
Diana Van Lancker Sidtis *

Department of Speech-Language Pathology and Audiology, New York University and Nathan Kline Institute for Psychiatric Research, 719 Broadway, Suite 200, New York, NY 10003, USA

Accepted 22 May 2006
Available online 12 July 2006

Abstract

Neurolinguistic research has been engaged in evaluating models of language using measures from brain structure and function, and/or in investigating brain structure and function with respect to language representation using proposed models of language. While the aphasiological strategy, which classifies aphasias based on performance modality and a few linguistic variables, has been the most stable, cognitive neurolinguistics has had less success in reliably associating more elaborately proposed levels and units of language models with brain structure. Functional imaging emerged at this stage of neurolinguistic research. In this review article, it is proposed that the often-inconsistent superfluity of outcomes arising from functional imaging studies of language awaits adjustment at both “ends” of the process: model and data. Assumptions that our current language models consistently and reliably represent implicit knowledge within human cerebral processing are in line for major revision; and the promise of functional brain imaging to reveal any such knowledge structures must incorporate stable correlates of the imaging signal as dependent variable.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Functional neuroimaging; Language; Cerebral laterality

1. From observation to interpretation

In the S. Harris cartoon (Fig. 1), two mathematicians stand before a proof on a blackboard, with the middle step, “THEN A MIRACLE OCCURS...” The second scientist has the line: “I think you should be more explicit here in step two.” Even without relying on miracles, the journey from experimental data to interpretation is seldom made without an occasional leap of faith. This cartoon has a particular relevance to scientific presentations of functional brain imaging results: it is not uncommon for a conference speaker to discuss the background and previous studies with a smooth confidence, and then, when slides of brain slices appear on the projection screen, to become visibly hesitant and halting. Yes, numerous brain areas show ranges of color and brightness, but what does it all mean? The technique of functional neuroimaging has had its detractors and concerns have been expressed (Bechtel & Stužek, 2001; Chertkow & Bub, 1994; Coltheart, 2000; Davis, Meunier, & Marslen-Wilson, 2004; Fellows et al., 2005; Poeppel, 1996a; Rugg, 1999; Sidtis, 2000; Uttal, 2001; Whitaker & Hockman, 1995), yet we consumers, like Sorcerer’s apprentices, must contend with “more and more currents (of imaging literature) flowing upon us”1. The purpose of this article is to explore the impact of the functional

☆ A publisher’s error resulted in this article appearing in the wrong issue. The article is reprinted here for the reader’s convenience and for the continuity of this special issue. Anyone wishing to cite this article should use the details of the original publication [Van Lancker Sidtis, D., 2006. Does functional neuroimaging solve the questions of neurolinguistics? Brain and Language 98, 276–290, doi:10.1016/j.bandl.2006.05.006].
* Fax: +1 845 398 5575.
E-mail address: drv1@nyu.edu

0093-934X/S - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/S0093-934X(07)00097-1

1 “Ach, und hundert Flüsse/Stürzen auf mich ein!!” From Goethe, Der Zauberlehrling, 1797.
imaging movement on a particular field, neurolinguistics, with a view toward informally bringing together issues, questions, and concerns that have arisen. The overview begins with a modest attempt to place the advent of functional imaging in a historical and intellectual context, turning next to the theme: What were the unresolved questions of neurolinguistics when neuroimaging arrived on the scene?

2. Theory and data

The relations of observation to theory, and theory to observation are delicate and fragmentary, requiring artful handling, and subject to sociopolitical trend (Fleck, 1979), as much in linguistics as any other scientific undertaking. Sometimes “the fox knows lots of little things,” and observations reign, having their own interest and power, such as in learning theory, genetics, microbiology, or descriptive linguistics; in other domains or in other times (to continue the allusion to the fable), “the hedgehog knows one big thing,” such that theory dominates and observations conform themselves to the umbrella idea, as in the era of Freudian analysis and, since 1957, the school of generative grammar (Chomsky, 1957, 1962, 1965, 1975). The history of science is a story of public opinion negotiating with scientific effort (Feyerabend, 1966).

What is less often talked about, but is crucially important in the language sciences, is the question: what constitutes data, what is a fact, and what leads to a scientific law? Facts are observed events that lead to scientific laws, and in the behavioral sciences, these laws are of the probabilistic variety (Nagel, 1961). Laws such as, “language is processed in the left cerebral hemisphere,” or “Parkinson’s disease is caused by depleted dopamine production by the substantia nigra,” contain numerous inferences, levels of interpretation, and induction from an the observational base—the set of facts. Any scientific undertaking embraces, in other words, imprecision, vagueness, and fragments of data that manifest themselves unevenly. Any scientific theorizing must contend with many degrees of freedom.

3. Constraints in the linguistic sciences

Experience with this truth about the scientific endeavor has promoted the development of constraints, or a convincing enumeration of conditions under which a scientific concept is scientifically meaningful (Nagel, 1961). With specific, universally agreed-upon constraints in place, types of models or types of theories will be limited, monitored by outside influences. Parsimony has often been proposed as a constraint for choosing between models or theories. This principle might well serve the physics of celestial bodies, but there is little in the description of the evolution of brain structures underlying behaviors to convince us that parsimony was significantly in play (Marcus, 2004).

A main complaint leveled against the structural-linguistic descriptivists of the forties and fifties (Francis, 1958) was their lack of theoretical constraint, and therefore, the inability to “know” whether any given structural description was accurate. To remedy this problem, e.g., Chomsky (1962, 1965, 1975) advocated specific guidelines to determine a “possible” human language.” As these guidelines already rose out of his theoretical apparatus—that language ability is innate and universal in humans and that this knowledge can generate a potentially infinite set of sentences, the imposition of constraints on possible models of language constituted a frank circularity. The dismissal of performance data in generative grammar as not pertinent led, in fact, to a continuous turning-over of models of language, and a spinning out of numerous approaches to the description of human language, with, again, no yardstick to choose among them (Harris, 1993; Steinberg, 1982). Psychology, with its empirical tradition, tried to join forces with the staunchly rationalist generative grammarians, and the field foundered (McCauley, 1987; Miller, 1990; Reber, 1987). As R. Jakobson often stated, theory without data is empty, and data without theory are meaningless (Jakobson, 1968). Much of linguistic science has involved finding the right balance between these two essentials.

A number of formal linguists, disappointed in the failure of psycholinguistics to lead to a convincing model of language, turned to neurolinguistics for grounding and for guidance. The attraction was that the examination of language impairment in the damaged brain would highlight structural elements not convincingly demonstrable in formal language analysis, and not discernible in normal language use. Here perhaps the constraints of hardware, the brain structures themselves, would help to chose between structural descriptions of language. The classic approach to neurolinguistic studies, correlation of lesion location with specified language disorder, consisted of efforts to validate or verify proposed “levels” and elements of language, and at the same time, to elucidate brain function (Table 1).
The existence of these levels—phonetics, phonology, morphology, syntax, semantics; and elements—phones, phonemes, morphemes, words, phrases, sentences, syntactic rules—has been relatively uncontroversial among language practitioners (Akmajian, Demers, & Harnish, 1997; Carroll, 1999; Martinet, 1970). The neurolinguistic program based itself on the assumption that brain structure and/or brain processing is/are organized according to these descriptors, and that specific damage would reveal disorders of the matching descriptor. Structural description was expected to have a natural constraint in neurological structure.

---

<table>
<thead>
<tr>
<th>Phonetics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmental</td>
<td></td>
</tr>
<tr>
<td>Prosodic (timing; tones)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phonology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEFT HEMISPHERE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morphology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td></td>
</tr>
<tr>
<td>Semantics</td>
<td></td>
</tr>
<tr>
<td>Linguistic</td>
<td></td>
</tr>
</tbody>
</table>

| Contextual |  |
| Familiar proper nouns |  |
| Emotional terms |  |

<table>
<thead>
<tr>
<th>Pragmatics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation</td>
<td></td>
</tr>
<tr>
<td>Topic, theme</td>
<td></td>
</tr>
<tr>
<td>Inference</td>
<td></td>
</tr>
<tr>
<td>Formulaic language</td>
<td></td>
</tr>
<tr>
<td>Prosodic (pitch)</td>
<td></td>
</tr>
</tbody>
</table>

| emotional meanings |  |
| attitudinal meanings |  |
| intonation in phrases |  |
| voice recognition  |  |

---

2 Their ontological status is queried in this article. The question is "where do these levels and elements exist?"
4. A brief history of brain and language relations

A workable marriage of theory and observation was brought about in what is now referred to as the Geschwindian model of language (Geschwind, 1970; Geschwind & Galaburda, 1985). This is the neurological or aphasiological model, further promulgated by Benson (1979), Goodglass and Kaplan (1972), Henderson (1987), Kertesz (1983) and many of their colleagues and students. Observations in two domains, language disorders and brain lesions, were correlated to describe a coherent account of how brain damage can be expected to affect language performance, and by extrapolation to the converse case, how language is organized in the normal brain. Here physical constraints were in place: it was possible to visualize the location of the lesions, and, theoretically, to describe and quantify the language deficits. But linguistic concepts did not enter in, basically, theoretically, or systematically, as the original parameters for this model. Instead, primary descriptors were gross performance modes: production and comprehension (e.g., nonfunt, fluent aphasia). As mentioned previously, this is in orthogonal conflict with the view of the role of performance data in models of language: “the attempt to account for knowledge in terms of ability is misconceived from the start” (Chomsky, 1988, p. 12) Another performance mode, one that has no status whatever in any kind of language model: structuralist, generative, performance, or otherwise—repetition—was used as a key pathognomonic parameter in determining aphasic category (e.g., conduction, transcortical). The notion of “grammar” as a basic, underlying ability that is lost in aphasia was indeed adduced—as in, for example, N. Geschwind’s classic case presentation of patient Franklin-, but without rigor of detail3; and confrontation naming ability (see Gardner, 1975), a process that has also absolutely no status in a model of language, was the most often used test of language function (Kaplan, Goodglass, & Weintraub, 1983).

With the advent of cognitive neuropsychology, another marriage was attempted: aphasiology and linguistic theory (see Basso, 2003, for review). There followed an impressive range of efforts to test, evaluate, and investigate the levels and elements proposed by linguists to describe normal language ability. Some forty years of neurolinguistic effort, using lesion and accompanying studies, such as EEG, ERPs, dichotic listening, tachistoscope, WADA and split-brain testing, and other adjunctive methods to study these basic linguistic properties, have not convincing validated these elements. It is here that one sees the unresolved questions of neurolinguistics. The major categorization principles in aphasia, used universally to classify language disturbance, still pertain not to structure or element, but to performance modes: production, comprehension, and repetition. This fact leaves consideration of structure and element in language, as reflected in brain damage, sliding about uncertainly on the second tier. And there, ongoing disagreements continue, for example, about such fundamentals as whether basic elements of phonology, grammar, or semantics are the more afflicted in Broca or Wernicke’s aphasia. Since these questions operationally have to be asked separately for production and comprehension, their impact on theoretical stances is further weakened. Proposals for an underlying agrammatism meet with counterproposals that grammatical comprehension is relatively intact or relatively impaired (Badecker & Caramazza, 1985; Caramazza & Zurif, 1976); Wernicke’s aphasia is said to be primarily a semantic disturbance, but grammatical functions are demonstrably also impaired (Caplan, Waters, DeDe, Michaud, & Reddy, 2004a), and phonology is seen as distorted in the nonfuent as well as the fluent aphasias (Levy & Kavé, 1999).

This is a frustrating state of affairs, as the levels and elements in the typical language model (Table 1) have seemed intuitively right and some can be shown experimentally to explain observations in natural language use. For example, speech error analysis yields an account of phonological elements; these can be switched, as in the classic example “the queer old dean,” a malapropism for “the dear old queen”; further, speech errors systematically reveal phrase constituency, the independence of morphemes (minimal units of meaning, such as prefixes and movable endings of words), the integrity of intonational contours, and the autonomy and relational structure of words (Clark, 1970; Cutler, 1982; Levelt, 1989; Shattuck-Hufnagel, 1983). While phonological integrity and grammatical phrasing in general can be demonstrated in various tasks involving normal speech production and language comprehension, reliable correlates with brain structures have not clearly followed.

Thus controversy remains about the psychological reality of linguistic elements. The intractability of these disputes was again revealed at a conference at MIT entitled “From sound to sense,” featuring prominent figures in speech science, when Peter Ladefoged, a world-renowned phonetician, evaluated Morris Halle’s (a world-renowned phonologist) assertion that phoneme is psychologically real as “utter nonsense”4 (Ladefoged, 2004).

5. Linguistic models, psychological reality, and brain codes

The success of neurolinguistics is at least in part dependent on the validity of the model of language utilized to study brain and language relationships, which has undergone considerable morphing (but from another perspective, the cerebral processing results are applied to verification of the language model). But the endeavor is anchored in the hardware of the brain. The problems here appear to arise not from our methodologies for measuring brain structure. It is increasingly suggested that the difficulty lies in the

---

3 Mr. Franklin’s aphasia diagnosis was “mixed.”

4 Dr. Ladefoged later confirmed this statement, adding “Ask any nonliterate native speaker of Chinese.”
assumption that our linguistic models have any orderly presence in functional structure or organizational principles in the brain (Ackermann & Riecker, 2004; Fodor, 2000; Hickok & Poeppel, 2004a, 2004b; Poeppel, 1996b). Production, repetition, and comprehension have remained valid and reliable as descriptors of neurogenic language disturbance, but these have no pertinence to linguistic models as they have been developed in the past century. The proposed linguistic levels—phonetics, phonology, morphology, lexicon, syntax, and semantics—may be more useful as an educational and analytic heuristic than for describing the biology of human language and language disturbance. Their status as autonomous, explanatory elements in mental and cerebral processing is no longer clear. These terms may not reflect the language of the mind or the brain.

It is assumed as a truism that speakers have little or no conscious knowledge of language structure and language rules as described by linguists. We give this phenomenon a name: implicit knowledge. Try teaching undergraduate students, native speakers of English, the meaning of the grammatical terms “active and passive voice” in their own mother tongue. Do they have difficulty because the knowledge is implicit, never to be brought forth to clear consciousness? Or do they have difficulty because English actually has a range of voices, including a middle voice? A similar notion was proposed by Ross (1973), who coined the term “nouniness,” and proposed that words we call nouns occur on a continuum, suggesting that the notions of disjunctive categories, “noun” and “verb” are constructs of the structuralist’s imagination. Or do students have difficulty learning these concepts because cerebral functionality does something completely different with the verbal system, using processes and categories that are opaque to us?

The German language has a three (gender dimension) by four (case dimension) by two (number dimension) matrix describing the grammar of nouns. There is little “sensible” order in this matrix: “der” occurs more or less randomly in six cells, “dem” in two, “den” in four, and so on. Second language speakers must learn the matrix if they are to have any hope of speaking a grammatical sentence, but native speakers, unless taught in school, have no sense of this. The matrix, however neat in tabular form, may have no status whatever in mental organization. Any second language speaker of German is grateful for the few morphological and phonological rules that cue nominal gender (German has masculine, feminine, and neuter, and most common nouns have arbitrary assignment), such as “heit,” and “schaft,” suffixes that always form a feminine noun. An informal survey revealed that no native German queried had knowledge of this rule. Of course, they know the genders of all the nouns, without the rule. Is this rule merely implicit? Or is it totally irrelevant except as a rule of thumb for GSLs (speakers of German as a second language)? The latter seems more likely. Because the German way of saying things sounds right, such rules are not in play. This type of observation, which is ubiquitous in native language competence, is usually attributed to the difference between implicit and explicit knowledge, and is thus dismissed as not reflective of whether or not these structural descriptions actually have a psychological, or neurological presence. But it IS beginning to be questioned: what does make it sound right?

Many of our most obvious categorical, structural, and rule-governed entities, however satisfying their graphic representation, and however parsimonious, rhetorically tidy, and thorough they have proven over centuries of philological, etymological, and linguistic lore, have failed to reveal convincing consistency in brain-behavioral studies. And indeed, interesting papers are appearing in the linguistic literature about how the levels and units fail to have the autonomy and independence formerly assumed (Marmaridou, Nikiforidou, Antonopoulou, & Salamoura, 2006). One example of this shift in viewpoint will suffice. Most of linguistic theory posits a logical delimitation between syntax and semantics, which in the past has been intuitively satisfying and useful in language analysis, but which in actual practice does not exist. These two “levels” commune and interact and signal each other incessantly (Bates & Goodman, 1997; Levelt, 1999; MacDonald, 1993; McClelland, 1987; Marslen-Wilson, 1987; Rayner, Carlson, & Frazier, 1983; Tyler & Marslen-Wilson, 1977). How can the brain be expected to clearly distinguish between these two “levels,” themselves demonstrably inseparable, when carefully viewed, even in our conceptual inventions? Perhaps it is becoming clear that much in our descriptive models, even in the most consensual formats as shown in Table 1, is not how the brain, or the mind, does it.

6. An overview of approaches to studying language processing in the brain

The linguists’ model of language endeavors to present the abstract, underlying structure of universal human language, using the principle of the intuition of the well-formed sentence as criterion (constraint). The psycholinguists’ model strives for formulation and confirmation of structure and process in language use, using performance data. The neurological/aphasiological model utilizes performance modalities to model relationships between brain and language, and its goal is to formulate typologies of language disturbance, using structure-function correlations. The cognitive neurolinguistic effort is to describe how properties of a model of language are represented in the brain, using performance measures from an impressive array of methodologies. As might be expected, the success of any of these disciplines is partial. As might be regretted, these approaches often “miss” each other in focus, goal, in theoretical basis, and in what is acceptable as relevant data.

The neurological model, as practiced in aphasiology, has been the most robust. People have lamented such nonconformities as “mixed” aphasia, the evolution of aphasic diagnoses in individual patients, and the considerable noncorrespondence between lesion site and aphasia type. Numerous anomalous cases, mentioned below, stand as vivid counterevidence. But this approach has had the
advantage of being constrained by observable brain lesions and measurable language disturbance.

There was reason to be hopeful that neurolinguistic studies, expanding from the aphasiological model to include linguistic categories, would bring us a major step forward in understanding language and brain relationships. First, neurolinguistic studies benefit from the constraints of neurology—observable destruction or stimulation of tissue. Second, the linguistic sciences had presented coherent models of language, which could be used to formulate specific questions about language representation in the brain. Yet in viewing the set of neurolinguistic questions, one finds little agreement. The competence-performance conundrum remains. Do persons with nonfluent aphasia, who show deficits in grammatical production, have a general, “underlying” grammar deficit? (Berndt, Mitchum, & Haendiges, 1996; Heim & Friederici, 2003). Attempts to identify common linguistic-grammatical deficits in the aphasias in a recent ten year study drawing upon considerable resource and expertise led to a result that the effect of complexity overrode all other variables tested (Caplan et al., 2004a, 2004b). Probes for more esoteric postulates in generative grammar, such as traces (an empty slot for a grammatical category left in the process of adding a relative clause) proposed to explain certain observations in some types of nonfluent aphasia (Grodzinsky, 1995), have resulted in mixed counterclaims (Caramazza, Capitani, Rey, & Berndt, 2001; DeBleser, Schwarz, & Burchert, 2006).

To account for grammatical performance in aphasic persons, such notions as a preferred canonical sentence type (Jacobs & Thompson, 2000), task demands (Friederici & Frazier, 1992), and rate of processing (Kolk, 1995; Waters, Caplan, Alpert, & Stanczak, 2003) have been added as significant variables. These depart in a major way from the format and spirit of generative grammar linguistics. Nonfluent aphasia is characterized by severely deficient word production. Is this a semantic deficit more severe than that seen in fluent aphasia? Who has the most severe phonological deficit, persons with fluent or nonfluent aphasia? Is anoma, present in all aphasic disturbances, a semantic deficit or a delay in processing? Of course, these questions pertain to clearly diagnosable aphasic syndromes, which occur in normal persons using PET or fMRI described a range of apparently active neural sites in association with a range of language tasks. In many of these early studies, right hemisphere sites were more “active” than left hemisphere sites for word recognition, speaking, understanding sentences, and so on. Strangely, at first, few authors remarked on this discrepancy with nearly 150 years of neurological teaching affirming that the left hemisphere is “dominant” for language (Herrmann & Fiebach, 2004; Obler & Gjerlow, 1999; Pulvermüller, 2002). As verbal tasks became more varied, more and more brain sites were reported, often with noncorresponding results (e.g., Pulvermüller, 2002, p. 47).

It is understood that a model of brain function underlies the data analysis of every imaging effort. Actually, theories both of language and of brain-language relationships must inform every experimental design that uses subtraction or contrast techniques, which constitute nearly all of the studies published in the first few decades. For example, an early, typical word comprehension study asks listeners to process words (for example, nouns) in task one, and then presents nonsense words in task two. The analysis is done by subtracting signals derived from task two from task one, and/or vice versa. In the meaningful words task, this presupposes that nonsense words are the “same” (to the neurophysiology of the brain) as natural words, only missing a feature of meaning; it further presupposes that the brain organizes “meaning” or “meaningful words” in an array of brain areas and “word shapes” or “nonsense words” in a subset of this array. But nonsense words not only lack conventional meaning, they are weird, bizarre; they may resonate during cognitive processing in interesting and unpredictable ways; they may stimulate Klang associations and mental images; and if there is little understanding of how words and/or meanings are processed in the brain, there is surely no good presupposition for what a particular brain will do with nonsense words. Therefore, taking the conglomerate of mental activities surrounding perception of a nonsense word and subtracting it from mental activities associated with a legal word in a language comprehension embraces compelling logic from one perspective, but from another, it is arguably absurd.

Variations on this theme pervade the imaging literature. More complex subtractions, or hierarchically based subtractions as they have been called, and especially those involving higher-order brain association areas, as pointed out by several authors (Bookheimer et al., 1997; Demonet, Wise, & Frackowiak, 1993; Friston et al., 1996; Jennings, McIntosh, Kapur, Tulving, & Houle, 1997; Poeppel, 1996a; Sidtis, Strother, Anderson, & Rottenberg, 1999; Small & Nusbaum, 2004; Wise, Hadar, Howard, & Patterson, 1991), are doubtless even less reliable. In characterizing the normal heuristics in scientific process of decomposition and localization (examination of the properties of portions of a system), Bechtel and Richardson (1993) describe in numerous ways how dynamic systems are seldom aggregative.

There are other disagreements with this policy and practice: arguments against looking for localized modules in the functioning brain/mind (Fodor, 2000; Uttal, 2001); the
requirement of a logical structure of “linkages” between scientific concepts or law and data of observation (Carnap, 1966; Nagel, 1961); the need for procedures for arriving at quantitative concepts (values) (Hempel, 1952). Yet many other language studies follow the same protocol, although with more complex and detailed pieces of an alleged language model. The difficulty is that the subtraction analysis further distorts elements and processes of a weakly validated model of language (Poeppe, 1996b), a model that has shown to have poor correspondence to any behaviors derived from neurological conditions, and independent constraints are lacking.

This problem is compounded by the arguments provided in this special issue that the current understanding of signals in functional brain imaging is inadequate. This adds up to using a theoretical model of language that has not been particularly fruitful in classical, structural neurolinguistics, being used to map brain representation as reflected in unconstrained, uncorroborated measures derived from the poorly understood signal characteristics of functional imaging data.

8. Hemispheric specialization research

During 20th century and before (Harrington, 1987), cerebral laterality studies flourished in neurolinguistic research and the several related disciplines (neurology and behavioral neurology, neuropsychology, electrophysiology, neurosurgery, psychology, communication disorders) (Bradshaw & Nettleton, 1983; Critchley, 1962; Espir & Rose, 1970; Friederici, 1999; Gazzaniga, Ivry, & Mangun, 2002; Geschwind & Galaburda, 1985; Loring, Meador, Lee, & King, 1992; Milner, 1980; Ojemann, 1983; Penfield & Roberts, 1959; Sperry, 1974: Zangwill, 1960). Progress in understanding was made. Numerous design methodologies were applied. Several assumptions served as firm foundation to this approach: the cerebral hemispheres tend to specialize in performing cognitive functions (Bever, 1975; Bogen, 1969; Cutting, 1990; Kupermann, 1991); most normally developing people show a similar profile of hemispheric specialization, and language, particularly phonology and syntax, in nearly all normal cases, as shown over and over by converging evidence from the many methodologies of neurolinguistic research, is processed in the left cerebral hemisphere (Jackson, 1874). Speaking is the “hallmark” of the left hemisphere, and language comprehensiveness of the right hemisphere is severely limited (Gazzaniga et al., 2002, pp. 413–414).

Ordinary language function, later identified more specifically as “ortholinguistic” competence, was held to be “intact” following right hemisphere damage (Lenneberg, 1967; Gardner, Winner, & Rehak, 1991).

Right hemisphere language representation was considered highly anomalous, such that an encounter with a patient with aphasia following a right hemisphere lesion could lead to a scientific article. A recent review (Marien, Paghera, De Deyn, & Vignolo, 2004) documented 180 such cases since 1975. Other occasional clinical observations challenged the accepted tenets about innate left hemispheric specialization of language function and localization of function in general (Marien, Engelborghs, Vignolo, & De Deyn, 2001), such as normally functioning individuals with only a thin rim of cerebral tissue due to extreme (more than 90%) hydrocephalus (Lorber, 1983; Lewin, 1980) or with one hemisphere (Bassser, 1962; Smith & Sugar, 1975; Van Lancker Sidtis, 2004b; Code, 1997; Vargha-Khadem & Mishkin, 1997). These rare but compelling exceptions kept the field vibrant, provoking studies of genetic influences (Marcus, 2004) and brain plasticity (Alexander & Annett, 1996). These debates have kept alive questions about how necessary and sufficient is extent and localization of neural tissue for behavior. But because of the predominance of converging evidence gathered from diverse sources using an array of methods—although many details remain unresolved—the basic idea of “left lateralization of language” has remained intact.

This earlier view of left hemisphere dominance for all of language function has seen modification in the past few decades. Now more is known about the right hemisphere contribution to the pragmatics of language (see overview by Van Lancker, 1997), sometimes termed, to provide a contrast to left hemisphere language function, “paralinguistic” competence. The broad domain of pragmatic functions, including nonliteral and emotional language (Van Lancker & Cummings, 1999), formulaic expressions (Van Lancker Sidtis & Postman, 2006; Van Lancker Sidtis, 2004a, 2006), inference (Foldi, 1987), theme, humor, and contextual relations (Joanette & Brownell, 1990; Myers, 1998; Young, 1983), prosody (Pell, 2006; Sidtis & Van Lancker Sidtis, 2003) and the like, remains separable in hemispheric specialization studies from the domain of “ortholinguistics,” the domain of language ability attributed to left hemisphere function (see Table 1). With these modifications that have been viewed by most workers in the neurolinguistic fields as advances, models of hemispheric specialization and lateralization of function have remained relatively stable, consistently allocating definable characteristics of language function to left or right hemispheres.

Functional neuroimaging has tipped over the laterality apple cart. In what way? Because for nearly every functional imaging language study—whether hearing and producing speech, picking out phonemes, selecting, classifying, or judging words or sentences, hearing stories or syntax—analysis reveals significant right hemisphere activation sites, sometimes “more” or “greater” than left hemisphere sites. A popular textbook refers to bilateral signal in neuroimaging studies of speech and nonspeech stimuli alike as “a big surprise” (Gazzaniga et al., 2002, p. 364). A few specific examples (a partial list) are seen in significant right hemisphere imaging signal reported for speech stimuli or tasks.

---

5 This term was suggested by J.E. Bogen to include phonetics/phonology, syntax, and linguistic semantics, and to designate these aspects separately from “paralinguistic” phenomena.
(Giraud et al., 2004); single words (Chee, O’Craven, Bergida, Rosen, & Savoy, 1999; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Roland, 1993), covert naming (Spitzer et al., 1998), animal naming (Laine, Rinne, Krause, Teras, & Sipila, 1999; Van Lancker, McIntosh, & Grafton, 2003), syntax (Indefrey, Hellwig, Herzog, Seitz, & Hagoort, 2004; Moro et al., 2001; Wartenburger et al., 2004), free speech (Tamas, Shibasaki, Horikoshi, & Ohye, 1993), syntactic rule learning (Tettamanti et al., 2002), irregular verbs (Beretta et al., 2003), pseudoword sentences (Friederici, Meyer, & von Cramon, 2000), silent word generation (Friedman et al., 1998); hearing speech (Giraud et al., 2004), semantic and phonological processing (Poldrack et al., 2001), consonant-vowel syllables (Jäncke, Wüstenberg, Scheich, & Heinze, 2002; Sidtis, Strother, & Rottenberg, 2003; Sidtis et al., 1999; Sidtis, Gomez, Groshong, Strother, & Rottenberg, 2006), animal semantic categories (Laine, Rinne, Hiltunen, Kaasinen, & Sipila, 2002), continuous speech, (Kircher, Brammer, Levelt, Bartels, & McGuire, 2004), meaning/syntax anomalies (Kuperberg et al., 2003), listening to stories (Papathanassiou et al., 2000; see Cabeza & Nyberg, 2000); for a review. A review of syntax processing studies in the past decade or so graphically displays a ratio of “activation” sites on the medial brain surfaces of left/right hemisphere of about 2/1, which the authors evaluate as “interesting” (Kaan & Swaab, 2002, p. 355). Having grown up with the classic laterality model of language, the neurolinguistic practitioner is perplexed.

Numerous papers appear with this presentation: more than, as much as, or half as much right hemisphere as left hemisphere activity for tasks long presumed to represent competences that are ortholinguistic in nature (e.g., speech, phonology, and syntax). In one other selected example, nearly half the reported sites for phonological tasks reported by Jacquemot, Pallier, LeBihan, Dehaene, and Dupoux (2003) appeared on the right hemisphere, but the authors state that the right hemisphere is involved “to a lesser extent.” Often no explanation of this frank deviation from the standard view of brain function for language is provided. Other authors have offered an array of speculative explanations accounting for right hemisphere involvement in their ortholinguistic tasks, including a putative role of prosody (Constable et al., 2004), on-line presentation (Clark & Wagner, 2003), left-handed button press or use of multiple talkers in stimuli (Callan, Jones, Callan, & Akahane-Yamada, 2004), cortical control of the diaphragm (Tamas et al., 1993), memory, long term memory, blocking mechanism, behavioral inhibition, or inhibiting application of a default rule (Beretta et al., 2003; Pihlajamäki et al., 2000), working memory (Heim & Friederici, 2003), interindividual variability of language dominance (Papathanassiou et al., 2000), coactivation when cognitive tasks become complex (Demonet et al., 1992), timing issues, complexity and length of sentence (Kaan & Swaab, 2002), speed of processing (Waters et al., 2003); attention shift (Fiez et al., 1995), auditory “anticipation,” (Wise, Greene, Buchel, & Scott, 1999) and listening to ones own voice (Price et al., 1996). On the other hand, a lack of right hemisphere activation in silent reading of metaphors is handled by an array of explanations from stimulus properties to task instructions (Rapp, Leube, Erb, Grodd, & Kircher, 2004). A failure to observe significant activity for both propositional speech and nursery rhymes in Broca’s area is attributed to the possibility that both tasks are automatic; counting showed a different activation allegedly due to having actual “syntactic structure” (Blank, Scott, Murphy, Warburton, & Wise, 2002). To account for widely distributed brain signal, a vaguely defined “network” notion is sometimes adduced (e.g., Demonet, Fiez, Paulus, Petersen, & Zatorre, 1996; Kaan & Swaab, 2002). This cornucopia of explanatory commentary adds up to more than an embarrassment of riches: it takes us into the unconstrained world of too many possibilities.

Early reports of bilateral regional cortical blood flow in automatic speech tasks (Larsen, Skinhoj, & Lassen, 1978; Ryding, Bravvik, & Ingvar, 1987), originally a good fit with a prevailing model, now have no meaning. The consistent appearance of right hemisphere signals in so many language studies of all kinds leads the reader to be wary when the right hemisphere is apparently not examined or mentioned in a language study (Caplan, 2001; Caplan, Alpert, Waters, & Olivieri, 2000; Crosson et al., 2001; Davis et al., 2004; Mason, Just, Keller, & Carpenter, 2003; Stromswold, Caplan, Alpert, & Rauch, 1996; Okada, Smith, Humphries, & Hickok, 2003), or, more ominously, where tables of activation in the body of the paper list right hemisphere sites, but this information is not included in the abstract or interpreted in the discussion. A meta-analysis, derived from 129 scientific reports on imaging in language tasks (Vigneau et al., 2006), proposes a model of left hemisphere language areas, coping with broad overlap of the 730 “peaks” analyzed for phonology, semantics, and “sentence” processing, expressly leaving right hemisphere results unconsidered. But are we then to assume that a coherent language model would emerge in like manner for reported right hemisphere “peaks”? Vigneau et al. (2006) posit the “close proximity” of the “site of the human voice” (p. 14) to language activation areas as support for their model, citing Belin, Zatorre, Lafaille, Ahad, and Pike (2000). But those cited authors (see also Belin & Zatorre, 2003) report greater or primary activation for human voice signals in the right temporal lobe.

The general impression is that it is not understood why there is significant brain activation in the right hemisphere for classically linguistic tasks. What have neurolinguists long considered quintessential left hemispheric linguistic tasks? Speaking, for one (mentioned previously as the “hallmark”), and phonology and syntax for two others. A current handbook states definitely that “generative syntax is present in only one hemisphere” (Gazzzanga et al., 2002, p. 411). A recent European review of neurolinguistic research in syntax attributes syntactic processes to “areas in the left perisylvian cortex” (Hagoort, Broon, & Osterhout, 1999, p. 305). Clinically, except for the 1% of dextral stroke patients with crossed aphasia (Marien et al., 2001), people do not have problems with speaking (other than occasional
transient dysarthria), phonology, or grammar following unilateral right hemisphere damage. Therefore, when bilateral frontal and parietal activation sites for phonological processes are alleged to be represent “underpinnings of phonological assembly” in regions alleged to be “previously associated with phonological control and on-line phonological representation” (Clark & Wagner, 2003, p. 313), the neurolinguist, familiar with longstanding traditions and experienced in clinical presentation, is skeptical. Observation of the dramatic failure of self monitoring in the fluent aphasic speaker, who generates neologisms, paraphasias, and jargon with no apparent awareness of these gross deviations from intelligible speech, leads to the question: where are the right frontotemporoparietal language areas reportedly “activated” in functional imaging studies in this type of disordered speaker? When irregular but not regular verbs in German are associated with right hemisphere activation (Beretta et al., 2003), and the clinically experienced neurolinguist cannot remember anyone with an impairment in producing or understanding a specific grammatically classed subset of verbs to the exclusion of others following right hemisphere brain damage, there is restiveness. The intellectual or clinical context to accommodate this kind of reported neuroimaging result is often lacking.

If the right hemisphere activation results reported in functional imaging studies are not readily interpretable, and are handled with a range of unconstrained explanations, how can the other reported signals—the multiple sites of “activation,” also not consistent across other imaging studies, often not consonant with previously believed notions of localization of function, be believed and assimilated? How, then, too, can the reader confidently consider reported findings of right hemisphere activation for pragmatic functions (Bookheimer, 2002) or metaphor (Bottini et al., 1994) as valid? As mentioned above, right hemisphere activation has already been reported for nearly every other language task. In Blank et al. (2002), bilateral activation was reported for three kind of speech tasks (one “propositional” and two “nonpropositional”—nursery rhymes and counting), but, in this case, no right hemisphere signal was reported for the two nonpropositional tasks, despite a literature that predicts such a finding (Code, 1997; Jackson, 1874; Van Lancker Sidtis, 2006). A consumer’s pique is well displayed when Coltheart (2000) accuses Price et al. (1998), who deny in their article that the right hemisphere signals in their study reflected right hemisphere reading, of misinterpreting their own results.

Rich interpretation, the practice of taking unconstrained data and launching into a world of theoretical possibilities, has been a common practice in functional imaging studies of language. Previous facts obtained from brain lesion studies, long the international standard for theories of brain-behavior relations, are either ignored, or selectively brought to bear on imaging results with a free reign. Interpretation of functional brain imaging results often manifests the virtues of the New Criticism School of literary analysis, in being precise and systematic, and aspiring to describe an organic unity of the disparate elements (Ransom, 1937; Richards, 1929), but the use of so much imagination in this venue leaves the reader a bit breathless.

Previous scholarship, especially in the earlier imaging papers, has often been lacking, as has already been lamented (Coltheart, 2000; Poeppel, 1996a). Many brain imaging studies appear to spring forth on a blank landscape of brain research, and many mainly reference each other (Fellows et al., 2005). Even the most brilliant and productive imaging researchers can err in this way, for example, asserting that right hemisphere communicative processes “above that of literal meaning of words and sentences”… have “received only a little attention in the lesion literature” (Bookheimer, 2002, p. 174). Actually there is a prodigious body of lesion literature on right hemisphere communicative processes since the 1970s, pursued with by such individuals as H. Gardner, H. Brownell, Y. Joanette, A. Young, J. Bradshaw, N. Foldi, E. Winner, A. Ellis, D. Kempler, C. Tompkins, J. Cutting, and P. Myers.

There are several plausible explanations for this state of affairs—for the practice of rich interpretation by the scientists, for the inconsistencies and multiple functionalities across reported studies (e.g., Müller & Basho, 2004; Poeppel, 1996a), for skepticism in the reader, for the failure to successfully relate imaging findings to other established neurolinguistic hypotheses and models. The first reason, one that has been mentioned in the imaging literature and is explored in this special issue, is that it is not really known what the signals mean. Viewed from notions of scientific structure, the brain activation signal falls short of certain standards expected of a dependent measure (Small & Nusbaum, 2004). Scientific statements are expected to have determinacy in the form of specific relational or structural properties that allow for explanatory hypotheses (Nagel, 1961). This entails that there are rules of correspondence linking theoretical ideas with experimental concepts. Proper measurement requires “a kind of isomorphism” between the empirical relations and the “properties of the formal game” used to quantify them (Stevens, 1951). In imaging, it is not known whether “activation” (the significantly identified signal) reflects brain incompetence (working harder) or competence (specialization) with a task. It is not known whether “more” or “stronger” activation means less competence at the skill (as suggested by Luke, Liu, Wai, Wan, & Tan, 2002), or more competence, as generally assumed; or whether activation means more or less involvement, or inhibition, or predilection to process, or relative automaticity of processing (McDermott, Petersen, Watson, & Ojemann, 2003; Wise et al., 1999), or some kind of passive working-in-tandem with actually responsible brain sites, or something completely different, or whether it has no inherent relationship to functional significance (Sidtis et al., 2003). In the domain of functional brain signals, consistent dependence of quantities and relations of things on other relational or structural processing characterizing a class of objects—the required condition for constructing an explanation—is missing (Nagel, 1961, p. 11). A study of
sentence comprehension describes the “complexity effects associated with posterior regions” as not as “consistent or large in magnitude” as in anterior areas (Constable et al., 2004), as though a correspondence existed, but what is the correspondence? What does a comparison of signal strengths mean, even within one data set, much less across studies? This implies like difficulties with the meaning of “abnormal” (Belin et al., 1996) or “decreased” activation (Demb et al., 1995).

Regarding concepts that figure in the evolution of a branch of science, it is difficult to see whether brain imaging signals lend themselves more successfully to classificatory, comparative, or quantitative scales (Carnap, 1966; Stevens, 1968). Of course, there is a deeply quantitative history in the way the signals are obtained and displayed. But how to classify, and/or how to compare the signals has not been worked out. It appears that the developmental trajectory for functional brain imaging has been upside down: a great deal of quantitative effort enters into the technological base, but the results as dependent measure barely achieve classificatory status, as the signals cannot be usefully named or placed into classes; they do not conceptually achieve the comparative conceptual level, as the meaning of “less” or “more” brain activation (or is it inhibition?) is unclear.

The indeterminacy of the signals is compounded by and complicitous with inadequately formulated questions, and by the imposition of complicated tasks to ask the questions (Small & Nusbaum, 2004). The problems with functional imaging might hammer another nail into the coffin of the likelihood that our posited language levels and categories have any correspondence in the “language” of the brain. The use of compound subtraction designs derived from current models of language, although this approach represents a respected heuristic of decomposition and localization in science (Bechtel & Richardson, 1993), often generates an uninterpretable kaleidoscope of patterns. This is because, first, the models are inadequate, and secondly, the meaning of the functional brain imaging signals is not understood.

Regarding neurolinguistic models, perhaps something very different from our conceptions is happening in the brain processing. It is often remarked, somewhat dismissively, that automatic machine speech or voice recognition achieves a recognition goal in ways completely differently from how humans process the speech signals. The methods of engineering depart entirely from the orderly bottom-up, top-down interactive process models of psycholinguistics (e.g., Levelt, 1999). Different engineering groups achieve their goals in different ways. Perhaps the brain achieves a cognitive goal in ways totally different from anything any of our abstract theoretical models posit, name, or propose.

With indeterminacy about the meaning of the functional brain imaging signals, and only weak independent validation of the linguistic models from structural brain imaging studies, the research efforts in functional imaging lack constraint on either end. This state of affairs brings to mind an academic story: At the beginning of each semester, a science professor announced to his students that half of the material in the assigned textbook, which he authored, was wrong, but, unfortunately, he did not know which half. Numerous writers have suggested that the field of neuroimaging has led to a revision of notions of language processing in the brain (e.g., Marcus, 2004; Pulvermüller, 2002; Stowe et al., 1994). Obviously, many suggestive results and provocative findings about human language are reported in the functional imaging literature, but due the lack of grounding at either end—either theory or dependent measure—evaluating these reports is not straightforward; we would like to believe it, but we do not know which half.

Solutions at both ends—what neurolinguistic questions are asked, and how the imaging data are handled—are desirable. One approach is to ground both in behavioral measures, using performance-based analyses (Sidtis, Anderson, Strother, & Rottenberg, 2001; Sidtis, 2006, this issue). Analyzing imaging data with respect to performance rather than using a contrast design establishes a more direct link between the behavior and the dependent variable, the signal (Sidtis, 2006, this issue). If task contrasts are used, the results can still be analyzed with respect to performance (Riecker, Kassuabek, Grüssel, Grodd, & Ackermann, 2006), but important information may be lost in the contrast (Sidtis et al., 2003, Sidtis, Strother, & Rottenberg, 2004). Tying imaging results to performance measures in a predictive manner provides an actual quantitative concept to the dependent variable, the signal.

A study comparing PET and fMRI “increases and decreases” found fair concordance between the two methods (Votaw et al., 1999), although the cognitive meaning of the signal remains uncertain. Another fruitful approach lies in the set of studies correlating imaging results with data from other methods (Demonet et al., 1993) such as electro-cortical stimulation (FitzGerald et al., 1997; Friederici, 1999) and Wada testing (Hunter et al., 1999). This grounding in observable behaviors, which themselves are not necessarily bound to an a priori theory or category, and effort at correspondence with other, better understood techniques, may help to realize the tremendous potential of functional brain imaging.

Over the course of the past few decades, important questions, and facts that can serve as constraints on data interpretation, have emerged. A few such points are how to assess brain responses to well-established functions, such as sensory and motor processes (Catani, Jones, & Ffytche, 2004; Formisano et al., 2003), how to distinguish inhibition from activation (Naeser, Martin, Baker, Hodges, & Sczerzenie, 2004); how to move from identifying separate sites to recognizing functional brain systems (Eidelberg, 2006, this issue; Sidtis, 2006, this issue), how to evaluate graded activation responses in the brain, what is the meaning of deactivated brain signals, and so on. The variety of responses seen in Broca’s area demand revision of previously held notions (Heim, Opitz, Müller, & Friederici, 2003; Müller & Basho, 2004), and may lead to better understanding of language function. Identifying coherent networks underlying cognitive functions will
continue to be challenging (Fuster, 2003; Lambon Ralph, Sage, & Roberts, 2000; Mesulam, 1990). Studies of language rehabilitation, using corresponding measures of recovery of aphasic patients to provide corresponding measures, hold promise for theory and practice in clinical science (Weiller & Rijntjes, 1999).

9. Conclusion

Functional imaging results for speech and language tasks reveal an array of brain sites, many of which are in conflict with long held assumptions about brain/language relationships. Many are inconsistent across and between like studies. Various sources of these discrepancies have been proposed: distortions and loss of information due to the subtraction and contrast analysis technique, lack of correspondence rules between the brain signals and other consistent measures, the signal as a dependent variable that is unduly qualitative or not well defined, and insufficient or flawed models of language. One concludes that the neurolinguistics of functional brain imaging is in its growth phase. A major proposal in this special issue advocates an experimental design that includes clearly delineated constraints that anchor the imaged signals in subjects’ performance data by finding defined and replicable correspondences.

It is likely, as many believe and some champion, that functional imaging will lead to a view of language processing in the brain that neither the neurological-aphasiological, nor the cognitive neuropsychological models could ever do, and that classic neurolinguistic studies, despite the benefit of reliable, structural brain measures, fell short of doing. Functional imaging studies may have added momentum to the growing recognition that many of the questions fomenting within neurolinguistics require reformulation. Better understanding of the functional imaging signals as dependent variables, coupled with the constraining influence of reliable performance measures, will aid in establishing viable neurolinguistic theories. Better theories, in turn, will provide clearer direction to functional imaging research.

Acknowledgments

Dora Katsnelson assisted capably with editing. The several people, who generously commented on this article, including three reviewers, remain anonymous, as I am responsible for its content. I regret if any offense is taken. It was my intention to express selected responses of consumers of functional imaging research in the neurolinguistic community. I honor and respect the hard work and dedication of the scientists mentioned here, and I hope for us all that the quality of mercy is not strained.

References


