obtained via the Aachen Aphasia Test, showed significant reductions in the N400 effect as compared to high comprehenders and normal controls. The latter two groups did not differ materially from one another; and 3) Despite reductions in the overall N400 effect for the aphasic group as whole, they did not differ significantly from the control group.

In a subsequent study, Swaab and colleagues again looked at the differences in the N400 effect between brain-lesioned and neurologically intact individuals (Swaab, Brown & Hagoort, 1997). In this study, however, the authors compared their aphasic subjects solely based on comprehension severity (i.e., mild vs. moderate-severe impairment) and not based on classification type as in the previous study. In addition, they added a group of non-aphasic patients with right hemisphere damage. Again, results indicated that aphasics with only a mild comprehension deficit show N400 effects that are not significantly different than normal, elderly controls. The non-aphasic, right hemisphere damaged patients also showed normal N400 effects. Only the group with moderate to severe deficits in comprehension showed significant reductions in amplitude and delayed peak latencies in the N400 waveforms.

Finally, Friederici and colleagues most recently (Friederici, Hahne, & von Cramon, 1998; Friederici, von Cramon, & Kotz, 1999) used ERPs to investigate assumptions about automatic versus controlled parsing processes in individuals with brain lesions. In the earlier of the two studies, Friederici and colleagues looked at the early left anterior negativity (ELAN), N400 and P600 effects in two patients: one with Broca's aphasia and the other with Wernicke's aphasia (Friederici, Hahne, & von Cramon, 1998). While the N400 effect appears to reflect semantic processes, the ELAN and P600 effects have been found to correlate with syntactic phrase structure violations and syntactically non-preferred structures or outright syntactic violations, respectively (ELAN: Neville, Nicol, Varss, Forster, & Garrett, 1991; Hahne & Friederici, 1999; P600: Osterhout & Holcomb, 1992; Hagoort, Brown, & Groothusen, 1993). Friederici, *et al.* found that the Broca's aphasic showed both an N400 and a P600 effect, however, did not show the early negativity (ELAN) typically seen in normal controls. On the other hand, the Wernicke's aphasic showed the early negativity and the P600 component, however, he did not show an N400 effect. Despite some variation from normal controls, their patients did show some normal patterns of implicit activation.

Though in its early stages of inception as an assessment tool for individuals with aphasia, the technique of using ERPs as a diagnostic index of processing deficits has already provided evidence of its benefits in the assessment and therapeutic processes. In revealing that a patient demonstrated implicit processing, despite an inability to comprehend and perform basic explicit tasks, one group of researchers was able to extend rehabilitation services for an individual, and ultimately achieve success with auditory comprehension tasks. Though yet to be investigated, this technique may also be useful in predicting recovery patterns. At a minimum, it has provided evidence that, at least initially, something of the signal is being processed in a relatively normal fashion. As the "severely-impaired" subjects tested by Hagoort and colleagues (1996) did show different patterns of brain activity for congruous versus anomalous stimuli, that are consistent with patterns observed in normal controls, and as the Wernicke's aphasic tested by Friederici, *et al*. (1998) showed an intact ELAN waveform, it may be reasonable to conclude that some recognition is taking place, and that the breakdown occurs somewhere further along in the processing stream.

Implicit Processing: Head-Mounted Eye Tracker

In the same vein as ERPs, eye tracking techniques have begun to be proposed as a way to index implicit processing in individuals with aphasia. The first empirical study to link eye movements to spoken-language comprehension was conducted by Cooper in 1974. The main finding of Cooper's work was that eye movements to pictures were closely time-locked to semantically relevant information in a simultaneously presented spoken story. Despite the obvious far-reaching potential of this method in investigating language processing, interest in tracking eye movements quickly died off until the mid-1990's. It was not until 1994 that the head-mounted eye tracker was first presented, at the annual CUNY conference on sentence processing, as a research tool for use in investigating language processing (Tanenhaus, Magnuson, Dahan & Chambers, 2000). From that time on, numerous studies have been conducted to investigate a wide variety of empirical questions, including: which domains listeners consider when interpreting reflexives and pronouns (Runner, Sussman, & Tanenhaus, 2000); how argument structure is used in comprehending filler-gap dependencies (Sussman & Sedivy, 2000); how listeners use lexical conceptual knowledge as a way to predict upcoming information (Altmann, Haywood & Kamide, 2000); and what is the time course of reference resolution (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). As with ERPs, eye movement trackers allow for continuous measurements without disrupting the speech stream and do not require metalinguistic judgments, thus making them well-suited for use with neurologically impaired populations.

A thorough investigation of the literature in the eye-tracking domain failed to turn up any studies to date that were carried out with individuals suffering from aphasia. However, a pilot study conducted by Brooke Hallowell and colleagues (Hallowell, Wertz & Kruse, 2002), demonstrated that for a group of healthy young adults, eye movement fixation times indexed accurate comprehension of target times as measured on the Revised Token Test, a standard assessment tool used with neurologically impaired populations. As the authors did demonstrate consistency in the pattern of eye movement responses, patterns that could be correlated with successful auditory comprehension, this tool appears to be quite promising in its ability to assess the same auditory comprehension in populations not capable of overt responses. It is also a technique with all of the same benefits as the ERP for measuring implicit processing in severely compromised patients.

Discussion

As our knowledge of brain organization and function advances, so must our standard tools of assessment. We design elaborate tools that assist in furthering our understanding of the intricate workings of the brain, however, they often do not find their way into everyday clinical practice. It is not surprising that we do not rigorously utilize these techniques - the clinical community at large continues to hold on to outdated views about brain organization and language processing in individuals with aphasia. Our medical books and linguistic texts continue to map out syndromes as discrete entities that have limited overlap in symptomatology, despite the abundance of research that has provided much more evidence to the contrary - a point often left out of such textbooks. Though beneficial in providing a hint about the potential deficits one may expect to see in a given patient, classifications and aphasia quotients (obtained via standardized tests) should be interpreted with caution. They are not a foolproof way of determining the specific lesion site, array of distinct deficits, or underlying cause of communication failure (i.e., loss of knowledge, or deficient access).

By using assessment measures that can tap into real-time implicit and explicit processing, we have broadened our understanding of the nature of the breakdown in aphasia that goes beyond traditional classifications. We have discovered that Broca’s aphasics are vulnerable to comprehension failures, even at the single word level. They show reductions in priming of the less frequent meanings of words as well as words that are not good exemplars. In addition, Broca’s aphasics show a protraction in the time-course of lexical activation. Hence, the breakdowns seen in this population at the sentence and conversation levels may have more to do with lexical activation problems than was once suspected. As such, the patient may benefit from rehabilitative work on lexical comprehension at the single word level despite no overt deficits: designing tasks that require a subject to process single lexical items in noise, in speeded presentation or when the meaning may be ambiguous might possibly facilitate the recovery process.

These new tools sensitive to online processing, e.g. ERPs and eye-tracking, have also uncovered sparing of implicit processing in individuals with global and Wernicke's aphasia. Patterns of brain activity in these patients, as measured by the N400 effect, do not appear to be significantly different than normal controls. This discovery has extraordinary clinical implications. The finding that a patient may have intact implicit processing can mean the difference between treatment discharge and further intervention. It also has an impact on how we counsel families regarding what the patient may or may not be understanding. Furthermore, it may prove useful in mapping out recovery patterns. ERPs and eye tracking allow for different types of linguistic processing to be measured, and therefore can provide greater information about the nature of the deficit. For example, the N400 requires stimuli to be designed in an expected versus an unexpected condition, i.e., "He takes his coffee with cream and *sugar,"* versus "He takes his coffee with cream and *dog*." This provides information about whether or not the subject has expectations about the target given the context. Eye-tracking, on the other hand, allows one to ask and answer a question such as, "Does the patient know which circle is the *brown* one as opposed to the *blue* one?" by stating, "Look at the *brown* circle," and monitoring eye-movements. These data may provide useful information about which types of processing may remain implicitly intact, and which do not, as well as providing predictions about the long-term prognosis for recovery from deficits in different domains of processing.

As our knowledge of brain dynamics has evolved, so should our standard assessment tools. We have the methods with which to better assess our clients, and ultimately design more effective treatment plans. It is imperative that we begin to implement them into our every day clinical practice.

References

Altmann, G., Haywood, S., & Kamide, Y. (2000). Anticipating grammatical function: Evidence from eye movements. Paper presented at the *13th Annual CUNY Conference on Human Sentence Processing.* La Jolla, CA.

Bates, E., Friederici, A., & Wulfeck, B. (1987). Grammatical morphology in aphasia: Evidence from three languages, *Cortex, 23,* 545-574.

Berndt, R., Caramazza, A. (1999). How "regular" is sentence comprehension in Broca's aphasia? It depends on how you select the patients. *Brain and Language, 67,* 242-247.

Blackwell, A. & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience, 7,* 228-257.

Blumstein, S., Milberg, W., Shrier, R. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language, 17,* 301-315.

Caplan, D. & Waters, G. (1999). Verbal working memory and sentence comprehension, *Behavioral and Brain Sciences, 22,* 77-126.

Caramazza, A., Berndt, R.S., Basili, A.G. & Koller, J. (1981). Syntactic processing deficits in aphasia. *Cortex, 17,* 333-348.

Connolly, J., Major, A., Allen, S., & D'Arcy, R. (1999). Performance on WISC-III and WAIS-R NI vocabulary subtests assessed with event-related brain potentials: an innovative method of assessment. *Journal of Clinical and Experimental Neuropsychology, 21,* (4), 444-464.

Connolly, J., Mate-Kole, C., & Joyce, B. (1999). Global aphasia: An innovative assessment approach. *Archives of Physical and Medical Rehabilitation, 80,* 1309-1315.

Connolly, J. & D'Arcy, R. (1999). Innovations in neuropsychological assessment using event-related brain potentials. *International Journal of Psychophysiology, 37,* 31-47.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory, & language processing. *Cognitive Psychology, 6,* 84-107.

Dick, F., Bates, E., Wulfeck, B., Utman, J., Dronkers, N., & Gernsbacher, M. (2001). Language deficits, localization and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review, 108,* 4, 759-788.

Dirks, D., Morgan, D. & Dubno, J. (1982). A procedure for quantifying the effects of noise on speech recognition, *Journal of Speech and Hearing Disorders, 47,* 114-123.

Eberhard, K., Spivey-Knowlton, M., Sedivy, J. & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of psycholinguistic Research, 24,* 409-436.

Folstein, M., Folstein, S., and McHugh, P. (1979). Mini-Mental State Examination. *Journal of Psychology Research., 12:* 196-198.

Friederici, A., Hahne, A., & von Cramon, D. (1998). First-pass versus second-pass parsing processes in a Wernicke's and a Broca's aphasic: Electrophysiological evidence for a double dissociation. *Brain and Language, 62:* 311-341.

Friederici, A. , & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research, 28,* 467-484.

Friederici, A., von Cramon, D., & Kotz, S. (1999). Language related brain potentials with cortical and subcortical left hemisphere lesions. *Brain, 122,* 1033-1047.

Goodglass, H. (1993). *Understanding aphasia.* San Diego, CA: Academic Press.

Goodglass, H. & Baker, E. (1976). Semantic field, naming and auditory comprehension in aphasia. *Brain and Language, 3*: 359-374.

Gordon-Salant, S. & Fitzgibbons, P. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research, 36,* 1276-1285.

Gordon-Salant, S. & Fitzgibbons, P. (1995). Recognition of multiply degraded speech by young and elderly listeners, *Journal of Speech and Hearing Research, 38,* 1150-1156.

Gordon-Salant, S. & Fitzgibbons, P. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech, Language & Hearing Research, 40,* 423-431.

Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language, 50,* 27-51.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences, 23,* 1-71.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439-483.

Hagoort, P., Brown, C. & Swaab, T. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain, 199,* 627-649.

Hahne, A., & Friederici, A. (1999). Functional neurotopography of syntactic parsing: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 193-204.

Hallowell, B., Wertz, R., & Kruse, H. (2002). Using eye movement responses to index auditory comprehension: An adaptation of the Revised Token Test. *Aphasiology, 16,* 587-594.

Helfer, K., & Wilbur, L. (1990). Hearing loss, aging and speech perception in reverberation and noise. *Journal of Speech and Hearing Research, 33,* 149-155.

Just, M. & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kean, M.L. (1985). *Agrammatism.* Orlando, FL: Academic Press.

Kertesz, A. (1982). *Western Aphasia Battery test booklet.* New York: The Psychological Corporation.

Kilborn, K. (1991). Selective impairments of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language, 41,* 275-288.

Kutas, M. & Hillyard, S. (1980). Event-related potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11,* 99-116.

Kutas, M. & Van Petten (1994). Psycholinguistics electrified: event-related brain potential investigations [Review]. In : Gernsbacher M., (Ed). Handbook of psycholinguistics. San Diego: Academic Press, pp. 83-143.

Laine, M., Goodglass, H., Niemi, J., Koivuselka-Sallinen, P., Tuomainen, J. & Marttila, R. (1993). Adaptation of the Boston Diagnostic Aphasia

Examination and the boston Naming Test into Finnish. *Scandinavian Journal of Logopedics and Phoniatrics, 18,* 82-93.

Milberg, W. & Blumstein, S. (1981). Lexical desision and aphasia: Evidence for semantic processing. *Brain and Language, 14,* 371-385.

Milberg, W., Blumstein, S., & Dworetzky, B. (1987). Processing of lexical ambiguities in aphasia. *Brain and Language, 31*, 138-150.

Milberg, W., Blumstein, S., & Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society, 26,* 305-308.

Miyake, A., Carpenter, P. & Just, M. (1994). A capacity approach to syntactic comprehension disorders: making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11,* 671-717.

Moineau, S. & Bates, E. (2001). Pushing the limits of word comprehension in normal and aphasic listeners. Unpublished manuscript.

Neville, H., Nicol, J., Barss, A., Forster, K. & Garrett, M. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3,* 151-165.

Osterhout, L. & Holcomb, P. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785-806.

Prather, P., Zurif, E., Stern, C., & Rosen, T. (1992). Slowed lexical access in nonfluent aphasia: A Case study. *Brain and Language, 43,* 336-348.

Prather, P. (1994). The time course of lexical activation in fluent and nonfluent aphasia. *Linguistics and Cognitive Neuroscience, 6,* 128-144.

Prather, P., Zurif, E., Love, T., & Brownell, H. (1997). Speed of lexical activation in nonfluent Broca's aphasia and fluent Wernicke's aphasia, *Brain and Language, 59,* 391-411.

Revonsuo, A. & Laine, M. (1996). Semantic processing without conscious understanding in a global aphasic: Evidence from auditory event-related brain potentials. *Cortex, 32,* 29-48.

Rugg, M. (1995). Event-related potential studies of human memory. In: Gazzaniga, M. (Ed.). The cognitive neurosciences*.* MIT Press, Cambridge, MA., pp 789-801.

Runner, J., Sussman, R., & Tanenhaus, M. (2000). Binding reflexives and pronouns in real-time processing. Poster presented at the *13th Annual CUNY conference on Human Sentence Processing,* La Jolla, CA.

Simpson, G. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96,* 316-340.

Sussman, R. & Sedivy, J. (2000). Using eyetracking to detect and describe failed gap effects. Poster presented at the *13th Annual Conference on Human Sentence Processing.* La Jolla, CA.

Swaab, T., Brown, C. & Hagoort, P. (1997). Spoken sentence comprehension in aphasia: Event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience, 9,* 39-66.

Swaab, T. (1998). Event-related potentials in cognitive neuropsychology: Methodological considerations and an example from studies of aphasia. *Behavior Research Methods, Instruments & Computers, 30,* 157-170.

Swaab, T., Brown, C. & Hagoort, P. (1998). Understanding ambiguous words in sentence contexts: Electrophysiological evidence for delayed contextual selection in Broca's aphasia. *Neuropsychologia, 36,* 737-761.

Swinney, D., Zurif, E., & Nicol, J. (1989). the effects of focal brain damage on sentence processing: An examination of the neurological organization of a mental module. *Journal of Cognitive Neuroscience, 1,* 25-37.

Tanenhaus, M., Magnuson, J., Dahan, D., & Chambers, C. (2000). Eye moevements and lexical access in spoken-language comprehension: evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29,* 557-580.

Utman, J. & Bates, E. (1998). Effects of acoustic degradation and semantic context on lexical access: Implications for aphasic deficits [Abstract]. *Brain and Language, 65,* 516-518.

Utman, J., Blumstein, S., & Sullivan, K. (2001). Mapping from sound to meaning: Reduced lexical activation in Broca's aphasics, *Brain and Language, 79,* 444-472.

Zurif, E., Caramazza, A., Myerson, R., & Galvin, J. (1974). Semantic feature representations for normal and aphasic language. *Brain and Language, 1*, 167-187.

Cognitive Science Online, Vol.1, pp.22-33, 2003 http://cogsci-online.ucsd.edu

A discussion and review of Uttal (2001) *The New Phrenology* [*]

Edward M. Hubbard

Department of Psychology and
Interdisciplinary Program in Cognitive Science
University of California, San Diego
edhubbard@psy.ucsd.edu

Abstract

In his recent book, William R. Uttal suggests that the modern attempts to understand the localization of cognitive functions are misguided. According to Uttal, attempts to localize cognitive functions will fail because the to-be-localized cognitive processes simply do not exist. In this review, I will defend the position that the mind can be meaningfully divided into cognitive parts and, given that the brain is non-homogenous, we may be able to localize these cognitive parts to brain parts. However, this view does not posit the simple localization that Uttal argues against. Instead, I argue that mental functions, broadly construed, are distributed, but are composed of multiple sub-components that can be localized to specific brain regions.

Introduction

I am neither a universalizer nor a localizer…In consequence I have been attacked as a universalizer and also as a localizer. But I do not remember that the view I really hold as to localization has ever been referred to. If it is, it will very likely be supposed to be a fusion of, or a compromise betwixt recent doctrines.

John Hughlings Jackson, (1882/1932)

One of the fundamental problems in philosophy, psychology, neuroscience and cognitive science is the extent to which mental functions can be decomposed, and once decomposed, how closely they can be tied to specific brain locations. Descartes famously argued that the mind was a unified entity and that it could not be divided. However, with increasing knowledge of the mind, Descartes' position has largely been abandoned in favor of the view that the mind can be decomposed into a number of functional parts and that these mental parts have their basis in different parts of the brain. The first (and most notorious) of these ideas was phrenology, first proposed by Franz Joseph Gall in 1796. This view has rightly been criticized on many grounds, but the essential insight, that the mind is composed of distinct functional parts, and that

---

[*] Uttal, W.R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain.* Cambridge, MA : MIT Press, 255 p., ISBN 0262210177.

these parts can be localized to specific parts of the brain, has remained an essential component of modern thinking about the relation between the mind and the brain. Since the time of Gall, there has been tremendous progress in our knowledge of both the mind and the brain, but whether, and how precisely, cognitive functions can be localized in the brain is still open to debate.

In *The New Phrenology*, William R. Uttal (2001) explores the fundamental assumptions underlying the attempt to localize cognitive functions in specific brain areas (what Uttal refers to as the “localizationist approach”). In his book, Uttal concludes, contrary to prevailing opinion, that the localizationist enterprise is beset by such deep fundamental problems that any attempt to localize cognitive functions in the brain is bound to failure. Uttal’s book is organized around three main questions:

1. Can the mind be subdivided into components, modules or parts?

2. Does the brain operate as an equipotential mass or is it also divisible into interacting but separable functional units?

3. Can the components, modules, or parts of the mind, if they exist in some valid psychological sense, be assigned to localized portions of the brain?

Uttal's answer to question 2 is a limited yes, the answer to question 1 is a decided no, and therefore the answer to question 3 must also be no. That is, there can be no meaningful relation between the parts of the mind and the parts of the brain because there are no meaningful parts of the mind to speak of.

To make his case that the answer to question 1 is “no” Uttal draws on both the history of neuroscience and psychology and current findings in cognitive science. Here, Uttal divides his critique into three main parts: In the first part, he reviews the history of localization attempts; In the second part, he critiques the idea that the mind can be divided into functional parts, especially for the higher cognitive processes, and in the third part, he lists a number of problems with the technologies that have been used to infer localization of mental function from localization of brain processes.

While many of Uttal’s criticisms are valid and should be borne in mind when evaluating extreme claims of localization, I find that Uttal's skeptical conclusion is unwarranted. Contrary to Uttal's claims, there has been progress in our understanding of both cognitive functions and their location in the brain, and there is every reason to believe that the problems involved with localization can be overcome, although not by the simplistic model of localization Uttal critiques.

Organization of the Book

Uttal’s book is organized around three broad themes, illustrated with numerous examples. In the first chapter, Uttal outlines the general problem, as he sees it, beginning with a discussion of historical attempts to localize brain function. To Uttal, the apparent unity of the mind poses a *prima facie* problem for theories of the mind that assume that it can be divided into functional parts that can then be localized to specific brain regions.

In chapter two, Uttal briefly digresses from his main theme to explain the technologies that have been used in cognitive neuroscience. Uttal describes methods from the level of single unit recording and surgical techniques to ERPs (which he refers to as "evoked brain potentials, EVBPs") computerized tomography (CT) scans, positron emission tomography (PET) and blood oxygen level dependent functional magnetic resonance imaging (BOLD fMRI). This section is relatively solid overall, and it is probably a good reference for introducing people to the techniques of modern cognitive neuroscience, but a number of minor flaws suggest that Uttal is either not careful in his reading, or that he is not completely prepared to discuss these issues. For example, the use of EVBP instead of ERPs suggests that Uttal is not reading current literature. Additionally, in his discussion of techniques to inactivate cortex, Uttal does not include recent developments (such as the use of the reversible GABA-agonist muscimol), which do not suffer many of the problems Uttal describes (see below). However, the most obvious error is that Uttal refers to "the central or Sylvian sulcus" (p. 30). However, the central sulcus is also known as the Rolandic fissure, while the Sylvian fissure is also known as the lateral sulcus. These are the two largest and best-known landmarks in the brain, and so Uttal should have taken extra care to get these details right. Overall, the chapter is adequate, but I wouldn't use it to study for an exam.

In chapter three, Uttal gets into the heart of his argument, laying out his reasons for being skeptical that a taxonomy of cognitive processes is even possible. Since a great deal of the burden of Uttal's argument lies here, I will dedicate the greatest part of this review to this portion of Uttal's book. To briefly summarize the argument here, Uttal claims that there has been, and can be, no progress on the problem of developing a taxonomy of cognitive processes, and therefore there can be no hope of localizing cognitive processes in the brain.

In the fourth chapter, Uttal focuses on technical problems that limit the inferences about localization that can be drawn on the basis of current methods in neuroscience. Many of the criticisms in this chapter are well-known to practicing neuroscientists, but the assemblage presented here is impressive. I will argue, however, that Uttal underestimates the power of converging research methods and the self-correcting nature of science to address problems in individual studies.

In the fifth and final chapter, Uttal argues that, instead of fractionating cognitive processes into parts, and then attempting to localize these parts to specific parts of the brain, we should return to a "molar level" analysis of behavior, focusing on input-output relations, instead of postulating unobservable cognitive processes. I will briefly discuss each of these sections, addressing some concerns that Uttal raises along the way, finally arguing that Uttal's concerns, while cause for caution, do not cause the entire enterprise of localization in the brain to unravel, and that, even more so, they do not necessitate Uttal's proposed return to behaviorism.

A Brief History of the Parts of the Mind

Uttal's main concern is that there has been, and can be, no progress on the problem of developing a taxonomy of cognitive processes, and therefore there can be no hope of localizing cognitive processes in the brain. "The preeminent problem in achieving a general solution to the localization issue lies in defining the psychological processes and mechanisms for which loci are being sought" (p.16).

To demonstrate this, Uttal dedicates over 40 pages to reviewing two thousand years of theories on the various different cognitive processes. While considerations of space do not allow me to reiterate Uttal's review here, anyone interested in psychology, cognitive science, or systems level neuroscience should consider reading this section.

In the fundamental taxonomy of cognitive psychology, Neisser proposed a number of categories of the mind, such as iconic storage, pattern recognition, focal attention, visual memory, speech perception and memory and thought (Neisser, 1967; Matlin, 1994). For Uttal, the fact that decades later, the main components of Neisser's taxonomy remained unchanged, simply serve to indicate that a pedagogical tool had become reified into the taxonomy of what could reasonably be sought in the brain. One alternative interpretation, not explored by Uttal, is the possibility that this taxonomy remains relatively unchanged not because it has been uncritically reified, but rather that it has provided a coherent account of mental activity that leads to testable predictions, and which has survived experimental disconfirmation. Compare this with other such theories, such as the ideas of earlier faculty psychology (the form of psychology that the early Phrenologists used), which of course are completely out of favor. Patricia and Paul Churchland have been especially vociferous advocates of the idea that our mental taxonomy is a corrigible theory, and that many ideas about the mind we currently hold will eventually turn out to be like the now discredited ideas of phlogiston or élan vital (see, e.g., P.S. Churchland, 1986; P.M. Churchland, 1989). Much future work remains to be done, but there is, again, no in principle reason to believe that we cannot develop an adequate taxonomy of the mind suitable for the task of localizing to parts of the brain.

It should also be noted that, while the fundamental taxonomy of the mind has remained relatively unchanged, the exact manner in which we have explained and modeled these mental abilities has undergone dramatic revision. In the early days of cognitive science, the mind was conceived of as a collection of black boxes, each autonomously performing computations on representations that were in some way a "language of thought" guided by sentence-like rules. With the advent of connectionism and alternative approaches to cognition, we now realize that mental computations are more likely carried out by interconnected networks of neurons, in which information is stored as patterns of synaptic strength, and computation can be thought of as transformations of high-dimensional neural vectors (see, e.g., Churchland & Sejnowski, 1992). These drastic changes of perspective at the microlevel have been mirrored by changes at the macrolevel by the realization that a great deal of cognitive activity takes place in situated groups of agents, interacting with their environment (Hutchins, 1995). One of Uttal's concerns (p. 143) is that the rigid, serial models of cognitive psychology are implausible. However, much work in cognitive science, building from neural networks, also denies this rigid, serial model, and yet attempts to localize cognitive functions to parallel distributed networks in the brain.

Another main concern for Uttal is the ever-greater number of cognitive parts that have been posited by localizationists. However, a careful analysis of the literature demonstrates that the path to progress in cognitive science lies in subdividing the components of the mind, which we can then attempt to localize to parts of the brain. For example, consider the case of memory. Early psychological investigations of memory treated it as a unitary construct. However, more recent analyses of memory, combining both psychological and neuroscientific approaches have suggested that "memory" really should be thought of as a variety of separable components with different proposed neural substrates (although the exact taxonomies of memory are subject to debate). One such classical example comes from Larry Squire's work on patients with hippocampal damage, in which the ability to form new "declarative" memories is impaired, while the ability to learn new "non-declarative" memory for skills and abilities is spared (for a review, see Squire & Kandel, 1999). While a great deal of work, both empirical and theoretical, remains to be done, the basic lesson that progress in cognitive science will come through subdivision of cognitive functions is one that seems to have been completely lost on Uttal.

Modularity and Localization

Uttal, following Fodor (1983), argues that certain input and output functions are modular, but central cognitive processes may not be. On Fodor's account, there are several important distinctions between the input and output processes and the cognitive processes. First, information in the input and output processes can be meaningfully encapsulated. Second, each of the input processes is clearly linked to a specific anchor of a specific class of physical stimulus. Third, the dimensions of the physical stimulus can be clearly identified, and therefore the dimensionality of the mental experience can also be clearly identified.

Based on these results, Uttal claims that, even if it were to turn out that certain input and output functions could be localized, it is not clear that cognitive functions could be localized because the taxonomy of mental processes seems to be open ended and there is no way to achieve closure in such a list of possible faculties. However, the discussion of memory, above suggests that it may be possible to localize certain cognitive capacities (e.g., memory) in the same way that we have localized certain perceptual capacities. Additionally, much recent evidence, such as that coming from embodied cognitive linguistics, has suggested that the strict distinction between perceptual-motor and cognitive functions may be illusory (see e.g., Lakoff & Johnson, 1999; Johnson, 1987) thus making it less clear exactly where Uttal would draw the dividing line between those processes that can be localized and those that cannot.

Uttal further argues, even if perceptual and motor functions can be identified with specific brain regions, "there is considerable evidence that…any but the simplest (sensory or motor) cognitive function involves large and distributed regions of the brain" (p. 155). However, the problem here for cognitive science is only apparent. One of the fundamental strategies in cognitive science is the technique of recursive decomposition (Palmer & Kimchi, 1986), in which a single component (e.g., the mind) is divided into smaller and smaller functional units (e.g., perception, attention, memory or action). These components can then be further recursively decomposed into smaller and smaller parts (e.g., processing of color, motion or orientation), which can then be localized to specific parts of the brain (e.g., cortical visual areas V4, MT and V1).

However, it is not only perceptual functions like vision that can undergo recursive decomposition. Consider the case of mathematics, a paradigmatic case of a cognitive function. To define the problem more carefully, where in the brain do arithmetic calculations occur? There is no one place in the brain where arithmetic calculations occur. Instead, arithmetic calculations activate a widespread network of brain regions, consistent with Uttal's arguments. However, it is also possible to decompose the process of performing arithmetic problems into several key components. First, a visually presented number (a grapheme) must be recognized. Second, the numerical magnitude of the grapheme must be recognized, and finally, computations must be invoked to determine the correct solution for the problem at hand. Additionally, this process may draw on stored math facts such as the overlearned knowledge that “2 + 2 = 4.” This model, dubbed the “triple-code” model, (Dehaene, 1992; 1997) was derived from a variety of different sources of evidence, including reaction time studies, neuropsychological deficits, and early imaging (PET) studies.

Not only does this model propose a specific taxonomy of the cognitive processes involved, but it also suggests that specific regions of the brain would be essential for the performance of these functions. Based on previous findings and the new, refined framework in which to study arithmetic calculation, the recognition of the visual grapheme could be expected to depend on structures in the inferior temporal lobe (fusiform and lingual gyri) while the ability to perform the numerical calculation would likely depend on the angular gyrus, and the ability to retrieve stored mathematical facts would depend on frontal lobe structures. This model has since received further verification from studies making use of a variety of techniques. For example, using cortical recording electrodes Allison *et al.* (1994) find that numbers and letters are represented in fusiform regions near other regions associated with the recognition of visual objects, while Whalen *et al.*, (1997) find that stimulating frontal cortex causes a temporary disruption of stored arithmetic facts. Göbel *et al.*, (2001) report that transcranial magnetic stimulation (TMS) over angular gyrus disrupts the ability to perform novel computations but does not impair the ability to report the answer to overlearned problems like 2 + 2. Finally, several recent fMRI studies have confirmed and extended this overall model (Pesenti *et al.* 2000; Rickard *et al.* 2000), although as in the case of memory, there remain unresolved questions.

One of the main points of this extended example is that it matters what you are attempting to localize. That is, there is an interaction between how cognitive processes are defined and how they are localized. Attempts to localize mathematical cognition would fail because performing arithmetic problems depends on a distributed network of brain areas. However, if we can appropriately define the subcomponents of mathematical cognition, it may be possible to identify different cognitive parts and localize them to specific brain parts. Returning to the epigraph at the beginning of this review, brain functions are neither localized, nor distributed throughout the whole brain, and this is the modal view in cognitive neuroscience today. By considering the interaction between the cognitive parts that we postulate and our ability to localize them, we see that claims that cognitive functions are distributed is an argument against something of a localizationist straw man, a view that very few (if any) serious researchers in this area actually believe.

Technical Problems in Lesion Analysis

In order to further support his claim that cognitive functions will never be reduced to specific neural substrates, Uttal describes a number of technical problems that limit the inferences about localization that should be drawn on the basis of the lesion method (both natural lesions in humans and surgical ablation techniques in animals).

For example, Uttal points out that the various ablation techniques (ways of removing brain tissue to see what happens when that tissue is removed) all suffer different forms of limitations, such as effects due to inadvertently severed fibers of passage or due to recovery of function. However, neuroscientists are aware of these difficulties and are always searching for newer and better techniques to overcome these difficulties. Recently, there have been a family of antagonists developed that are able to reversibly, selectively inactivate certain populations of neurons (for example, the GABA agonist muscimol described above), overcoming some of the difficulties Uttal has discussed with this class of experiments.

Another problem for the localizationist is the heterogeneity of lesions and the individual variability in brain organization. Although it has been at the core of the localizationist enterprise since the earliest days, the lesion method still suffers from numerous problems of interpretation. After the downfall of phrenology, the idea of localization of function was revived by Paul Broca's studies of patients with brain lesions in the 1860s (Greenblatt, 1995; Star, 1989). Broca studied numerous patients with deficits in their ability to produce language (aphasia), independent of any motor problems (dysarthria) and concluded that the left inferior frontal cortex was critically involved in language production. Similarly, Carl Wernicke proposed that an area in the left posterior superior temporal region was critical for comprehension of language.

Most people are familiar with this basic taxonomy, which is still in use today: Broca's area is critical for language production and grammar while Wernicke's area is crucial for comprehension. However, more recent evidence has demonstrated that lesions to insula lead to nonfluent aphasias more so than Broca's area (Bates *et al.*, submitted; Dronkers, 1996) and that many different lesions can cause agrammatism (Dick *et al.*, 2001). Additionally, recent research suggests that Broca's area may house the human mirror neuron system (Iacoboni *et al.*, 1999), although this may not necessarily be a serious contradiction, as the mirror neuron system may be critical for the evolution of language (Arbib & Rizzolatti, 1996). Similarly, Wernicke's area is believed to be critical for speech comprehension and phonological encoding, but others have found (Saygin *et al.*, in press) that Wernicke's area lesions can also cause deficits in the comprehension of nonverbal environmental sounds (even significantly more than speech sounds) and thus processing in this region is probably not specific language, but instead may be involved in any complex auditory analysis. Again, these problems demonstrate that how you divide up and refine the processes within a cognitive task/domain is important, and we should always be aware of these problems when pursuing the localizationist goals. However, a look at the study of language localization also demonstrates that conceptual revision and progress is still occurring, sometimes top-down, as we come to better understand language, and sometimes bottom-up, as we explore its neural basis (see also, Bates & Dick, 2000).

Technical Problems with Functional Imaging

In addition to the problems with the lesion method, Uttal lists a number of concerns about one of the primary methods used in cognitive science, functional imaging. This concern about the interpretation of imaging data is, of course, not new (for a review of the problems with fMRI and the inferences that one can and cannot make on the basis of the BOLD signal, the interested reader is referred to Cabeza and Kingstone, 2001 or to any of the numerous online texts discussing potential MRI artifacts see, e.g., Hornak, 1999).

One of the concerns that most clearly relates to our current concerns about localization is the problem of the arbitrary threshold in fMRI. As Uttal notes, "Evidence of sharply defined and highly localized artificial boundaries arising from a poor choice of a threshold could easily lead to an erroneous conclusion about the cerebral localization or nonlocalization of a psychological process" (p. 167-168). This is especially true in the case of certain classes of statistical analyses.

However, these types of threshold analyses are not the only methods for analyzing fMRI data available, and as new and more powerful techniques are developed, new analyses (such as independent components analysis, principal components analysis or network connectivity analysis) will be performed that will provide an alternative to the arbitrary threshold setting methods. For example, in much current vision research, one technique is to use a "reference" scan against which to have a baseline or standard and then to compare activation relative to that standard. This does not entirely avoid the problem of thresholds, but can make them a little less arbitrary. As these types of techniques find their way into higher cognitive processes, the problem of the baseline will become less arbitrary.

An additional change in imaging methods is a return to region of interest (ROI) analyses. In the early days of functional imaging, ROI analyses were common, as researchers were interested in simply confirming the validity of the new techniques. However, as researchers became more confident in their techniques, imaging became more and more commonly used as an exploratory tool. When used as an exploratory tool, researchers often simply look for the "hottest" spot in the brain and leave their analysis at that. More recently, however, there has been a return to the ROI technique to explore activity in specific brain regions, on the basis of a prior hypothesis, and to quantify the manner in which the brain responds to various different types of stimulation and in different tasks. This hypothesis driven usage of fMRI can be used to avoid many of the problems associated with the use of arbitrary thresholds.

The Scientist in a Vacuum

There is, however, a deeper problem with Uttal's analysis, that I will call the "scientist in a vacuum fallacy." Throughout the book, Uttal raises valid concerns about the extent to which we can infer localization from any one experiment or any one methodology. For example, from the fact that damage to one region of the brain causes a particular deficit, we cannot infer that the part of the brain damaged is necessarily the seat of that function (if you remove a transistor from a radio and it starts to squawk, it doesn't follow that the transistor is a squawk suppressor). Or, finding a particular peak of activity in the brain doesn't imply that the locus of that particular peak is "the" region of the brain that performs this particular function.

No scientist works in a vacuum. Rather, every scientist has colleagues and competitors who reinforce or challenge his or her results (see, for example, Latour 1987). These colleagues and competitors may replicate (or fail to replicate) any of the results that a given researcher may arrive at. So, for example, if a given researcher were to find a unique spot of activity by chance alone that actually had no bearing on the cognitive functions being investigated, then other future researchers would fail to find a similar spot of activity, and the results would therefore be regarded as an error or an anomaly. That is, although Uttal has raised the specter that noise may contaminate the results of a single experiment, or even a family or experiments, he has not demonstrated that there is a clear bias in the methods of cognitive neuroscience, writ large, that would lead to *systematic* errors regarding whether cognitive functions are localized or not.

Conversely, if the same localization was to be supported by similar results from deficits after brain lesions, deficits induced by TMS, and increased activity as indicated by the fMRI BOLD signal, we would then feel justified in concluding that there was in fact a real neural basis for the proposed localization. This is because different techniques have different strengths and weaknesses and different potential artifacts. The possibility that each method would independently hit upon the same location simply due to these technical issues is vanishingly small. For example, in the case of mathematical cognition above, many researchers, making use of many different techniques have reached similar conclusions regarding the localization of the cognitive processes involved in arithmetic cognition. The collaborative, self-correcting nature of scientific research and the use of multiple converging research methodologies make it likely that future localizationist ventures will be more and more accurate.

Conclusions

The history of localizationist research has long been fraught with bitter battles as to whether the brain, like the other organs in the body, can be divided into subparts, which perform different functions. Uttal would have us believe that the answer to this question is no. Instead, he argues that the mind is a unitary phenomenon, and that attempts to subdivide the mind suffer from severe logical difficulties, especially since the mental faculties that we are attempting to localize can only be observed indirectly. However, this problem is no greater than the problem faced by other domains of science. For example, particle physicists cannot directly observe sub-atomic particles. Instead, they must infer their existence from the behavioral effects on other physical measuring devices, such as a cloud chamber. Therefore, there is nothing uniquely privileged about mental states.

Since the mind can be meaningfully decomposed into subprocesses (e.g., memory, language comprehension), and can be inferred from the behavior of the system, there is no logical, in principle, barrier to developing an appropriate taxonomy of the mind even though there is still a great deal of debate about the details of this taxonomy. Uttal's historical arguments, contrary to his conclusions, do demonstrate that significant progress has been made in our understanding of the taxonomy of the mental, and further suggest that a growing consensus may be at hand. With this high level description of the mind in hand, the attempt to localize these processes in specific parts of the brain may move forward rather rapidly.

Uttal's specific concerns about the techniques used to explore the neural basis of these cognitive parts are hardly unique. These concerns are shared (and indeed have often been raised) by the very people that are attempting to localize cognitive processes to specific parts of the brain. I would argue that Uttal's concerns are, in part due to an outdated view of what the localizationist enterprise is attempting to accomplish. To localize macrolevel cognitive processes like memory or mathematics to specific neural substrates is clearly an untenable goal (as Hughlings-Jackson noted over 100 years ago) and very few researchers in cognitive neuroscience are still pursuing this simplistic goal. Additionally, Uttal's arguments suffer from a very unrepresentative view of a scientist in a vacuum, working without benefit of colleagues and competitors and restricted to only one methodology. If this were the true situation of the scientist, it would be a dire situation indeed. However, given that scientists do not work in a vacuum, these concerns should not limit the possibility of the localizationist enterprise in toto. Instead, it should merely limit our enthusiasm for specific experimental results.

It should also be noted that Uttal's approach assumes a purely top-down answer to the question of localization. That is, once we have an accurate cognitive taxonomy in hand, then, and only then, will it be possible to even begin to attempt to localize mental parts in brain parts. However, the study of brain parts can also provide a bottom-up constraint on our theories of the mental (cf. P.S. Churchland, 1986). Just as placing a couch in a living room begins to constrain the possible locations of the rest of the furniture in the room, so too does an understanding of the brain begin to constrain the theories about cognitive parts that we entertain.

In the final analysis, the conclusions of this book are truly disappointing. Uttal argues that, instead of fractionating cognitive processes into parts, and then attempting to localize these parts to specific parts of the brain, we should return to a "molar level" analysis of behavior, focusing on input-output relations, instead of postulating unobservable cognitive processes (in essence, a thinly veiled return to behaviorism). Uttal's concerns, while cause for caution, should not cause the entire enterprise of localization in the brain to unravel, and that, even more so, they do not necessitate a return to behaviorism.

In contrast to Uttal's claims, I believe that it is quite probable that we will, piece by piece, come to an understanding not only of the parts of the mind, but also of their localization to the parts of the brain. In contrast to Uttal's conclusion that we need to return to a molar level description of behavior focusing on input-output relations, I believe that progress in localizing cognitive functions to brain areas will occur when we divide the mind more and more finely, not less. Currently, we have but the dimmest glimmer of what such a theory might look like, but its outlines are beginning to take shape. As the form becomes clearer and we part the mists that obscure our view of the relation between the mind and the brain, we will one day be able to answer that ultimate question of who we are, and what makes us think.

References

Allison, T., McCarthy, G., Nobre, A., Puce, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 4* (5):544-554.

Arbib, M. A. & Rizzolatti, G. (1996). Neural expectations: A possible evolutionary path from manual skills to language. *Communication & Cognition*, *29*(3-4):393-424.

Bates, E., & Dick, F. (2000) Beyond phrenology: Brain and language in the next millennium. *Brain and Language, 71*, 18-21.

Bates, E., Wilson, S., Saygin, A.P., Dick, F. Sereno, M., Knight, R. T. & Dronkers, N. F. (submitted). Assessing brain-behavior relationships using voxel-based lesion-symptom mapping.

Binder, J. R. & Rao, S. M. (1994). Human brain mapping with functional magnetic resonance imaging. In A. Kertesz (Ed.) *Localization and neuroimaging in neuropsychology*. (pp. 185-212) San Diego, CA: Academic Press.

Cabeza, R. & Kingstone, A. (Eds, 2001) *Handbook of functional neuroimaging of cognition*. Cambridge, MA: MIT Press.

Churchland, P. M. (1989). *A neurocomputational perspective : the nature of mind and the structure of science.* Cambridge, MA: MIT Press.

Churchland, P. S. (1986). *Neurophilosophy: toward a unified science of the mind-brain.* Cambridge, MA: MIT Press.

Churchland, P. S. & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44* (1-2):1-42.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dick, F., Bates, E., Wulfeck, B., Utman, J., Dronkers, N. & Gernsbacher, M.A. (2001). Language deficits, localization and grammar: evidence for a distributive model of language breakdown in aphasics and normals. *Psychological Review, 108*(4), 759–788.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*(6605): 159-161.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (1998). *Cognitive neuroscience: The biology of the mind*. New York: W.W. Norton.

Göbel, S., Walsh, V. & Rushworth, M. F. S. (2001). The mental number line and the human angular gyrus. *NeuroImage, 14*(6): 1278-1289.

Greenblatt, S.H. (1995). Phrenology in the Science and Culture of the 19th Century. *Neurosurgery*, *37*:790-805.

Hornak, J.P. (1999) *The basics of MRI*. http://www.cis.rit.edu/htbooks/mri/mri-main.htm.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449):2526-2528.

Jackson, J.H. (1882/1932). On some implications of dissolution of the nervous system. *Medical Press and Circular 2*; Reprinted in J. Taylor, (Ed. 1932) *Selected Writings of John Hughlings Jackson*. (pp. 29-44) New York: Basic Books.

Johnson, M. (1987). *The body in the mind : The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh : The embodied mind and its challenge to Western thought.* New York : Basic Books.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.

Matlin, M. W. (1994). *Cognition* (3rd ed.). New York: Harcourt Brace.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts

Palmer, S. E. & Kimchi, R. (1986). The information processing approach to cognition. In T.J. Knapp & Lynn C. Robertson, (Eds.) *Approaches to cognition: Contrasts and controversies*. (pp. 37-77) Hillsdale, NJ: Lawrence Erlbaum Associates.

Pesenti, M., Thioux, M., Seron, X., De Volder, A. (2000). Neuroanatomical substrates of Arabic number processing, numerical comparison, and simple addition: A PET study. *Journal of Cognitive Neuroscience*, *12*(3):461-479.

Rickard, T. C., Romero, S. G., Basso, G., Wharton, C., & Grafman, J. (2000). The calculating brain: An fMRI study. *Neuropsychologia*, *38*(3):325-335.

Saygin, A. P., Dick, F., Wilson, S., Dronkers, N. & Bates, E. (in press). Neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain*, *126*.

Star, S. L. (1989). *Regions of the mind: brain research and the quest for scientific certainty*. Stanford, CA: Stanford University Press.

Squire, L. R. & Kandel, E. R. (1999). *Memory: From mind and molecules*. New York, NY: Scientific American Books.

Uttal, W. R. (2001). *The new phrenology: the limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.

Whalen, J., McCloskey, M., Lesser, R.P. & Gordon, B. (1997). Localizing arithmetic processes in the brain: Evidence from a transient deficit during cortical stimulation. *Journal of Cognitive Neuroscience*, *9*(3):409-417.