NOTES AND DISCUSSION

Functional MR Imaging during Auditory Word Perception: A Single-Trial Presentation Paradigm

Gregory Hickok

University of California, Irvine

and

Tracy Love, David Swinney, Eric C. Wong, and Richard B. Buxton

University of California, San Diego

Functional magnetic resonance imaging (fMRI), while still in relatively early stages of its development, has become a widespread tool for addressing issues in the neurobiology of language and other cognitive domains (e.g., Belliveau et al., 1991; Binder et al., 1994; Cohen et al., 1994; Hickok et al., 1995; Shaywitz et al., 1995; for a review of the technique see Bandettini et al., 1995). In a typical fMRI experiment, stimuli are presented in blocks of say 20–40 sec that alternate through several cycles with an equally long baseline/control period. So for example, a subject may see individual words presented one-at-a-time for 30 sec, then non-words for 30 sec, and so on for five cycles. This presentation scheme provides a serious constraint on experimental design in fMRI studies and consequently limits the kinds of questions one can address with this methodology. One cannot, to cite just one example, use lexical decision as a behavioral measure in the magnet because the item type (word/non-word) is fully predictable from the experimental design. Relatedly, under these experimental conditions it is unclear to what extent fMRI activations are due to expectation-driven (controlled) processes versus lower-level perceptual (automatic) processes.

Here we present data from an alternative fMRI presentation scheme used with auditory speech perception. The paradigm is modeled on standard methodologies in event related potential (ERP) studies where test items are ran-

domly intermixed with control items and the signal is selectively averaged over multiple presentations of individual items in one or the other condition. Because the fMRI signal is dependent on hemodynamic changes rather than electrical potential changes, and because hemodynamic changes are sluggish (on the order of seconds), the interstimulus interval in the fMRI version has to be significantly longer than in the ERP version, but otherwise the presentation schemes are identical. Previous work showing that brief (less than a second) ''bursts'' of motor activity (Bandettini et al., 1995) and visual stimulation (Savoy et al., 1995) can yield a detectable signal suggested that such a presentation scheme may be feasible.

The subject was a right-handed woman age 23 who volunteered her participation. Her native language is English. The stimuli consisted of 25 mono- and bisyllabic concrete nouns (e.g., *doctor, tree, desk;* mean frequency = 67.96 (Francis & Kucera, 1982), range 5–224) randomly intermixed with 25 non-words. The non-words were created by digitally reversing each of the 25 nouns and therefore were acoustically matched with the word stimuli in terms of duration, amplitude, and frequency spectrum. The non-words were not recognizable as modified English words. Stimulus duration averaged 536 msec. Auditory stimuli were presented binaurally via air conduction through plastic tubing connected to insert earplugs. We presented one item every 12 sec for a total study duration of 10 min. The 12-sec inter-trial interval was selected on the basis of previous observations of the time period required for the signal to return to baseline after the stimulus offset (Bandettini et al., 1995). The subject was instructed to listen and ''pay attention'' to the items; no explicit task was required. Acclimation to the MR scanner was achieved during several prior runs (part of another experiment) and during the acquisition of structural images. Imaging was carried out on a GE Signa 1.5T system fitted with a three axis local head gradient coil (Wong, Bandettini, & Hyde, 1992) using gradient-echo, echo planar imaging (EPI) (TR = 1000 ms, TE = 40, FOV = 24). Two 10-mm axial slices covering portions of the lateral fissure and superior temporal sulcus were imaged. In-plane resolution was 3.75×3.75 mm. We selected the six most activated pixels from each of four regions of interest (ROI): left primary auditory cortex, right primary auditory cortex, left posterior superior temporal sulcus, and right posterior superior temporal sulcus. No negatively correlated pixels were identified in these ROIs. Pixels from a given ROI were contiguous so that each ROI represents a single blob of activation. These regions were identified using (i) 3D surface reconstructions (Damasio & Frank, 1992) from co-registered high resolution MR images (T1-weighted, 3D Volume SPGR, slice thickness = 2 mm, FOV = 24) and (ii) anatomical detail on the EPI images themselves. Activated pixels were identified by correlating signal changes (collapsed over the two stimulus conditions) with a reference waveform (Bandettini, Jesmanowicz, Wong, & Hyde, 1993); the correlation coefficient was defined as our measure of activation. Those pixels with the highest correlation within a given ROI also had the greatest percentage signal
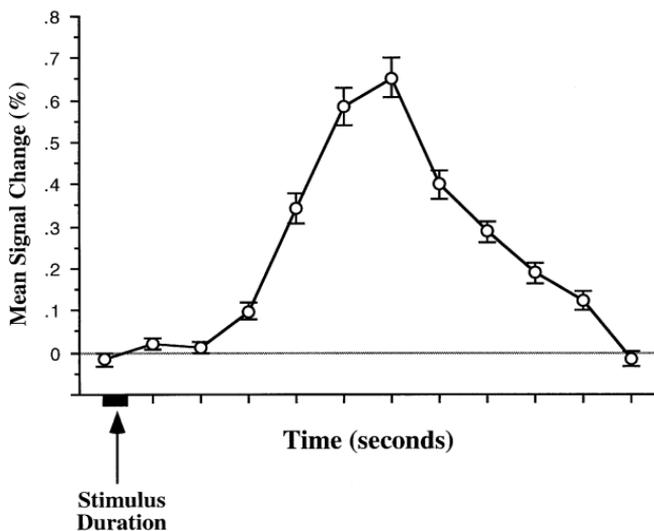
FIG. 1. Mean signal change as a function of time averaged over the two stimulus types and the four regions. Bars represent standard error.

change values. The resulting dataset was then divided into 12-sec epochs separated by condition. Signal values were converted to reflect percentage change from baseline where baseline was defined as the mean value of image numbers 1, 2, and 12 in each epoch.

The results were clear: Looking at the data collapsed over conditions, we noted a very robust signal change beginning 3 sec post stimulus onset and peaking 5–6 sec post onset (see Fig. 1)—a time-course consistent with previous observations of the hemodynamic response (e.g., Bandettini et al., 1995). (It is worth noting that while the MR signal change is significantly delayed relative to the onset of the underlying neural response, there appears to be a relatively constant temporal relation between these two events (Savoy et al., 1995), which in turn renders the time-course of the hemodynamic response a neuroscientifically significant dependent variable.) Maximum average signal change was between 0.6 and 0.7%. There was no *overall* difference between the word and non-word conditions, but stimulus-type did interact with brain region where the mean response difference between word and non-word stimuli (averaged over the entire 12-sec epoch) went in opposite directions in primary auditory regions versus superior temporal sulcus regions (see Fig. 2a). This effect is seen more clearly if we look at the time-course of this signal change. Figure 2b shows the response curves for words versus non-words in primary auditory and superior temporal sulcus regions. The major difference in the response to words versus non-words in primary auditory areas shows up on the ramp-down with non-words producing a more gradual return to baseline, whereas the major difference in superior temporal sulcus
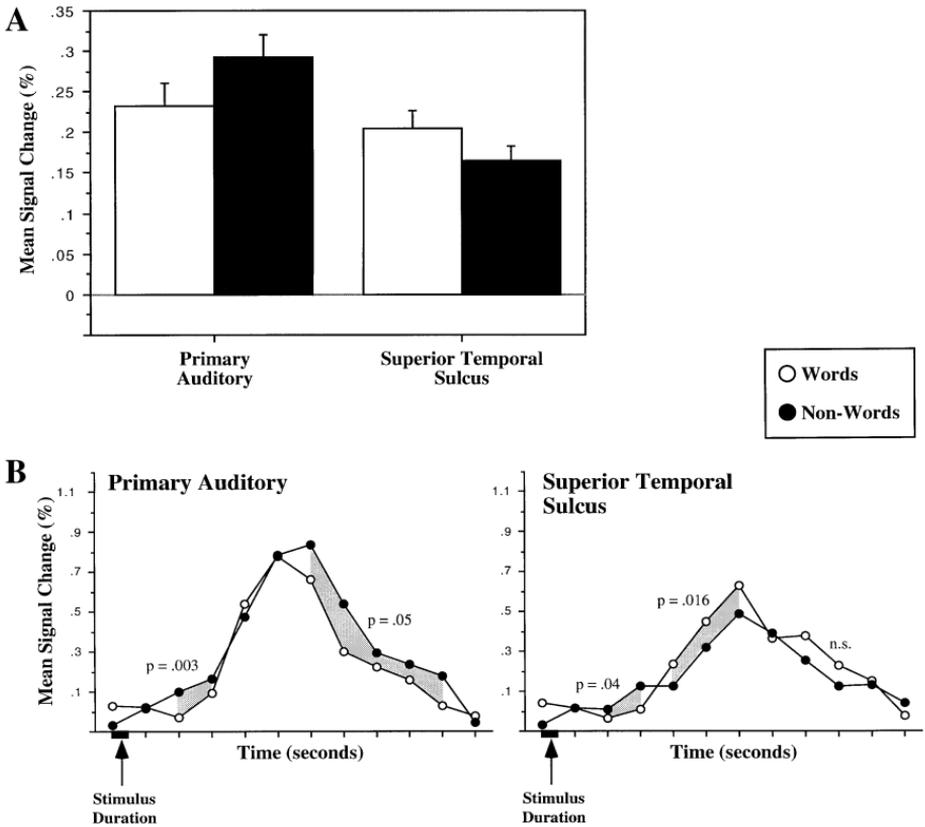
FIG. 2. Interaction between ROI and stimulus type (A) in terms of signal change averaged over the entire 12-sec epoch and (B) as a function of time. Shaded areas in (B) represent regions in the response curve that differ significantly between stimulus conditions. p-Values correspond to two-tailed t tests calculated over the group of values in each shaded region.

areas shows up on the ramp-up with words producing a steeper response and a higher peak.

We also observed very robust differences between regions and between hemispheres, but it is difficult to rule out the possibility that these effects are simply due to regional differences in flow patterns. For this reason we will simply summarize these results here. Looking at the data collapsed across word/non-word conditions, in primary auditory areas the response was faster (the ramp-up occurred earlier) than in superior temporal regions. The peak amplitude of the response was similar across all regions except in the left primary auditory region which was greater by 0.4 percentage points. Whether these differences turn out to be related to differences in the neural response or an artifact due to regional differences in flow patterns will have to await further work.

In sum, the points we would like to make are that (i) reliable changes in

the fMRI signal can be detected in response to individually presented speech stimuli averaging over as few as 25 items, (ii) this measurement technique is, at least potentially, sensitive to differences in the neural response to different types of speech stimuli which can show up in terms of the amplitude and/or latency of the response, and (iii) the nature of the differences that one can observe are such that they would likely be missed if the traditional blocked-design fMRI experiment were used. Finally, the fact that stimuli can be presented individually in a randomized format rather than in the standard blocked format will allow for more flexibility in experimental design in functional MR imaging.

REFERENCES

Bandettini, P. A., Jesmanowicz, A., Wong, E. C., & Hyde, J. S. 1993. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine,* **30,** 161–173.

Bandettini, P. A., Wong, E. C., Binder, J. R., Rao, S. M., Jesmanowicz, A., Aaron, E. A., Lowry, T. F., Forster, H. V., Hinks, R. S., & Hyde, J. S. 1995. Functional MR imaging using the BOLD approach: Dynamic characteristics and data analysis methods. In D. L. Bihan (Ed.), *Diffusion and perfusion magnetic resonance imaging.* New York: Raven Press.

Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J., & Rosen, B. R. 1991. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science,* **254,** 716–719.

Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., Estkowski, L. D., Goldstein, M. D., Haughton, V. M., & Hyde, J. S. 1994. Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology,* **35,** 662–672.

Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., & Noll, D. C. 1994. Activation of the prefrontal cortex in a nonspatial working memory tasks with functional MRI. *Human Brain Mapping,* **1,** 293–304.

Damasio, H., & Frank, R. 1992. Three-dimensional in vivo mapping of brain lesions in humans. *Archives of Neurology,* **49,** 137–143.

Francis, W. N., & Kucera, H. 1982. *Frequency analysis of English usage.* Boston: Houghton Mifflin Company.

Hickok, G., Clark, K., Erhard, P., Helms-Tillery, K., Naeve-Velguth, S., Adriany, G., Hu, X., Tomaso, H., Bellugi, U., Strick, P. L., & Ugurbil, K. 1995. Effect of modality on the neural organization for language: An fMRI study of sign language production. *Society for Neuroscience Abstracts,* **21,** 694.

Savoy, R. L., Bandettini, P. A., O'Craven, K. M., Kwong, K. K., Davis, T. L., Baker, J. R., Weisskoff, R. M., & Rosen, B. R. 1995. Pushing the temporal resolution of fMRI: Studies of very brief visual stimuli, onset variability and asynchrony, and stimulus-correlated changes in noise. *Proceedings of the Society for Magnetic Resonance, Third Scientific Meeting.*

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweiler, D. P., Katz, L., & Gore, J. C. 1995. Sex differences in the functional organization of the brain for language. *Nature,* **373,** 607–609.

Wong, E. C., Bandettini, P. A., & Hyde, J. S. 1992. Echo-planar imaging of the human brain using a three axis local gradient coil. *Proceedings of Society of Magnetic Resonance in Medicine 11th Annual Meeting,* **105.**