

Conscious and Unconscious Processing of Nonverbal Predictability in Wernicke's Area

Amanda Bischoff-Grethe,¹ Shawnette M. Proper,² Hui Mao,³ Karen A. Daniels,⁴ and Gregory S. Berns^{2,5}

¹Department of Neurology, ²Department of Psychiatry and Behavioral Sciences, and ³Department of Radiology, Emory University School of Medicine, Atlanta, Georgia 30322, ⁴School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332, and ⁵Georgia Tech/Emory Biomedical Engineering Department, Atlanta, Georgia 30322

The association of nonverbal predictability and brain activation was examined using functional magnetic resonance imaging in humans. Participants regarded four squares displayed horizontally across a screen and counted the incidence of a particular color. A repeating spatial sequence with varying levels of predictability was embedded within a random color presentation. Both Wernicke's area and its right homolog displayed a negative correlation with temporal predictability, and this effect was independent of individuals' conscious awareness of the sequence. When individuals were made aware of the underlying sequential predictability, a widespread network of cortical re-

gions displayed activity that correlated with the predictability. Conscious processing of predictability resulted in a positive correlation to activity in right prefrontal cortex but a negative correlation in posterior parietal cortex. These results suggest that conscious processing of predictability invokes a large-scale cortical network, but independently of awareness, Wernicke's area processes predictive events in time and may not be exclusively associated with language.

Key words: functional imaging; Wernicke's area; predictability; sequences; nonverbal grammar; awareness; entropy

The prediction of future events is a problem faced by nearly every organism and occurs across a wide variety of contexts. In humans, prediction appears in multiple domains, ranging from predicting what the stock market will do to predicting what a colleague might say. Some processes clearly are more predictable than others, and an assessment of the inherent predictability of a given process may have significant relevance to an individual. Whereas the stock market may be relatively unpredictable, other processes, like language, are highly predictable. Language, by definition, is constrained by the underlying rules of syntax, which effectively constrain the associated statistics of language (Chomsky, 1957; Pinker, 1994; Seidenberg, 1997). In the context of information transmission, this inherent predictability will limit the rate at which information can be reliably transmitted (Shannon and Weaver, 1949; Cover and Thomas, 1991). In this article, we describe the human neural circuitry associated with monitoring temporal predictability on a nonverbal task both with and without conscious awareness.

Predictability, or its converse, entropy, can be quantified by global measures of information transmission. Fundamentally, these rely on the probabilities with which events occur. Previous information may also influence predictability, in which case one may consider conditional, or Bayesian, statistics. A simple way of controlling temporal predictability is to design an artificial grammar, which is simply a set of rules governing the probability of transition from one state to another (Reber, 1967, 1993; Stadler,

1989; Cohen et al., 1990; Cleeremans and McClelland, 1991). In linguistics, a state might represent a word, with complex chains of transition probabilities capturing the statistics of the underlying grammatical rules. This type of statistical description is not limited to linguistics but can be applied to any system with discrete states, e.g., chemical reactions. More generally, these grammars, or Markov chains, can be used to generate temporal sequences with precisely defined statistics. By varying only the transition probabilities between states, the statistics can be changed without altering the underlying rules. More importantly, an overall measure of statistical uncertainty, and therefore predictability, can be defined by the entropy of the sequence. We used functional magnetic resonance imaging (fMRI) to study the neural response to nonverbal predictability and the effect of conscious awareness on this response.

MATERIALS AND METHODS

Experimental design. Thirty-six right-handed participants aged 20–47 (average age, 29.2 years) gave written informed consent for the study. Four squares were displayed horizontally across a screen (Fig. 1). The squares were individually illuminated in one of three randomly selected colors (blue, red, or yellow), and each square was illuminated for 1 sec. Participants were instructed to keep a mental count of the total number of blue squares presented during the entire session. To avoid an unbalanced experimental design because of the varying levels of difficulty in counting single digits versus either double or triple digits, participants were instructed to begin their count at 100. At the end of the scan, subjects were asked for the total number of blue squares counted so as to confirm that they performed the task correctly. A nonmotor measure, the mental count, was used for determining task performance because we were interested in the effects of uncertainty on a mental task but did not wish to confound the effect with the potential uncertainty of a motor response. The spatial order in which the squares were illuminated was determined by one of three conditions: (1) a four-element repeating spatial sequence, 1-3-2-4- . . . where the number designates the box position beginning from the left (entropy = 0); (2) a probabilistic version of this sequence (entropy ~ 1); and (3) a randomly ordered presentation (entropy = 2).

Received Sept. 16, 1999; revised Dec. 22, 1999; accepted Dec. 23, 1999.

This work was supported by the Departments of Psychiatry and Radiology, Emory University School of Medicine, the Stanley Foundation (G.S.B.), National Institute on Drug Abuse (Grant K08 DA00367 to G.S.B.), and a National Science Foundation Markey fellowship (A.B.G.).

Correspondence should be addressed to Gregory S. Berns, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 1639 Pierce Drive Suite 4000, Atlanta, GA 30322. E-mail: gberns@emory.edu.

Copyright © 2000 Society for Neuroscience 0270-6474/00/201975-07\$15.00/0

Figure 1. Experimental design of the presented task. Four squares were displayed horizontally across a screen. The squares were individually illuminated in one of three randomly selected colors (blue, red, or yellow) for 1 sec. Beginning at 100, participants kept a mental count of the total number of blue squares presented during the entire session. The sequence of spatial positions illuminated was determined by an artificial grammar with the entropy of the sequence varying between condition blocks. The given example illustrates 4 sec of a scan where the spatial sequence was presented (1-3-2-4). The occurrence of functional scans and the subject's mental count of the blue squares are also shown.

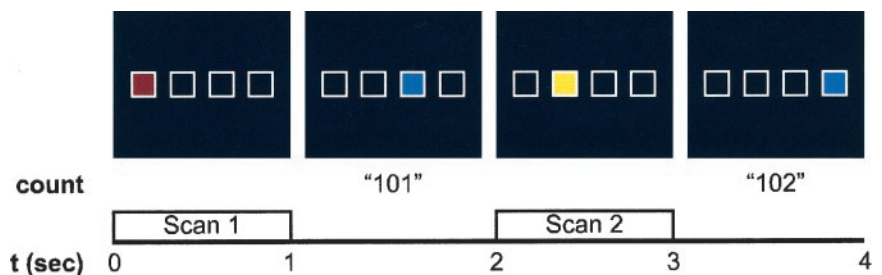


Table 1. Markov sequence conditional probability matrix

X_{i+1}	1	2	3	4
X_i				
1	0	0.15	0.7	0.15
2	0.15	0	0.15	0.7
3	0.15	0.7	0	0.15
4	0.7	0.15	0.15	0

The matrix defines the probability relationships among the different elements of the conditional sequence for a Markov chain with entropy = 0.92 (where X_i represents the spatial position at trial i , and X_{i+1} the spatial position at the next trial).

The experimental design was separated into two studies. In the first study, we were interested in the main effects of both predictability and awareness as well as the interaction between them. In other words, which brain regions were correlated with sequential predictability, and how did explicit knowledge of the sequence change this pattern? Subjects participating in the first study were subdivided into two groups. The first group was instructed simply to keep a mental count of the blue boxes and did not have any practice before the functional session. They were not informed of the possible existence of a spatial sequence. After the session, none of these individuals reported awareness of the underlying spatial sequence, and they are therefore designated the UN-AWARE group [$n = 14$; 4 males (M), 10 females (F)]. The second group performed the identical task, except that the zero-entropy spatial sequence was explicitly shown to them before the scanning session. To ensure that they fully encoded the spatial sequence, they also received a 2 min practice session on the zero-entropy condition. Because they were shown the spatial sequence before performing the task, they were called the AWARE group ($n = 12$; 3 M, 9 F).

In the second study, we were interested in examining the time course of acquisition of sequential predictability, i.e., a time by condition interaction. Again, participants in this study were divided into two groups. Unlike the previous study, subjects in the first group received no underlying spatial sequence throughout the session; the spatial target presentation was completely random (entropy = 2). They are therefore referred to as the RANDOM group ($n = 5$; 2 M, 3 F). The second group, called SEQUENCE ($n = 5$; 2 M, 3 F), performed the zero-entropy condition for the entire session, but without any prepractice, nor were they informed of the existence of a spatial sequence (entropy = 0). Like the UN-AWARE group, these subjects did not show evidence of explicit awareness when debriefed after the session.

Entropy. We define temporal uncertainty by the conditional entropy (Shannon and Weaver, 1949; Cover and Thomas, 1991). Using a Markov chain: $H(X_{i+1} | X_i) = -\sum p(x_i) \sum p(x_{i+1} | x_i) \log_2 p(x_{i+1} | x_i)$, where $H(X_{i+1} | X_i)$ is the first-order conditional entropy, $p(x_i)$ is the probability of event x_i occurring (e.g., which spatial position is chosen), and $p(x_{i+1} | x_i)$ is the probability of x_{i+1} , given that x_i occurs previously. Predictability, I , is defined by the mutual information theorem: $I(X_{i+1} | X_i) = H(X) - H(X_{i+1} | X_i)$, and represents the decrement in uncertainty provided by the preceding stimulus. Using the entropic measure, the three conditions were (1) the full sequence, 1-3-2-4... ($H = 0$); (2) a Markov sequence with the conditional probability matrix ($H = 0.92$), as outlined in Table 1; and (3) a fully random sequence ($H = 2.0$). Each condition was maintained for a block of 90 stimulus presentations (45 scans) before switching to the next condition block (Fig. 2). The task proceeded continuously without any breaks or notations between the conditions. Three repetitions of each block were given during a single 13.5 min study

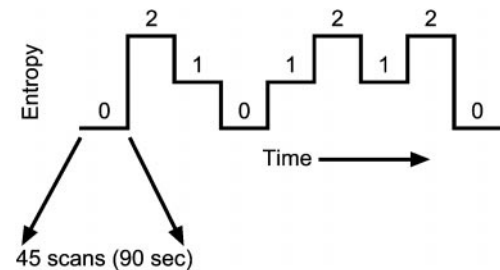


Figure 2. Overall block design of the functional scan session. Each of the three entropy levels (0, 1, and 2) were presented three times during the course of the functional scan. Each entropy block lasted for 45 scans (90 sec).

session, and the block order always began and ended with the zero-entropy condition. The remaining conditions were counterbalanced across time both within and between individuals by reversing the condition order for some participants. In the forward order, the different entropy levels were ordered as 0-2-1-0-1-2-1-2-0 (Fig. 2). The reverse condition order was 0-2-1-2-1-0-1-2-0. This variation in conditional presentation was used to reduce the possibility of confounding linear temporal effects with the experimental design. Participants in the UN-AWARE and AWARE groups were randomly assigned to receive either the forward conditional order (UN-AWARE, $n = 9$; AWARE, $n = 6$) or the reverse conditional order (UN-AWARE, $n = 5$; AWARE, $n = 6$).

The RANDOM group served as another control for possible unknown nonlinear temporal effects that might be inadvertently confounded with the task design and not handled by the counterbalancing. The SEQUENCE group was similar to the RANDOM group in that no variation between blocks occurred and was used to examine the time course of learning. The protocol was approved by the Emory Human Investigations Committee.

MRI. A single fMRI session with 405 scans was obtained during the continuous performance of the task. Functional MRI was performed with gradient-recalled echoplanar imaging [repetition time (TR) = 2000 msec; echo time (TE) = 40 msec; flip angle = 90°; 64 × 64 matrix; 10 × 8 mm contiguous axial slices] on a Philips 1.5 T scanner (Kwong et al., 1992; Ogawa et al., 1992). Structural, T1-weighted MRIs were obtained for subsequent spatial normalization (spin-echo, TR = 500 msec; TE = 20 msec; flip angle = 90°; 256 × 256 matrix; 24 × 5 mm contiguous axial slices).

Statistical analysis. The data were analyzed on a voxel-by-voxel basis using an ANOVA with both conditional entropy (0, 1, 2) and group (UN-AWARE, AWARE, RANDOM) as main effects. Data were analyzed using Statistical Parametric Mapping (SPM99b; Wellcome Department of Cognitive Neurology, London, UK) (Friston et al., 1995). Motion correction to the first functional scan was performed within subject using a six-parameter rigid-body transformation. The mean of the motion-corrected images was then coregistered to the individual's 24-slice structural MRI, using a 12-parameter affine transformation. Spatial normalization to the Montreal Neurological Institute template (Talairach and Tournoux, 1988) was performed by applying a 12-parameter affine transformation followed by a nonlinear warping using basis functions (Ashburner and Friston, 1999). All transformations were computed sequentially with one reslice operation at the end. The spatially normalized scans were smoothed with an 8 mm isotropic Gaussian kernel to accommodate anatomical differences across subjects.

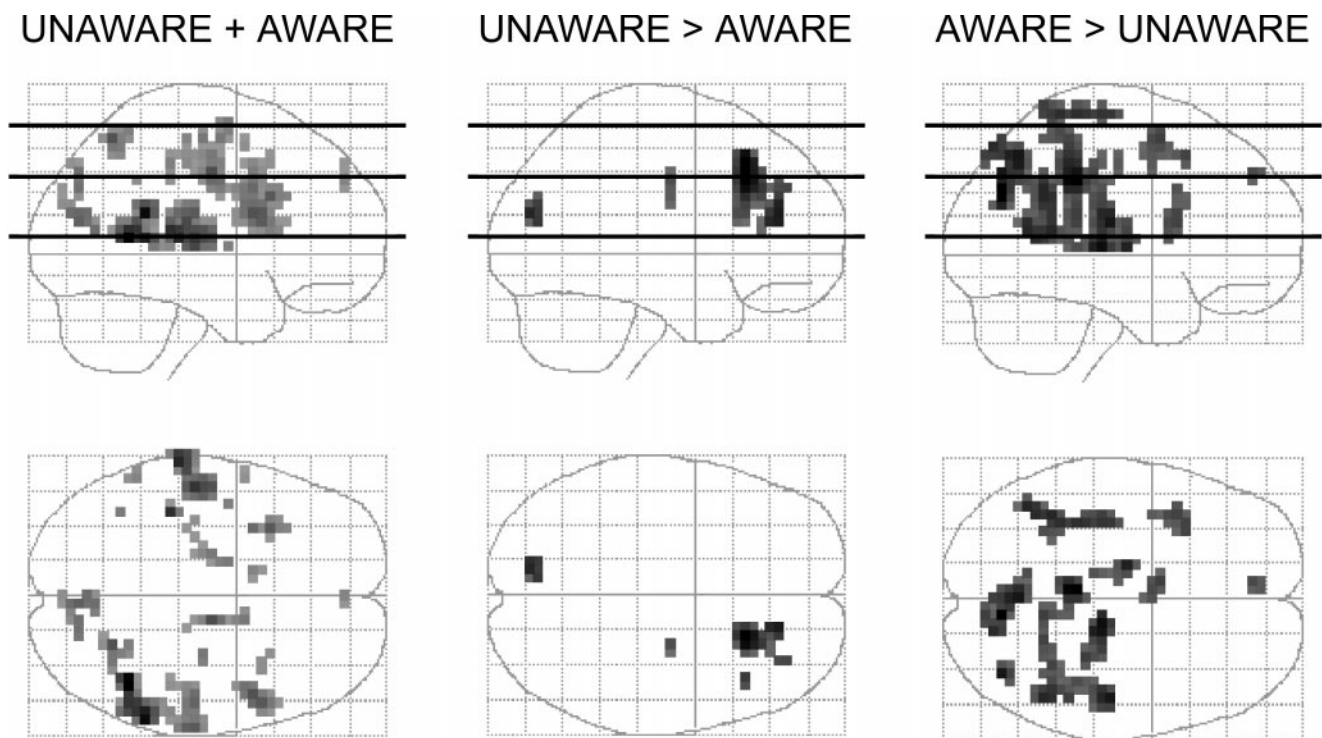


Figure 3. Sagittal and transverse projections of the statistical comparison for the UNAWARE + AWARE, UNAWARE > AWARE, and AWARE > UNAWARE groups across all three entropy conditions. UNAWARE + AWARE shows regions common to both of these groups that positively correlated with entropy. The other two columns show the interaction of awareness and entropy. Horizontal lines in the sagittal images indicate planes in Figure 4. All groups were analyzed with a significance threshold of $p < 0.01$ (uncorrected).

A random-effects model was used to make statistical inferences (Friston et al., 1999). This was done by first high-pass filtering each time series (cutoff = 500 sec) and then computing three adjusted mean images for each participant, one per entropy level, using a subjectwise ANCOVA to remove any global signal intensity differences. Although the RANDOM and SEQUENCE groups contained only one entropy level throughout the entire scan, we assigned dummy entropy levels to blocks of 45 scans in the same order as shown in Figure 2 when calculating the three adjusted mean images. Thus, the RANDOM and SEQUENCE groups were treated temporally as having been presented with the forward order (0-2-1-0-1-2-1-2-0) and their adjusted mean images were calculated accordingly. This allowed for direct comparison to both the UNAWARE and AWARE groups. A multigroup design matrix was specified with three adjusted mean images per subject using the subject groups AWARE, UNAWARE, and RANDOM. Linear contrasts between the two extreme entropy levels were examined for each group, with a threshold for significance of $p < 0.01$ (uncorrected for multiple comparisons). This relatively liberal threshold was used because of the decreased statistical power in a random-effects design. Using SPM99b, we implemented this using the contrast vector, $[-1, 0, 1]$, which corresponded to the ordered entropy levels, $[0, 1, 2]$. This was done separately for both the UNAWARE and AWARE groups, and the conjunction of these contrasts identified the brain regions with a significant positive correlation to entropy in both groups (Price and Friston, 1997). The interaction of awareness X entropy was assessed by the two contrast vectors: $[-1, 0, 1, 1, 0, -1]$ (UNAWARE > AWARE) and $[1, 0, -1, -1, 0, 1]$ (AWARE > UNAWARE).

RESULTS

When debriefed, subjects reported a mean count of blue boxes of 259 ± 5 (SE). This represented a mean accuracy of $96 \pm 2\%$ (SE), indicating that subjects maintained sufficient attention throughout the scan session to perform adequately.

Both the UNAWARE and the AWARE groups displayed activation in Wernicke's area and its right homolog that correlated

with the entropy of the underlying spatial sequence (Figs. 3, 4; Table 2). The conjunction between these two groups identified areas of activation associated with a common process (UNAWARE + AWARE). The RANDOM control group did not show evidence of such a relationship (Fig. 4, bottom row). The relationship of entropy to adjusted activation in these regions was nonlinear, with the mid- and high-entropy conditions having similar levels of activation but different from the zero-entropy condition. The magnitude of this relationship was apparently greater in the AWARE group in the right posterior temporal cortex, although the group X entropy interaction was not statistically significant there.

In addition to performing a conjunction analysis, we also examined the interactions between the AWARE and UNAWARE groups; this was used to determine which activations were dependent on awareness versus unawareness. In Figure 4, the UNAWARE > AWARE interaction revealed an activation change within right prefrontal cortex (RPFC). In examining the relationship of adjusted BOLD response to entropy in right PFC, we found that the AWARE group had more activation in the zero-entropy condition, i.e., when a deterministic sequence was present. Neither the UNAWARE nor RANDOM groups displayed this effect (Fig. 4, middle row). Thus, the AWARE group displayed a negative correlation to entropy in RPFC, whereas both the UNAWARE and RANDOM groups did not. The AWARE > UNAWARE interaction indicated significant effects in many regions but with a particularly prominent one in right posterior parietal cortex (Table 2).

Because fMRI measurements are relative to each other, there

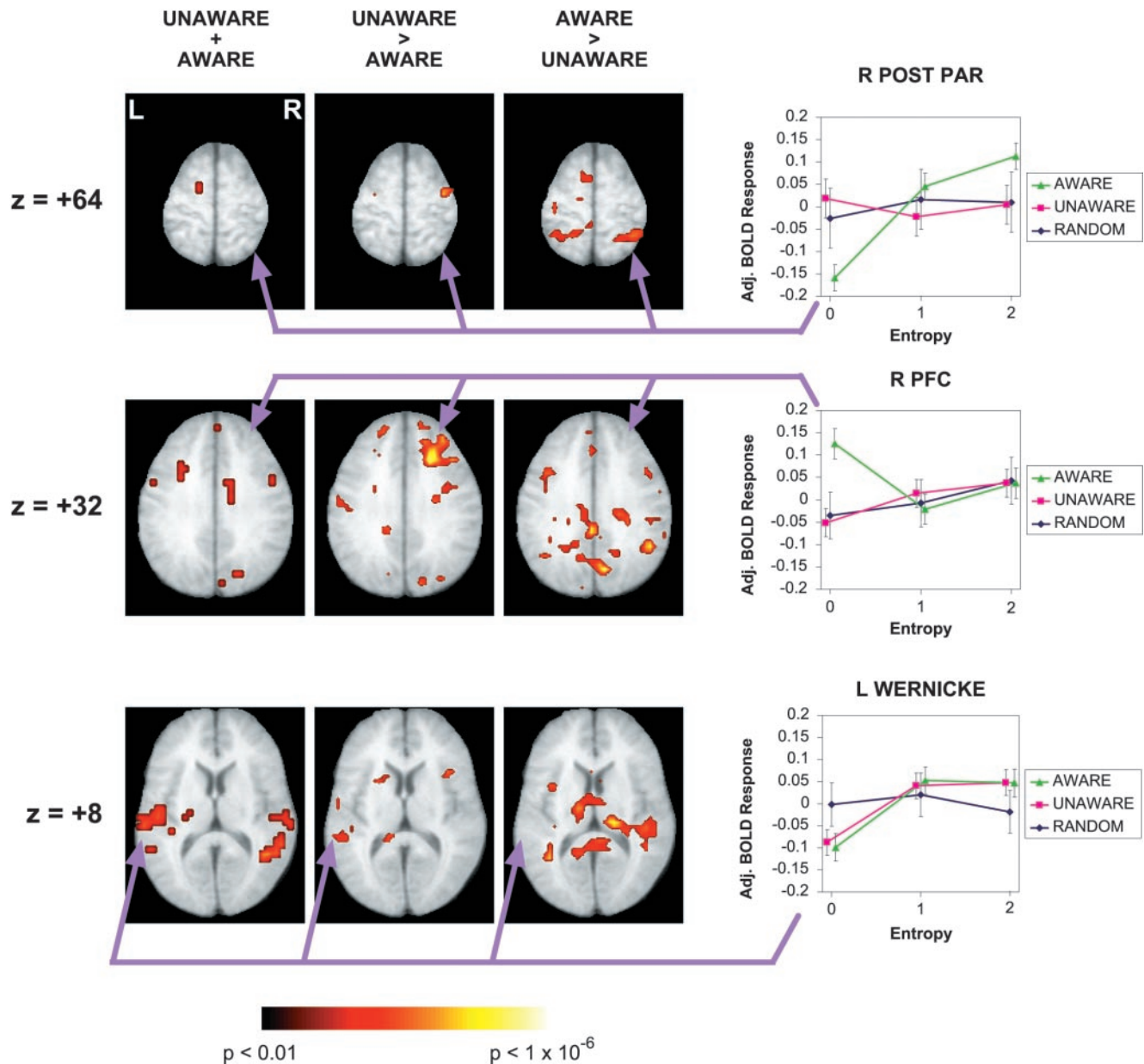


Figure 4. Statistical parametric maps showing regions that correlated with conditional entropy. Transverse planes (Talairach, $Z = +64$, $Z = +32$, and $Z = +8$) for both the conjunction (UNWARE + AWARE) and the interactions (UNWARE > AWARE, AWARE > UNWARE) between the groups. The relationship of the adjusted BOLD response to entropy is shown for selected regions in the AWARE, UNAWARE, and RANDOM groups (right). Wernicke's area (coordinates: $-52, -20$ to $-40, 4-8$) showed a significant correlation to entropy independent of awareness. The right posterior parietal cortex displayed a significant positive entropy relationship only in the AWARE group (coordinates: $28, -48, 64$). The right prefrontal cortex displayed a negative relationship to entropy in the AWARE group (coordinates: $20, 24, 32$). The RANDOM did not display any of these relationships.

are two possible interpretations of these results: increased conditional entropy is associated with increased activation of these regions, or increased predictability (the converse of entropy) is associated with decreased activation. Examination of the repetition \times entropy interaction in the UNAWARE group suggested that Wernicke's activation in the zero-entropy condition decreased with each repetition more quickly than either the mid- or high-entropy conditions. To verify this, we compared the monotonic temporal drifts of this region in both the RANDOM and the SEQUENCE groups. Both Wernicke's area and its right homolog displayed significant decreases in activation with time, but only in the SEQUENCE group (Fig. 5). This suggested that acquisition of predictability was associated with decreased activation

in these regions, not increased activation in response to entropy.

DISCUSSION

Temporal predictability appears in a variety of contexts, so if Wernicke's area is associated with a generic predictability function, then there should be similar findings across the imaging literature. In a study of language complexity, Just et al. (1996) reported an increase in voxel recruitment in Wernicke's area as the complexity of visually presented sentences increased, but their results could also be interpreted as a decrease in activation with sentence predictability. Similarly, deviant syllables and tones have been associated with left posterior superior temporal acti-

Table 2. Regional activations associated with comparisons of task performance

Region	Cluster (voxels)	MNI coordinates (x,y,z)			Z score
UNAWARE + AWARE					
Right middle temporal gyrus (BA22)	49	44	-52	8	3.45
Left superior temporal gyrus (BA22)*	56	-64	-28	8	2.81
Right inferior frontal gyrus (BA44)	42	48	8	20	2.23
Left inferior frontal gyrus (BA44)	20	-32	16	28	2.20
Right postcentral gyrus (BA40/43)	36	64	-20	20	2.11
UNAWARE > AWARE					
Right cingulate/prefrontal gyrus (BA8/32)	86	20	24	44	4.33
AWARE > UNAWARE					
Right parieto-occipital sulcus (BA19)	63	8	-72	32	4.16
Left cingulate gyrus (BA31)	58	-4	-36	36	3.98
Right pulvinar thalamus	55	12	-24	4	3.84
Right lateral sulcus (BA40)	34	48	-20	16	3.65
Left superior temporal gyrus (BA22)	21	-40	-48	12	3.45
Right supramarginal gyrus (BA40)	36	48	-52	32	3.43
Left postcentral gyrus (BA3/4)	22	-36	-36	68	3.30
Left frontal medial gyrus (BA6)	22	-8	4	56	3.05
Right posterior parietal lobe (BA40/7)	30	48	-40	52	3.05

Coordinates are from the Montreal Neurological Institute (MNI) template and use the same orientation and origin as found in the Talairach atlas (Talairach and Tournoux, 1988). The table shows at most three maxima > 8.0 mm apart per cluster ($k \geq 20$).

*Wernicke's area.

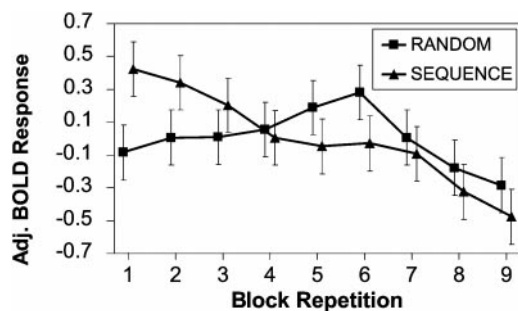


Figure 5. Differing temporal effects between the RANDOM group ($n = 5$) and the SEQUENCE group ($n = 5$) in Wernicke's area (coordinates: -56, -28, 12). Activation decreased as the grammar statistics were acquired in the SEQUENCE group, but not in the RANDOM group (ANCOVA, $p = 0.004$).

vation (Celsis et al., 1999). In nonhuman primates, neurons in the superior temporal cortex show a graded response to vocalization complexity, and lesions impair animals' ability to discriminate both their coos and other auditory patterns (Heffner and Heffner, 1984; Javitt et al., 1994; Rauschecker et al., 1995; Colombo et al., 1996). Several motor sequencing studies have reported results consistent with ours. Typical sequencing studies may involve a comparison of a sequential to a random task. With the inclusion of an attentional distractor, positron emission tomography studies of motor learning have shown decreased regional cerebral blood flow (rCBF) in the left middle temporal and left inferior temporal areas as subjects learned implicitly (Grafton et al., 1995; Hazeltine et al., 1997). Interestingly, these areas were not involved in the absence of a distractor, although a decreasing rCBF was seen in the left superior temporal gyrus as task familiarity increased. In our study, both the UNAWARE and the AWARE subject groups showed a decreased response to predictability in Wernicke's area, with the effect being greater in the AWARE subjects than in the UNAWARE subjects. It is possible

that this activity may reflect a subvocalization process attributable to the mental counting, but since this was constant throughout the task, the spatial predictability must have at least modulated the activity in this region. We suggest that Wernicke's area may not be restricted to language aspects per se, but rather is responsive to probabilistic features in time.

The observation that decreased Wernicke's activation occurred in both the UNAWARE and AWARE groups suggests the existence of a process independent of awareness in this region. The conjunction analysis showed that Wernicke's and its right homolog were the only areas common to both groups showing a parametric relationship to entropy. Because the UNAWARE group did not receive any prepractice, we can assume that the acquisition of the grammar statistics occurred during the scan session. In contrast, the AWARE group had both explicit knowledge and practice with the zero-entropy grammar before the scan session and consequently should have had significantly less learning during the scan. Although the UNAWARE group had a significant correlation in Wernicke's area to entropy, the magnitude of this relationship was slightly less than that of the AWARE group. This may be attributable to the fact that implicit acquisition results in less activity change than explicit, or it may be attributable to the fact that the unaware group began the scan in an undifferentiated state. Averaged over the course of the scan, the effect would then appear smaller. The fact that this process can occur implicitly is consistent with evidence that language acquisition is also an implicit process (Pinker, 1994; Elman et al., 1996; Morgan and Demuth, 1996; Saffran et al., 1996; Marcus et al., 1999), and Wernicke's area has long been associated with language syntax.

The AWARE group showed significantly more regions correlated with entropy than the UNAWARE group, and these extra regions corresponded to the known attentional systems in the human brain. The posterior parietal system has classically been associated with spatial attention (Posner and Petersen, 1990; Pardo et al., 1991; Corbetta et al., 1993; Ungerleider and Haxby,

1994; Corbetta et al., 1995; Courtney et al., 1996; Le et al., 1998; Carpenter et al., 1999; Rosen et al., 1999). Posterior parietal correlation with entropy in the AWARE group implies that increased entropy was associated with increased attention for spatial position. This most likely represents a “top-down” effect because it did not occur in the UNAWARE group. Both functionally and anatomically distinct neural systems for implicit and explicit learning have been suggested (Squire, 1987), but because we did not scan subcortical regions, we could not evaluate these dissociations in this study.

Right prefrontal cortex has been implicated in both grammar learning and working memory, particularly spatial working memory. Our observation that awareness of the spatial sequence was associated with increased activity in the zero-entropy condition is consistent with the hypothesis that these subjects were actively using working memory, but only when the sequence was recognizable as such. This is consistent with several previous imaging studies of working memory (Jonides et al., 1993; Petrides et al., 1993; McCarthy et al., 1994; Swartz et al., 1995; Courtney et al., 1996; Cohen et al., 1997) as well as a putative specialization for spatial working memory on the right (Smith et al., 1996) and physiological data suggestive of neuronal storage (Goldman-Rakic, 1987; Miller et al., 1996; Fuster, 1997). Our findings are also consistent with a suggested role for right PFC in the explicit learning of individual items in a sequence (Fletcher et al., 1999), which would only be observable in the AWARE group.

Our results suggest a role more general than language processing for Wernicke's area, and consequently, an expanded interpretation of the neurological basis of language. Although there is ample evidence that Wernicke's is intimately related to both spoken and written language (Petersen et al., 1988; Fiez et al., 1996; Fiez and Petersen, 1998), whether language acquisition is rule-based or probabilistic is debatable (Chomsky, 1957; Pinker, 1994; Elman et al., 1996; Morgan and Demuth, 1996; Saffran et al., 1996; Seidenberg, 1997; Marcus et al., 1999), and the close anatomical relationship of this region to visuospatial processing has been noted (Serenio et al., 1995). Infants have been shown to segment nonwords based on statistical relationships between sounds (Saffran et al., 1996), and our results suggest that a grammar need not be language-specific, only probabilistic, to involve Wernicke's area. Critically, this can occur implicitly, which is a requirement for language acquisition. In dyslexia, where both the phonological and visual systems appear disrupted, Wernicke's area fails to systematically increase its activity as the difficulty of a reading judgement task is increased (Wagner and Torgesen, 1987; Stanovich, 1988; Eden et al., 1996; Shaywitz et al., 1998). If Wernicke's area performs a generic predictability function, then dysfunction would be consistent with this observation. The richness of human language likely arises from a combination of bottom-up statistics and top-down syntactical rules. From both a developmental and evolutionary perspective, processing of temporal predictability in Wernicke's area would seem to be a logical starting point to acquire language.

REFERENCES

- Ashburner J, Friston KJ (1999) Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7:254–266.
- Carpenter PA, Just MA, Keller TA, Eddy W, Thulborn K (1999) Graded functional activation in the visuospatial system with the amount of task demand. *J Cogn Neurosci* 11:9–24.
- Celsis P, Boulanour K, Doyon B, Ranjeva JP, Berry I, Nespoulous JL, Chollet F (1999) Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage* 9:135–144.
- Chomsky N (1957) *Syntactic structures*. The Hague: Mouton.
- Cleeremans A, McClelland JL (1991) Learning the structure of event sequences. *J Exp Psychol Gen* 120:235–253.
- Cohen A, Ivry RI, Keele SW (1990) Attention and structure in sequence learning. *J Exp Psychol Learn Mem Cogn* 16:17–30.
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–608.
- Colombo M, Rodman HR, Gross CG (1996) The effects of superior temporal cortex lesions on the processing and retention of auditory information in monkeys (*Cebus apella*). *J Neurosci* 16:4501–4517.
- Corbetta M, Miezin FM, Shulman GL, Petersen SE (1993) A PET study of visuospatial attention. *J Neurosci* 13:1202–1226.
- Corbetta M, Shulman GL, Miezin FM, Petersen SE (1995) Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science* 270:802–805.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1996) Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb Cortex* 6:39–49.
- Cover TM, Thomas JA (1991) *Elements of information theory*. New York: Wiley.
- Eden GF, VanMeter JW, Rumsey JM, Maisog JM, Woods RP, Zeffiro TA (1996) Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature* 382:66–69.
- Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K (1996) *Rethinking innateness. A connectionist perspective on development*. Cambridge: MIT.
- Fiez JA, Petersen SE (1998) Neuroimaging studies of word reading. *Proc Natl Acad Sci USA* 95:914–921.
- Fiez JA, Raichle ME, Balota DA, Tallal P, Petersen SE (1996) PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb Cortex* 6:1–10.
- Fletcher P, Buchel C, Josephs O, Friston K, Dolan R (1999) Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. *Cereb Cortex* 9:168–178.
- Friston KJ, Holmes AP, Worsley KJ, Poline J-B, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210.
- Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? *NeuroImage* 10:1–5.
- Fuster JM (1997) *The prefrontal cortex, anatomy, physiology, and neuropsychology*. New York: Raven.
- Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: *Handbook of physiology, the nervous system, higher functions of the brain* (Plum F, ed), pp 373–417. Bethesda: MD: American Physiological Society.
- Grafton ST, Hazeltine E, Ivry R (1995) Functional mapping of sequence learning in normal humans. *J Cogn Neurosci* 7:497–510.
- Hazeltine E, Grafton ST, Ivry R (1997) Attention and stimulus characteristics determine the locus of motor- sequence encoding. A PET study. *Brain* 120:123–140.
- Heffner HE, Heffner RS (1984) Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science* 226:75–76.
- Javitt DC, Steinschneider M, Schroeder CE, Vaughan Jr HG, Arezzo JC (1994) Detection of stimulus deviance within primate primary auditory cortex: intracortical mechanisms of mismatch negativity (MMN) generation. *Brain Res* 667:192–200.
- Jonides J, Smith EE, Koeppe RA, Awh E, Minoshima S, Mintun MA (1993) Spatial working memory in humans as revealed by PET. *Nature* 363:623–625.
- Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996) Brain activation modulated by sentence comprehension. *Science* 274:114–116.
- Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R, Cheng HM, Brady TJ, Rosen BR (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA* 89:5675–5679.
- Le TH, Pardo JV, Hu X (1998) 4 T-fMRI study of nonspatial shifting of selective attention: cerebellar and parietal contributions. *J Neurophysiol* 79:1535–1548.
- Marcus GF, Vijayan S, Bandi Rao S, Vishton PM (1999) Rule learning by seven-month-old infants. *Science* 283:77–80.

- McCarthy G, Blamire AM, Puce A, Nobre AC, Bloch G, Hyder F, Goldman-Rakic P, Shulman RG (1994) Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. *Proc Natl Acad Sci USA* 91:8690–8694.
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16:5154–5167.
- Morgan JL, Demuth K (1996) Signal to syntax: bootstrapping from speech to grammar in early acquisition. Mahwah, NJ: Lawrence Erlbaum.
- Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K (1992) Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci USA* 89:5951–5955.
- Pardo JV, Fox PT, Raichle ME (1991) Localization of a human system for sustained attention by positron emission tomography. *Nature* 349:61–64.
- Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331:585–589.
- Petrides M, Alivisatos B, Meyer E, Evans AC (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci USA* 90:878–882.
- Pinker S (1994) *The language instinct*. New York, NY: W. Morrow.
- Posner MI, Petersen SE (1990) The attention system of the human brain. *Annu Rev Neurosci* 13:25–42.
- Price CJ, Friston KJ (1997) Cognitive conjunction: a new approach to brain activation experiments. *NeuroImage* 5:261–270.
- Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.
- Reber AS (1967) Implicit learning of artificial grammars. *J Verbal Learning Verbal Behav* 6:317–327.
- Reber AS (1993) *Implicit learning and tacit knowledge. An essay on the cognitive unconscious*. Oxford: Oxford UP.
- Rosen AC, Rao SM, Caffarra P, Scaglioni A, Bobholz JA, Woodley SJ, Hammeke TA, Cunningham JM, Prieto TE, Binder JR (1999) Neural Basis of endogenous and exogenous spatial orienting. A functional MRI study. *J Cogn Neurosci* 11:135–152.
- Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.
- Seidenberg MS (1997) Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275:1599–1603.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Shannon CE, Weaver W (1949) *The mathematical theory of communication*. Chicago: University of Illinois Press.
- Shaywitz SE, Shaywitz BA, Pugh KR, Fulbright RK, Constable RT, Mencl WE, Shankweiler DP, Liberman AM, Skudlarski P, Fletcher JM, Katz L, Marchione KE, Lacadie C, Gatenby C, Gore JC (1998) Functional disruption in the organization of the brain for reading in dyslexia. *Proc Natl Acad Sci USA* 95:2636–2641.
- Smith EE, Jonides J, Koeppel RA (1996) Dissociating verbal and spatial working memory using PET. *Cereb Cortex* 6:11–20.
- Squire LR (1987) *Memory and brain*. New York: Oxford UP.
- Stadler MA (1989) On learning complex procedural knowledge. *J Exp Psychol Learn Mem Cogn* 15:1061–1069.
- Stanovich KE (1988) Explaining the differences between the dyslexic and the garden-variety poor reader: the phonological-core variable-difference model. *J Learn Disabil* 21:590–604.
- Swartz BE, Halgren E, Fuster JM, Simpkins E, Gee M, Mandelkern M (1995) Cortical metabolic activation in humans during a visual memory task. *Cereb Cortex* 5:205–214.
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the brain*. New York: Thieme.
- Ungerleider LG, Haxby JV (1994) “What” and “where” in the human brain. *Curr Opin Neurobiol* 4:157–165.
- Wagner RK, Torgesen JK (1987) The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol Bull* 101:192–212.