NOTES AND DISCUSSION

Some Remaining Questions about Studying Phonological Processing with PET: Response to Demonet, Fiez, Paulesu, Petersen, and Zatorre (1996)

DAVID POEPEL

Biomagnetic Imaging Laboratory, University of California, San Francisco, 513 Parnassus Avenue, S-362, San Francisco, California 94143-0628

SELF-DEFENSE

One cannot but be grateful for such a detailed and spirited response to a paper (Poeppel, 1996), even when the outcome is such that one is deemed “naive,” “superficial,” “unrealistic,” and so on. In spite of my personal shortcomings, the article has at least served to unify several different groups using PET and motivated them to clarify the rationale underlying some of their studies. Since Demonet, Fiez, Paulesu, Petersen, and Zatorre and I are presumably on the same side with respect to the big picture, perhaps it would be useful to try to derive some common ground from all this.

The structure of my argument was as follows. (1) There are several studies in the literature that make explicit claims about the neural basis of phonological processing. (2) The experimental paradigms used in these studies typically involved subjects executing some form of rhyme judgment. Therefore it might be helpful to compare these studies. (3) There is a fair amount of discrepancy among the results of these studies (no overlap/sparse overlap). (4) The discrepancy among the data is probably due to the complexity of phonological processing: different tasks engaged different aspects of phonological processing. (5) Therefore I reached what I thought to be a conservative and uncontroversial conclusion (which was stated multiple times): it is too early to attribute phonological processing to specific brain areas on the basis of these PET data.

1Fax: (415) 502-4302. E-mail: poeppel@itsa.ucsf.edu.
Although Demonet et al. state that the basic observation (no overlap) is correct, their response is predicated on the conviction that one should not expect the experiments reviewed in Poeppel (1996) to converge because of two observations: first, the experiments were not similar in the relevant respects, and second, phonological processing is not unitary but a complex process. I could not agree more—although I would still expect to see some agreement across studies designed to engage phonological processing. Be that as it may, I tabulated the degree to which the studies were similar or different (Tables 1 and 2 in my original paper) in order to highlight how differences among results might arise. Second, not only did I not suggest that phonological processing is a unitary process, but—again, explicitly—I took the fact that phonological processing is a nonunitary process with multiple subroutines to lie at the basis of the no-overlap observation. Since Demonet et al. agree with these two points that form the basis of the argument, their critical reply is more of a restatement than a rebuttal.

In considering what factors can influence the outcome of PET studies, the discussion was focused specifically on how the behavioral paradigms employed could lead to ambiguous conclusions. That there exist purely technical reasons that might account for certain discrepancies is also uncontroversial. Many of the (technical and nontechnical) limitations have been discussed by other authors (Chertkow and Bub, 1994; Sergent, 1994; Whitaker and Hochman, 1995). They include problems such as how to standardize (normalize) across brains (the issue of “warping”), how to decide among statistical methods to optimally tease out effects, how to coordinate the spatial information obtained by PET measurements with the time course of information processing (Heinze et al., 1994), and so on. Many of these issues have been confronted with success in the literature, for example in work by Friston and his colleagues (e.g., Friston et al., 1995). In contrast to these technical problems—which can naturally also lead to considerable discrepancies among results—I deliberately concentrated on how to interpret blood flow data in light of the behavioral paradigms used to elicit responses (see also Sergent, 1994; Whitaker and Hochman, 1995). In this brief response I want to again remain focused on the interpretation of the PET results. Two

---

2 For an important result relevant to replicating PET measurements, see a recent study by Chertkow and colleagues (1996). Chertkow et al. performed the same experiment with four groups of subjects and, curiously, observed a surprising amount of variability across the groups despite identical execution of the studies.

3 Paulesu complains that I did not use Tailarach coordinates to make my comparisons but rather used Brodmann areas. As I made clear (see Table 3 in my original paper), Brodmann designations were used merely as mnemonics; I am aware that data are reported in Tailarach space. Indeed, if the discrepancies among studies had been that subtle, it would have been beneficial to use a standardized coordinate space. However, the variation in the data ranges across lobes, so that I think mnemonic names suffice to illustrate the no-overlap effect.
points are emphasized. First, whereas I agree that a convincing model of verbal working memory can be derived from the data, it is important not to confuse working memory (the articulatory loop) with phonological processing in general. Second, the reason that it matters that the tasks (and, consequently, the subtractions) have a clear connection with established properties of natural language is that the goal of functional neuroimaging is presumably the development of a cognitive neuroscience of basic cognitive processes, not a cognitive neuroscience of performing particular experimental tasks.

A POINT OF AGREEMENT: THE ARTICULATORY LOOP

Demonet et al. disagree with my conclusion and believe that a coherent framework of phonological processing does emerge from the studies discussed in Poeppel (1996). Based on those data as well as new findings, I think they are correct in asserting that a coherent story emerges, but I think it is important to be clear that it is Baddeley’s verbal working memory model (specifically, the articulatory loop) that is turning out to provide a unifying framework for the data (Baddeley, 1986). It is satisfying that such a well-established cognitive psychological theory is supported by imaging results. But insofar as claims about speech and language processing are articulated, one must bear in mind that it is a theory about verbal working memory, not phonological processing as it pertains to speech perception or production.

Like Demonet et al., I am impressed with the convergence of data with regard to the activation in Broca’s area (and also left posterior superior temporal gyrus, and the left supramarginal gyrus). I suggested that this aspect of the results converges across studies (for example, in my discussion of the experiments by Zatorre, Paulesu, and Demonet) and that the verbal working memory interpretation makes good sense. Demonet et al.’s Table 1 summarizes the data for Broca’s area activation across studies and illustrates that there is good agreement in terms of Talairach coordinates. A persuasive model of the neural basis of the articulatory loop that has thus been developed, with (at least part of) Broca’s area subserving the subvocal rehearsal component and (at least part of) left supramarginal gyrus subserving phonological storage.

This model raises two important challenges for the cognitive neuroscience of speech and language processing. First, it requires that a much more elaborate model of Broca’s area be developed to account for the range of findings now associated with Broca’s area. For example, Broca’s area can also be activated in absence of subvocal rehearsal tasks (Price et al., 1994; McLaughlin et al., 1992) and is typically associated with explicitly syntactic tasks (Stromswold et al., 1996). In order to arrive at a model that integrates (a) the articulatory loop findings, (b) the non-memory-driven activations of Broca’s area, and (c) the classical deficit-lesion results, one is
compelled to postulate either multiple subdivisions of the area or an extremely complex functional role for a unitary area.

ARE WE AFTER BASIC COGNITIVE PROCESSES OR EXPERIMENTAL TASKS?

A second challenge associated with the articulatory loop model is to now connect the verbal working memory results as closely as possible with phonological processing in speech perception and production. Span tasks and rhyming tasks recruit phonological knowledge in a very particular way that—while related to real-life speech processing—is not equivalent. Since one of the goals of this research is presumably to understand how speech and language are organized, it would be interesting to see what kind of role the implicated areas play in more typical situations using spoken language.

The issue of ecological (linguistic) validity was raised in the general commentary as well as in the replies by Petersen and Fiez and by Zatorre, and I want to clarify my position. First, it is not at stake whether phonological processing has been engaged in the studies—clearly it has been, since presumably every linguistic stimulus engages phonological processing to some extent. Second, it is not argued that to study language one is required to use naturalistic tasks (cf. Poeppel and Johnson, 1995). However, Demonet et al.’s assertion that ‘the degree to which an experimental task mimics a real-life event is only of marginal relevance’ is debatable. Tasks need not mimic real-life events, but if one is attempting to understand real-life events such as speech and language processing, the tasks should presumably be connected to real-life events in a way that allow inferences to be drawn about real life. What I take to be the relevant question, then, is this: Do the experiments (and, by extension, the subtractions that are set up) engage phonological processing in a way that allows one to draw conclusions about how people process phonological information in more canonical circumstances, i.e., when listening to or producing words or sentences? If the tasks and subtractions are not connected to speech sound processing in a clearly interpretable way, this raises the concern that one is studying the neurobiology of task execution rather than the neurobiological basis of phonological processing. While both are interesting, they are not coextensive, and understanding how, for instance, span and rhyming tasks are executed does not generalize in a straightforward manner to understanding how spoken language is processed independent of memory tasks.

CONCLUSION

Demonet et al. evidently interpreted my paper as arguing that PET as a methodology is subject to principled problems that preclude its usefulness for understanding the neural basis of speech and language processing. Of course, nothing to that effect was stated. Some recent studies successfully
use new analysis techniques in studying auditory processing (Fiez et al., 1995); other studies integrate data from EEG or MEG to understand the time course of PET data (e.g., Heinze et al., 1994). That PET studies will generate important data to elucidate the neural basis of cognitive processes is therefore not at stake. What I hoped to illustrate was the following: To obtain results that have compelling implications for how the speech and language systems are organized, it is crucial to consider what kind of language paradigms are chosen and how they are implemented (cf. Poeppel and Johnson, 1995). It is not controversial that functional imaging is a powerful tool, but I agree with the positions articulated by Ungerleider (1995) and Raichle (1996), who argue that the success of the functional brain imaging in cognition depends not so much on technical refinements as it depends on the appropriate use of behavioral paradigms and independently motivated hypotheses. In conclusion, then, I stand by my original assertion that for imaging results to be interpretable it is essential for the research to be tightly connected with cognitive psychological, psycholinguistic, and linguistic theories.

REFERENCES


