6 Defining the relation between linguistics and neuroscience

David Poeppel and David Embick
University of Maryland College Park and University of Pennsylvania

The popularity of the study of language and the brain is evident from the large number of studies published in the last 15 or so years that have used PET, fMRI, EEG, MEG, TMS, or NIRS to investigate aspects of brain and language, in linguistic domains ranging from phonetics to discourse processing. The amount of resources devoted to such studies suggests that they are motivated by a viable and successful research program, and implies that substantive progress is being made. At the very least, the amount and vigor of such research implies that something significant is being learned. In this article, we present a critique of the dominant research program, and provide a cautionary perspective that challenges the belief that *explanatorily significant progress* is already being made. Our critique focuses on the question of whether current brain/language research provides an example of interdisciplinary *cross-fertilization*, or an example of *cross-sterilization*. In developing our critique, which is in part motivated by the necessity to examine the presuppositions of our own work (e.g. Embick, Marantz, Miyashita, O'Neil, Sakai, 2000; Embick, Hackl, Schaeffer, Kelepir, Marantz, 2001; Poeppel, 1996; Poeppel et al. 2004), we identify fundamental problems that must be addressed if progress is to be made in this area of inquiry. We conclude with the outline of a research program that constitutes an attempt to overcome these problems, at the core of which lies the notion of computation.

PROBLEMS

In principle, the combined study of language and the brain could have effects in several directions. (1) One possibility is that the study of the brain will reveal aspects of the structure of linguistic knowledge. (2) The other possibility is that language can be used to investigate the nature of computation in the brain. In either case, there is a tacit background assumption: namely that the combined investigation promises to generate progress in one of these two domains. Given the actual current state of research, these two positions — rarely questioned or, for that matter, identified in studies of language and the brain — lack any obvious justification when examined carefully. If asked what to study to learn about the nature of language, surely one would not send a student to study neuroscience; rather, one might recommend a course in phonetics or phonology or morphology or syntax or semantics or psycholinguistics. Similarly, if asked about neurobiology, one typically does not recommend the study of linguistics, or even neurolinguistics. Thus the idea that neuroscience is in a position to inform linguistic theory, and vice versa, is clearly open to question. (3) A third option is that the cognitive neuroscience of language should be pursued as an end in itself. To the extent that this option can be coherently formulated as a program of research (what point is there to a science of language and brain that contributes to the understanding of neither?), results in this domain run the risk of being effectively sui generis; that is, isolated from other research programs in such a way that they do not form the basis for progress beyond the immediate question addressed in any given study. At the very least, then, it is clear that current neurolinguistic research has not advanced — in an *explanatorily significant* way — the understanding of either linguistic theory or of neuroscience. While this failure is by no means necessary, we contend that it will continue until certain fundamental problems are identified, acknowledged, and addressed.

Here we concentrate on two problems. The first problem, which we call the *Granularity Mismatch Problem* (GMP), is that there is a mismatch between the 'conceptual granularity' of the elemental concepts of linguistics and the elemental concepts of neurobiology and cognitive neuroscience (which are, relative to the corresponding linguistic primitives, coarse-grained). This mismatch prevents the formulation of theoretically motivated, biologically grounded, and computationally explicit linking hypotheses that bridge neuroscience and linguistics. Naturally, the GMP applies not just to the linguistics-neuroscience interface, but equally to other experimental disciplines that operate with objects of different sizes.

> **Granularity Mismatch Problem (**GMP): Linguistic and neuroscientific studies of language operate with objects of different granularity. In particular, linguistic computation

> involves a number of fine-grained distinctions and explicit computational operations. Neuroscientific approaches to language operate in terms of broader conceptual distinctions.

The second problem is called the *Ontological Incommensurability Problem* (OIP): the fundamental elements of linguistic theory cannot be reduced or matched up with the fundamental biological units identified by neuroscience. This problem results from a failure to answer the question of how neurological structures could be specialized to perform specific types of computations, linguistic or otherwise. That is, while our particular focus here is on language, the GMP and OIP could be applied to the entire range of areas in which the relationship between cognition and biology is examined, and thus are general 'interface problems' for the study of cognition

Illustrating what we take to be the 'contact-problems' or 'interface-problems' between linguistics and neuroscience, consider the central dilemma, illustrated in Figure 1. The figure enumerates aspects of the architecture of each domain and directly exemplifies the conceptual mismatches. The natural move given these two distinct sets of categories is to attempt a reduction or a direct mapping between one set of categories and the other.

Linguistics	Neuroscience
Fundamental elements of representation (at a given analytic level)	
distinctive feature	**dendrites, spines**
syllable	**neuron**
morpheme	**cell-assembly/ensemble**
noun phrase	**population**
clause	**cortical column**
Fundamental operations on primitives (at a given analytic level)	
concatenation	**long-term potentiation (LTP)**
linearization	**receptive field**
phrase-structure generation	**oscillation**
semantic composition	**synchronization**

FIG. 1 Some primitives for representation and processing. The two unordered lists enumerate some concepts canonically used to explain neurobiological or linguistic phenomena. There are principled ontology-process relationships within each domain (i.e. vertical connections). However, if we take these lists seriously, the interdisciplinary (i.e. horizontal) connections remain, at best, arbitrary.

A direct reduction would involve connecting linguistic categories on the left to neurobiological categories on the right with an arrow that implies a direct computational connection between the two. To our knowledge, there is not a single case of a successful reduction in these terms in the domain of language; it appears that that the categories on the two sides are simply listed using different alphabets (or 'currencies'):

> **Ontological Incommensurability Problem (OIP):** The units of linguistic computation and the units of neurological computation are incommensurable.

The OIP does not suggest that no progress is being made in either the linguistic or neurobiological ontology; clearly, each of these is becoming increasingly refined, with improved empirical coverage. Rather, the OIP encapsulates the observation that these ontologies are developing independently of each other, with no solid connections linking them. In part this is the result of the fact that the objects/processes in each column (Figure 1) have been introduced in order to allow for certain types of generalizations. But the generalizations that these notions permit are different in kind. For example, the *morpheme* is introduced to capture regularities concerning the terminal elements of the syntax, i.e. the minimal pieces of word- and sentence-structure; *linearization* operations are introduced to characterize the required process that transforms hierarchical representations into representations suitable for our available input-output machinery; and so on. In contrast, *neuron* is an anatomic unit that can encompass numerous distinct processing subroutines, and *synchronization* is postulated as a hypothesis about how spatially and temporally distributed neural activity might be coordinated in the generation of unified perceptual experience. It is evident that a direct mapping is extremely problematic. Indeed, it is conceivable that the conceptual architecture of linguistics and neurobiology as presently conceived will never yield to any type of reduction, requiring instead substantive conceptual change in one or both of the disciplines (in the sense of Carey 1985) that might enable unification (in the sense of Chomsky 2000). This problem, once again, is a more general challenge in the cognitive neurosciences and is exemplified here on the basis of the linguistics-neuroscience interface, although all approaches with interfaces of differing character face these issues.

We suggest a straightforward solution to the GMP and OIP, namely spelling out the ontologies and processes in computational terms that are at the appropriate level of abstraction (i.e. can be performed by specific neuronal populations) such that explicit interdisciplinary linking hypotheses can be formulated. Based on our discussion, we suggest a program of research that pursues the second strategy mentioned above, namely that the use of linguistically motivated categories can support the study of computation in the brain. In other words, rather than pursuing the standard approach in which linguistically postulated categories must be validated by biological data, a position which we argue to be fundamentally flawed, we recommend taking linguistic categories seriously and using them to investigate how the brain computes with such abstract categorical representations. Importantly, our perspective advocates an integrated approach to the study of linguistic computation, in which linguistic theories must be accountable to all forms of evidence, including psycho- and neurolinguistic results. The integrated approach has direct implications both for the cognitive neuroscience of language and for linguistic theory, implications that are identified as the discussion proceeds below. In this and other ways, this approach stands in contrast to the prevailing view in neurolinguistics, to which we now turn.

THE STANDARD RESEARCH PROGRAM IN THE COGNITIVE NEUROSCIENCE OF LANGUAGE

The canonical assumption of the Standard Research Program about research on brain and language is that neurobiological methods are used to validate concepts and categories introduced to the experimental research program by linguistic theory. For example, theoretical linguistic research deals with elemental concepts such as ‘root,’ ‘functional category,’ or ‘ head movement,’ and the neurolinguist is supposed to set out to obtain correlative biological measures that provide support for the concept in question. On this view, the data generated by the range of techniques that are used in neurolinguistic research -- i.e. the neuropsychological deficit-lesion method, EEG, MEG, PET, or fMRI -- provide evidence for concepts, representations, and processes that are independently motivated by linguistic research, and the neurolinguistic data give the theoretical-linguistic conceptual apparatus the imprimatur of hard science methodology. This approach constitutes a form of reductionism in which biological evidence is ‘better’ or more fundamental than other evidence.

Research in this vein has a long and respectable tradition and, to be sure, many important results have been obtained. Indeed, the observation that localized brain lesions or brain activation correlate with specific linguistic domains has been foundational for modern neuroscience research (for review and new perspectives, see Hickok &

Poeppel 2004). Modern studies using contemporary recording techniques show that some of the relatively broad distinctions one can draw in linguistics (e.g. syntax versus lexical semantics versus phonology) are reflected in biological data. While such insights are certainly scientifically interesting, clinically relevant, and receive considerable popular attention, there are clear limitations to this methodology that dampen our enthusiasm about this approach as a comprehensive research program. Although this type of research provides the field with important correlative datapoints, one learns little of explanatory depth about language and little about the brain. That is, while such results might indicate the existence of some correlation between linguistic and biological objects, there is no theory of such correlations, nor do such correlations necessarily lead to any further understanding of how brain structures or linguistic computations operate.

The level of computational detail present in studies of linguistic representations and processes far exceeds our knowledge of how to detect such distinctions in the physiological measurements we understand, as well as our know-how about what to look for in the data. As a result, the (often implicit) belief that linguistic categories are not 'real' until detected in the brain subjects linguistic investigations to a kind of methodological stricture that simply cannot be taken seriously. It is unreasonable to expect that all distinctions relevant to linguistic computation must have visible reflexes in *current* imaging (or lesion, or psycholinguistic) data. For instance, the fact that the sentences *The cat is on the hat* and *The hat is on the cat* are different grammatical objects, each requiring distinct representation/computation by the grammar, is a fact whether or not these sentences can be shown using current techniques to be different in terms of neuronal activation. An explanatory theory of linguistic computation in the brain should employ linguistic categories as a means of exploring neural computation; but the failure to detect distinctions in any particular case does not necessarily imply that the linguistic distinctions are incorrect. The latter type of inference might be possible in the context of an articulated theory of neurolinguistic computation; but we have nothing like that at present.

It is quite generally the case that contemporary linguistic research investigates fine-grained and subtle distinctions among representations and processes, whereas neurobiological data that are concerned with speech and language probe coarser distinctions, for instance, questions such as *Are there differences between phonological and syntactic processing*? In other words, there is a compelling mismatch (GMP) in what we can learn about language by studying language (a lot, judging by the progress of the last 50 years of linguistic research) and what we can learn about language by studying the brain (not as much, judging by the progress of the last 150 years of neurolinguistic research). Similarly, neurolinguistic research per se rarely leads to principled neurobiological insights. To learn something substantive about brain structure and function, it is necessary that we develop a focused research program that explicitly formulates hypotheses about how particular brain areas execute the complex functions they support. In the specific case of language, it is clear that the standard research program offers relatively little in the necessary direction, and for this reason an alternative research program must be developed.

PROGRESS IN THE STANDARD RESEARCH PROGRAM? IMAGING BROCA'S AREA.

The discussion above concentrates on the fact that the distinctions made in neurological study of language are coarse in comparison with the distinctions made by linguistics. *Syntax*, *semantics*, and *phonology* are not the names of explicit computational tasks, as is often implicit in standard research; rather, these terms refer to (often vaguely defined) general domains ('phrase structure'; 'meaning'; 'sound structure), each of which consists, of course, of numerous computations and representations in any coherent linguistic theory. One consequence of the failure to recognize the coarseness of the categories employed in the cognitive neuroscience of language is that there are instances of false convergence. In the particular case that we briefly examine in this section, the false convergence is one that suggests that 'Broca's Area' is a (more or less monolithic) cortical area whose function is to compute syntax (the latter construed as a more or less monolithic task). While many functional imaging studies have argued for such a conclusion, closer examination reveals that this interpretation is not tenable (Haagort, this volume, and Thompson-Schill, this volume, discuss Broca's complex and its putative functions) and that the difficulties in this area arise, among other reasons, from the failure to analyze neurolinguistic computation at the correct level of granularity. (We ignore here the additional more technical problems that confront such functional imaging studies, including issues associated with more fine-grained anatomic distinctions, experimental design, analysis, as well as implicit assumptions about the relationship between loci of activation and cognitive systems.)

The activation of Broca's area has been reported in many studies of both syntactic comprehension and production, leading many researchers to conclude that this area has a privileged status in syntax. Elsewhere we review this work in more detail (Embick & Poeppel, in press); here we limit ourselves to a few examples from different techniques (PET, fMRI), designs (block versus single trial), and sensory modalities (auditory versus visual) to illustrate the generality of the issue. Turning to specific studies, Dapretto and Bookheimer (1999), used fMRI in a block design, and presented sentences auditorily to subjects who performed one of two tasks. In a condition labeled 'syntactic', participants were asked to judge whether two sentences -- one active (*The policeman arrested the thief*), and one passive (*The thief was arrested by the policeman*) -- were the same or different. In the 'semantic' condition, subjects judged whether two sentences in which a single word differed were the same (*The lawyer/attorney questioned the witness*) or different (*The lawyer/driver questioned the witness*). This study reported activation in BA 44 for the comparison syntax minus semantics (as well as syntax minus rest), and activation in BA 47 for semantics minus syntax. Auditory presentation was also used in the event-related fMRI study performed by Ni et al. (2000), in which subjects performed syntactic and semantic oddball tasks, in which a sequence of grammatical sentences contained an occasional deviant oddball (syntactic: **Trees can grew*; semantic: *#Trees can eat*). A subtraction of *semantics* from *syntax* showed activation in BA 44/45. A block design with visual presentation was employed in the PET study of Moro et al. (2001). The study employed silent reading and acceptability judgments on four types of Italian sentences: a baseline of Jabberwocky; word-order violations; morphosyntactic violations; and phonotactic violations. Activation for the syntactic and morphosyntactic conditions minus the phonotactic condition was found in left BA 45, and Right BA 44/45. An fMRI study by Kang, Constable, Gore, and Avrutin (1999) used an event-related design in which subjects were presented visually with phrasal stimuli containing syntactic and semantic violations. The stimuli were verb phrases like *drove cars* (the normal condition). There were two deviant conditions: syntactically deviant, e.g. **forgot made*; and semantically deviant, e.g. **wrote beers*. Relative to the normal condition, activation was found for both the syntactically and semantically deviant stimuli in BA 44/45; the activation in left BA 44 was greater for syntax than for semantics. In addition to the studies using anomaly detection/judgment outlined above, activation in Broca's area has also been reported in studies of the syntax of artificial language learning (Musso et al. 2003), as well as in studies of syntactic complexity (Caplan et al. 1998).

Despite the different tasks and designs in these studies, the fact that Broca's area (defined as BA 44/45) was consistently active in a number of 'syntax' studies seems at first glance to be confirmation of the claim that this area is specialized for syntax. Even limiting ourselves to the imaging literature, however, there are considerations that suggest that this conclusion is at best an oversimplification.

The first additional consideration is that Broca's area has been reported to be active in a number of linguistic tasks that are not (overtly) syntactic. These other tasks range from sub-lexical and lexical tasks, for instance auditory lexical decision (Zatorre et al. 1992; Poeppel et al. 2004) to studies of minimal pairs in tone languages (Gandour et al. 2000) to phonetic tasks such as the the processing of rapid phonetic transitions or phoneme sequences (Fiez et al. 1995; Gelfand and Bookheimer 2003). Burton (2001) reviews imaging studies that implicate BA 44/45 in phonetics and phonology. From that review it can be concluded that the claim that Broca's area is exclusively devoted to syntax is incorrect, although it leaves open the possibility (examined below) that Broca's area is specialized for language in some broader sense.

The second consideration that complicates the simple view of a straightforward syntax-Broca's area mapping is the fact that Broca's area is active in a number of entirely non-linguistic tasks; naturally these findings also challenge the more general claim that this area is specialized for language, and not simply syntax. The tasks include motor activation, motor imagery, and rhythmic perception (see Embick & Poeppel, in press, for discussion).

The interpretation that identifies Broca's area as responsible for syntax is, naturally, informed by sources of evidence other than imaging studies, including deficit-lesion studies and electrophysiological studies. Concentrating on imaging studies, to which much recent energy has been devoted, it is clear that a simple mapping between 'Broca's area' and 'syntax' cannot be maintained (cf. Haagort, this volume; Thompson-Schill, this volume). While these results generate an apparent contradiction, this situation cannot be surprising given a realistic view of how cognitive functions such as the construction and manipulation of a syntactic representation are computed. In linguistic domains other than syntax, for instance, a complex internal structure is clearly required for processes such as phonetic and phonological analysis, lexical analysis, and so on. Therefore the expectation that syntax should be a simplex, unstructured computation associated with a single undifferentiated cortical region is unrealistic, and probably hopeless as a hypothesis for guiding future research. It is clear that one, or perhaps several of the computational subroutines that are essential for syntactic processing/ production are computed in the Inferior Frontal Gyrus (IFG). But these are not 'syntax' per se -- they are computational subcomponents of syntax. What is required is a theory that identifies these operations at the correct level of abstraction or granularity and seeks to associate them with different subparts of 'Broca's complex' (Haagort, this volume) and other implicated brain areas. For example, two components essential to syntax are the creation of hierarchical structures and a process that linearizes these hierarchical structures. These are the kinds of computations that can be abstracted from syntax in the broad sense, and which are perhaps associated with different subparts of the IFG. The natural assumption is that the differently structured cortical areas are specialized for performing different types of computations, and that some of these computations are necessary for language but also for other cognitive functions. For instance, the activation of 'mirror neurons' (Rizzolatti and Arbib 1999) in the IFG has a role in motor action/imitation, but also finds a natural place in the linguistic domain in the context of 'forward' models of speech perception (Halle 2002). Thompson-Schill (this volume) attributes to at least one part of 'Broca's Complex', specifically BA 47, the generic role of "selection between competing sources of information". While this type of operation is so general that it must hold for virtually any cognitive process, one might be able to work out for what specific aspects of language an operation of that type could be relevant.

Based on this brief summary, we cannot conclude that major insights have been obtained concerning the structure of language or our understanding of the brain. This negative conclusion holds in spite of the fact that not all discussions of Broca's Area are subject to the criticisms leveled above (Hagoort, this volume; Thompson-Schill, this volume; Horwitz et al. 2003). That is not to say that the imaging work is not (a) clinically helpful and (b) potentially informative to theory construction. On the contrary, in conjunction with an appropriately granular theory of the computations performed in the brain, the spatial information provided by imaging has the potential to illuminate aspects of the biological foundation of language by providing the critical link between specialized cortical areas and cognitively relevant types of computations. However, in the broader context of the issues addressed in this paper, it is clear that what look like results linking linguistic and neurological categories in the case of Broca's area are actually *problems*; and these problems result from the limitations that are inherent to the standard research program.

STEPS TOWARDS PROGRESS: ELECTROPHYSIOLOGICAL STUDIES?

We have argued that the imaging literature, although rich with important correlative information, is, for the moment, unsatisfying as a source of information likely to enrich explanatory models. What is the status of electrophysiological research? In fact, most of the work builds on the same assumptions as most imaging studies. One aspect of standard electrophysiological work on language processing that underscores this perspective is that the experiments reflect a 'reification' of ERP components. Specifically, many (probably most) studies on the LAN, N400, and P600/SPS components interpret each component as reflecting *syntax* or *semantics* or *phonology*. Indeed, a major goal of many studies, much like in imaging, is to dissociate syntactic from semantic and phonological processing. This may be a useful goal (of an intermediate type), but it again highlights the mismatch between the granularity of linguistic versus neurolinguistic concepts. An ERP component cannot reflect syntax per se, because syntax is not a single computation. Moreover, by not looking to the subroutines involved, it misses the overlap that might occur because computational subroutines are shared by different processes (say, for example, linearization).

There are, of course, numerous exceptions, i.e. studies that attempt to probe in detail how linguistic categories and computations are executed. We merely point out that, typically, the main distinctions being drawn in such electrophysiological studies using EEG or MEG are *syntax* versus *semantics* versus *phonology*, and the standard interpretation is that the LAN 'is' syntactic structure building, the N400 'is' lexical semantic integration, and the P600 'is' syntactic error detection (and perhaps reanalysis and repair processes). In this way, there is no substantive distinction, at the conceptual level, of studying linguistic representation and computation between imaging and electrophysiological approaches.

PROSPECTS: REDEFINING A RESEARCH PROGRAM

Putting aside simple associations like 'syntax is in Broca's area', the next move is to appeal to a finer-grained set of categories derived from ongoing research in linguistic theory and in neuroscience. We take it that the central question of neurolinguistic research is the question of how the grammar of human language is computed in the human brain. Our revised research program diverges from a familiar assumption in linguistic theory, which often proceeds as if experimental evidence -- whether from neuroscience or psycholinguistics -- is *in principle* irrelevant to theories of how language works. This assumption, which is often tacit in linguistic theory, is made manifest in the idea that there might be notions of 'psychological' or 'neurological' reality that are distinct from the reality that linguistic theory addresses. This view of linguistic reality is incompatible with our approach to language and the brain.

The grammar consists of representation and computations. We assume that linguistic computations are executed in the brain in real time. There is no need for terms like 'psychologically real' or 'neurologically real.' These terms, because they are qualified, imply that there is some other type of reality to linguistic computations beyond being computed in the brain. If a linguistic analysis is correct -- i.e. identifies something real -- it identifies computations/representations that are computed in the minds/brains of speakers. How these computations are implemented at different levels of biological abstraction is the primary analytical question for neurolinguistics. As noted, our perspective requires an integrated theoretical and experimental perspective, something that runs contrary to a current trend in linguistic theory. The tendency in generative syntax, for example, is to speak as if the computations proposed in syntactic analyses need not be regarded as computations that are performed in real time. But why should the null hypothesis be that there is some notion of grammar that is not computed in the brain in real time? This assumption simply makes the link between linguistics and neuroscience harder to bridge, for reasons that are ultimately historical, and not necessarily principled. Just as the research program of neurolinguistics must be informed by linguistic theory, linguistic theory cannot proceed in a way that systematically ignores experimental results. Even if specific instances in which experimental data resolve questions of theory are difficult to come by at present, this is a fact that reflects technical and methodological difficulties and a non-integrated research program; in principle, the forms of evidence on the language faculty that are provided by these methodologies are just as relevant to linguistic theory as, say, native speaker intuitions are.

At the level of the computations referred to in the preceding discussion, our revised research program insists that we restrict our attention to computations that are actually performed by the human brain. That is, the notion of computation that is central to our research program is not an abstract model of computation; we are interested in the question of what computations are performed in the brain, and not some way of modeling behavior. Ultimately if we discover restrictions on the types of abstract computations the brain can perform, we might discover as a result the nature of some of the properties of human language. But this linking is only possible given our assumptions about the grammar and the nature of computation outlined above.

One way to proceed is to stand typical neurolinguistic research on its head. Suppose one abandons the central concern with identifying correlations between biological measurements and previously hypothesized elements of language processing and aims, instead, to explicitly use elemental linguistic units of representation and computation to investigate how the brain encodes complex information. More colloquially, suppose we use language to learn how the brain works. Based on established and empirically well supported distinctions drawn in linguistics (say the notion of 'constituency' or the notion of 'distinctive feature'), we work on the problem of how the brain encodes complex and abstract information, in general, and linguistic information, in particular. Insofar as we learn additional facts about the language system (that were not visible to linguistic or psycholinguistic research per se), we are delighted – and happily take credit for any serendipitous findings. However, the basic assumption is that we study aspects of brain function by relying on a system whose cognitive architecture is well understood (like the visual system, for example).

There are many levels of analysis at which one could proceed from this perspective. In some of our own research, we are beginning at the beginning, i.e. with the process of speech perception. Speech perception is of interest because it forms the basis for the transformation of physical signals into the representations that are used for computation in the brain (see Scott, this volume). One fundamental challenge for the system is how to transform continuous physical signals (acoustics) into the abstract, discrete representations that form the basis for further linguistic computation. We can build on the theoretical position that the elementary linguistic constituent is the 'distinctive feature' (e.g. Halle 2002), and from that perspective the computational challenge is to go from sound to feature. This transformation of information is non-trivial: no automatic speech recognition system comes anywhere close to the performance of a human.

Preliminary results have demonstrated that it is possible to probe neural representation by using linguistically motivated categories like distinctive feature. For instance, Phillips et al. (in prep) investigate the neural response to stimuli that differ in terms of a phonological feature [±voice]. The study employs a paradigm in which all stimuli differ acoustically. Despite these acoustic differences, all stimuli fall into the major categories defined by a phonological feature. The results of this study suggest that the brain can employ phonological (as opposed to acoustic) categories like [±voice] for computation by 180ms. Thus, by making use of *distinctive feature*, motivated by linguistic research, the experimental study is able to derive claims about the time-course of auditory processing in the brain. Eulitz & Lahiri (2004) take the relevance of abstract features further, providing neurophysiological evidence that the hypothesized abstract primitives at the basis of lexical representation can be probed with such an approach.

Moving from the phonetic/phonological level to the domain of syntax, matters become more complex. The general strategy we have outlined calls for a separation of cognitively realistic computations from more general areas, such as *syntax*. One potentially promising operation of this type is the operation of linearization. The hierarchical representations motivated by syntactic theory must have a linear order imposed on them, because of the requirement that speech be instantiated in real time. In addition to being necessary for syntax, it is quite plausible that linearization operations of this type are also required in other linguistic and cognitive domains (e.g. for phonological sequencing, or for motor planning/execution, respectively). Extracting the computational operation (or operations) of linearization from these different domains amounts to approaching the problem at the correct level of granularity, in the manner we have stressed above: linearization operations of a specific type have uniform computational properties, and it might be expected that certain brain regions are specialized to perform this type of computation. Ultimately it is possible that the use of (a family of) linearization operations in different cognitive tasks broadly construed is in part responsible for the apparently puzzling activation of Broca's area reviewed above.

There is much work to be done in these areas. To the extent that we have made progress in clarifying a research program that promises to yield substantive results, we still have not come close to the problem of how specific computations are executed by specialized brain regions. But the agenda we have outlined makes it possible to move closer to such questions, by highlighting the importance of concentrating on the nature of computational operations in language at the correct level of granularity.

CONCLUSIONS

The joint study of brain and language -- cognitive neuroscience of language -- has achieved some basic results correlating linguistic phenomena with brain responses, but has not advanced to any explanatory theory that identifies the nature of linguistic computation in the brain. Results from this area are therefore in some ways both confused and confusing. The absence of an explanatory theory of this type is the result of the conceptual granularity mismatch and the ontological immensurability between the foundational concepts of linguistics and those of neurobiology: the machinery we invoke to account for linguistic phenomena is not in any obvious way related to the entities and computations of the biological systems in question. Consequently, *there is an absence of reasonable linking hypotheses* by which one can explore how brain mechanisms form the basis for linguistic computation.

If this critical perspective is on the right track, there is significant danger of (long-term) interdisciplinary cross-sterilization rather than cross-fertilization between linguistics and neurobiology, or, for that matter, linguistics and other empirical disciplines. To defend against being subjected to a poverty-of-the-imagination argument, we suggested a substantive alternative research program. The critical link between disciplines should come from computation, specifically, from computational models that are made explicit at the appropriate level of abstraction to create an interface for linguistics and neurobiology. By hypothesis, in such computational models the primitives and operations must (i) be of the type that they can plausibly be executed by assemblies of neurons -- thereby providing the neurophysiological grounding -- and (ii) reasonably be constitutive subroutines of linguistic computation – thereby providing the theoretical foundation.

Acknowledgments

During the preparation of this manuscript, DP was a Fellow at the Wissenschaftskolleg zu Berlin. DP is supported by NIH R01 DC05660.

References

Burton, M. W. (2001). The role of inferior frontal cortex in phonological processing. *Cognitive Science*, 25, 695–709.

Caplan D., Alpert, N., and Waters, G.. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience*, 10, 541-52.

Carey, S. (1985). Conceptual Change in Childhood. Cambridge, MA: Bradford Books, MIT Press.

Chomsky, N. (2000), Linguistics and Brain Science. In A. Marantz, Y. Miyashita, and W. O'Neil eds., Image, Language, Brain. MIT Press, Cambridge, MA, 13-28.

Dapretto, M., and Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, 24, 427-432.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., and Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proc. Natl. Acad. Sci. USA* 97, 6150-6154.

Embick D, Hackl, M., Schaeffer, J.,. Kelepir, M. and Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency, *Cognitive Brain Research*, 10, 345-348.

Embick, D and Poeppel, D (in press). Mapping Syntax Using Imaging: Prospects and Problems for the Study of Neurolinguistic Computation. In K. Brown (ed.), Encyclopedia of Language and Linguistics 2nd ,Oxford: Elsevier.

Eulitz, C. and Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition, *Journal of Cognitive Neuroscience*, 16, 577-583.

Fiez, J., P.Tallal, M. Raichle, W. Katz, F. Miezin, and S. Petersen (1995). PET studies of auditory and phonological processing: Effects of stimulus type and task condition. *Journal of Cognitive Neuroscience*, 7, 357-375.

Gandour J., Wong D., Hsieh L., Weinzapfel B., Van Lancker D., Hutchins G. (2000). A crosslinguistic PET study of tone perception. *J Cogn Neurosci* 12(1), 207-22

Gelfand, J. and Bookheimer, S. (2003). Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron*, 38, 831-42.

Halle, M. (2002). From Memory to Speech and Back. Mouton de Gruyter.

Hickok, G & Poeppel D, Eds. (2004). Towards a new functional anatomy of language. *Cognition*, 92 (1-2).

Horwitz B., Amunts K., Bhattacharyya R., Patkin D., Jeffries K., Zilles K., Braun A. (2003). Activation of Broca's area during the production of spoken and signed language: a combined cytoarchitectonic mapping and PET analysis. *Neuropsychologia* 41(14), 1868-76

Kang A., Constable R., Gore J., Avrutin S. (1999). An event-related fMRI study of implicit phrase-level syntactic and semantic processing. *Neuroimage*, 10, 555-61.

Moro, A. Tettamanti, M., Perani, D., Donati, C., Cappa, S., Fazio, F. (2001) Syntax and the brain: grammar by selective anomalies, *NeuroImage*, 13, 110 - 118.

Musso M., Moro A., Glauche V., Rijntjes M., Reichenbach J., Buchel C., Weiller C. (2003). Broca's area and the language instinct. *Nat Neurosci*, 6, 774-81.

Ni W., Constable R., Mencl W., Pugh K., Fulbright R., Shaywitz S., Shaywitz B., Gore J., Shankweiler D. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *J Cogn Neurosci*. 12:120-33.

Phillips C, Pellathy T, Yeung, H., Marantz A (in preparation). Phonological feature representations in auditory cortex.

Poeppel, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language* 55:317-351.

Poeppel D, Guillemin A., Thompson J, Fritz J, Bavelier D, Braun AR (2004). Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia* 42, 183-200.

Rizzolatti, G., Arbib MA (1998). Language within our grasp. *Trends Neurosci* 21: 188-194.

Zatorre R., Evans A., Meyer E., Gjedde A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256: 846–9.