Imaging the Grammatical Brain

Yosef Grodzinsky
Department of Psychology
Tel Aviv University and
Aphasia Research Center
Department of Neurology
Boston University School of Medicine
Yosef1@post.tau.ac.il

1. INTRODUCTION

What do students of language do? Linguists seek a characterization of the nature of linguistic knowledge; psycholinguists are after modeling the algorithms that implement this knowledge when language users speak and understand; and neurolinguists are interested in neural mechanisms that realize these algorithms, and their cerebral localization. One can imagine a research program in which these perspectives all cohere, in an attempt to understand and account for knowledge of language, its acquisition, processing mechanisms, and neural computation. This is the neurobiological project to characterize human language. This chapter is about attempts to reconstruct an image of language mechanisms through the analysis of lesion data (mostly from aphasic syndromes), and through functional neuroimaging. I will attempt to demonstrate that a correct choice of unit for functional analysis of behavior leads to a clearer image of the linguistic brain.

2. PROBLEMS IN THE IMAGING OF GRAMMAR

The project to discover the neural basis of language is ambitious, but not new: the interest in brain/language relations is age-old. News comes in the form of innovative technologies that make the quest for a coherent, focused picture of language more within our reach than ever before. Linguistic theory provides sophisticated technology for the analysis of the linguistic signal; instruments that measure neural activity are less invasive than ever, and their resolution in both time and space is high; and with some additional experimental ingenuity, new solutions to old problems (and to new ones, too) can be found.

   This research enterprise must thus define brain/language relations in the form of an equation, both sides of which contain complex terms: on the one side there is linguistic behavior, described in the best theoretical vocabulary one can find, and on the other side there are brain mechanisms, accounted for by whatever neuroscience can offer. The exact relation between the sides is also extremely complex. And it is here that disagreements arise. Some neurolinguists study words, others sentences, and yet others investigate not linguistic units, but activities, such as speaking, listening, reading and writing. It is quite difficult to home in

on a unit of analysis on which consensus (one that would hopefully reflect understanding) exists. In this respect, the study of language is unique. Compare it to the study of the visual system – a research program that has been – uncontroversially, I believe – a success story. Debates in vision exist, yet none regards the basic unit of analysis. In low-level vision, lines are lines, angles are angles, and edges are edges – elemental parts that quite clearly play a constitutive role in forming our visual experience. Likewise, in visual object recognition (where a heated controversy is currently taking place) the basic units of analysis – objects organized in hierarchically structured categories – also seem consensual. In the study of language, by contrast, little is agreed upon. This weakness threatens to hinder the effort to image the neural basis of language.

   The present perspective sees linguistic capacity as critically involving the pairing of sound sequences and meanings, aided by inventories of combinatorial rules, and stores of complex objects of several types, over which these rules operate as language is practiced through its various modalities. An examination of the current state of knowledge, based on traditional and new techniques and approaches, indicates that this is the best approach. The remainder of this chapter is an attempt to make some order in the available sea of data, in an effort to formulate a coherent research program that will hopefully lead to a precise characterization of the neural basis of the human language faculty. For reasons of space, the focus will be on linguistic combinations at the sentence level, leaving aside fairly successful areas of research such as the cerebral representation of word structure and meaning, phonology and phonetics, and compositional semantics (e.g., Phillips *et al*., 2001; Posner & Pavese, 1998).

3. RECONSTRUCTING THE IMAGE OF GRAMMAR: LESION DATA

Some neuropsychology first, briefly: The old Connectionist school – founded by Broca, Wernicke, and Lichtheim, and revived in our time by Geschwind (1979) – fortified belief in the existence of cerebral language centers. As clinicians, they mostly emphasized patients' communicative skills, viewing language as a collection of ***activities***, practiced in the service of communication: speaking, listening, reading, writing, naming, repetition, etc. The characterization of the language centers derived from this intuitive theory – each activity was associated with a cerebral locus. The resulting theory of localization uses these activities as building blocks and takes them as the essence of human linguistic capacity.

Against this background, psycholinguists have since the 1960s attempted to devise a new perspective. They challenged the old outlook on the centers, on the basis of theoretical and experimental tools borrowed from linguistics and psycholinguistics (e.g. Goodglass & Berko, 1960; Zurif & Caramazza, 1976). Not denying the validity of the old approach, they focused on the distinction between linguistic levels of analysis, and placed new issues at center stage. Language was no longer a set of activities, but a structure-dependent piece of knowledge, divided into ***levels of representation***, mainly phonological, syntactic and semantic. A variety of 1970s' experiments proved this approach worthwhile. Surprising results were obtained, showing that the brain makes distinctions between types of linguistic information. Such results could not be couched in the standard view, and an attempt was thus made to "redefine" the centers (Zurif, 1980): the anatomical, center-based conception was retained, yet each center was now said to contain devices used for the analysis and synthesis of language, rather than activities. Roughly, Broca's region (Brodmann's Area BA 44,45,47) was said to house syntax (harnessed in the service of both receptive and productive language), while semantics was assumed to reside in Wernicke's area (BA 22, 42, 39). Neuroanatomy also witnessed

parallel advances. As large samples of patients became available, it became increasingly clear that language occupies larger areas than previously supposed. As analytic and experimental tools improved, the involvement of both hemispheres in aspects of linguistic activity was documented (cf. Ojemann, 1991).
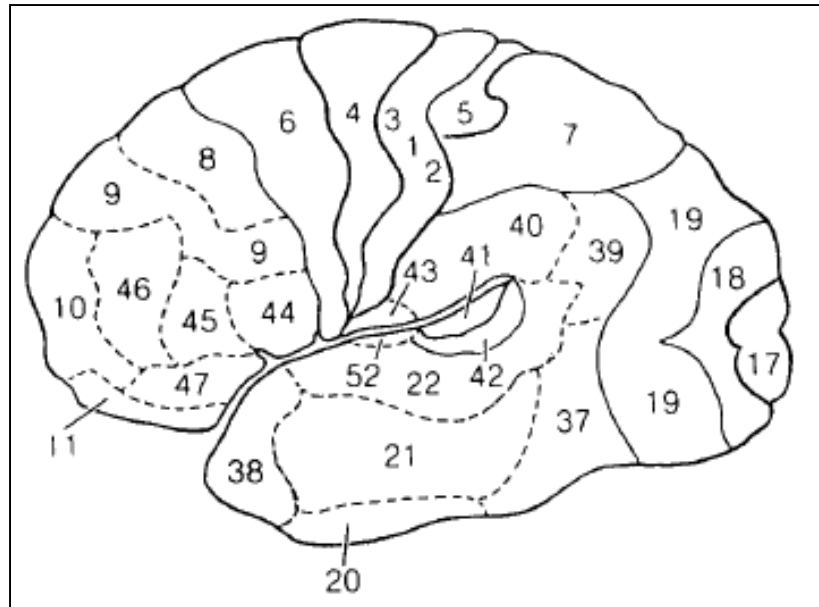

Figure 1: Brodmann's division of the left cortical surface into areas

Yet as findings accumulated – from different tasks, languages, stimulus types, and laboratories – contradictions within the behavioral data began to surface: In some cases, Wernicke's aphasics showed syntactic disturbances; Broca's patients, on the other hand, while failing certain tasks that probe syntactic abilities, succeeded in others. Serious doubts were cast on the new model, in which Broca's area (but not Wernicke's) supports receptive syntactic mechanisms. Subsequent attempts to reconcile the findings with the prevailing view were called for. So it was claimed that regions are organized not just by activities and linguistic levels, but also by *tasks*. Claims like "syntactic comprehension is compromised", and "grammaticality judgment is intact" began to be heard.

Upon examination, a common thread could be noticed in these analyses: while being detailed in the description of tasks and activities, they were all rather "holistic" in their approach to the linguistic signal. Gross distinctions between form and meaning seemed sufficient, and hence, less attention was paid to detailed structural properties of linguistic stimuli. Still, the amended neurological model of language could continue and prevail.

An exception to this description is the study of phonology and phonetics (cf. Blumstein, 1994). In these areas fine, theoretically-motivated distinctions have long been used, and landmark discoveries of subtle distinctions that the brain makes among informational types have been made. Students working in other domains of language, however, were slower in making connections between matters neuropsychological and linguistic. Yet, when it turned out that the task-oriented approach was incorrect, the next move was to try and argue that the inconsistencies in results discussed above were just apparent, due our failure to make distinctions among linguistic types. It was argued that systems of grammatical knowledge are complex, and as such, can experience partial

breakdown subsequent to focal brain damage. The next step, then, was to seek linguistic frameworks within which patterns of impairment and sparing in aphasia could be couched, and which in turn would give rise to a finer-grained theory of brain/language relations. This resulted in an investigation into the cerebral localization of ***grammatical rule systems***. It was shown that despite the importance of channels through which language is practiced, the correct (in fact most telling) unit of analysis for the interpretation of lesion data is the particular rule type.

It is here that considerations pertaining to the structure of language began to matter heavily. It seems uncontroversial that the most important aspect of our "mental organ for language" is its combinatorial nature, that is, the knowledge base and algorithms for the concatenation of linguistic sequences at all levels – phonetic, phonological, morphological, syntactic and semantic. Activities and tasks no doubt play a mediating role in linguistic communication, yet the defining characteristic of the language faculty is its being composed to rule systems. How these rule systems are instantiated in neural tissue thus seems the central question in neurolinguistics.

Thus it was shown that in the domain of language production, the brain makes fine distinctions among rule types: Broca's aphasics are deficient in producing Tense inflection, but intact in Agreement inflection (as shown in a wide variety of languages). Cross-linguistic studies further indicated that not only inflection type, but also, the position of the verb in the sentence determines its appearance subsequent to a lesion in Broca's region (Friedmann, 1998). In receptive language, the distinction between transformational and non-transformational sentences, yields a big performance contrast: aphasics with lesions in Broca's region understand active sentences, subject relatives, subject questions and the like normally, yet fail on their transformational counterparts: passives, object relatives and questions, etc. This led to the claim that in receptive language, Broca's aphasics are unable to compute transformational relations. This generalization helps localize this grammatical operation in the brain (Trace-Deletion Hypothesis, TDH, cf. Grodzinsky, 2000 for a review). Furthermore, the highly selective character of this deficit has major theoretical ramifications to linguistic theory and the theory of sentence processing. A particularly compelling argument that supports the localization of transformations in Broca's region comes from cross-linguistic comparisons: Chinese, Japanese, German and Dutch, Spanish and Hebrew have different properties, and the performance of Broca's aphasics is determined by the TDH as it interacts with the particular grammar of each language. In English, aphasics comprehend active sentence properly. Yet the finding are different in Japanese, which has 2 types of actives: *Taro-ga Hanako-o nagutta* (Taro hit Hanako*)* – Subject Object Verb], and ***Hanako-o*** *Taro-ga nagutta* – **Object** Subject Verb. These constructions are simple, they mean the same and are identical on every dimension, except in that the latter is derived transformationally, with the bolded element fronted to the left edge of the sentence. Remarkably, Broca's aphasics comprehension splits: the handle the SOV type properly, and are at chance level on the OSV.

In Chinese, an otherwise SVO language like English, (bolded) heads of relative clauses (1a, 2a) follow the (parenthesized) relative, unlike English (1b, 2b) in which they precede it. Remarkably, this reversed order correlates perfectly with the cross-linguistic results in aphasia: subject relatives (1) are comprehended at chance in Chinese and above chance in English, whereas object relatives (2) yield the opposite pattern:

(1) a. [_ zhuei gou] de **mau** hen da   *chance*
        chase dog that cat very big
    b. **The cat** that [_ chased the dog] was very big   *above chance*
(2) a. [mau zhuei _] de **gou** hen xiao   *above chance*
        cat chased that dog very small
    b. **The dog** that [the cat chased _] was very small   *chance*

    English and Chinese thus yield mirror-image results, which correlates with a relevant syntactic contrast between the two languages. Other intriguing cross-linguistic contrasts also exist, providing further evidence that Broca's region is critically involved in transformational analysis. Moreover, reflections of the same disruption are also found in the domain of real-time processing (Zurif, 1995). This rich database is further augmented by results regarding grammatical aspects of the mental lexicon. These are also localizable, as they appear retained in Broca's aphasia, but severely disrupted after a lesion in Wernicke's area.

        In sum, over the past decade or so, a new, intriguingly complex model of grammar/brain relations has emerged: Aspects of receptive syntax – those dedicated to the computation of transformational relations – are represented in Broca's, and to some extent in Wernicke's region; the linguistic lexicon is in the latter region, whereas other parts of receptive syntax, while clearly residing in the left hemisphere, are not localizable as of yet; in productive language, Broca's region is dedicated to extremely limited aspects of structure, that pertain the upper, leftmost end of the syntactic tree (see Friedmann & Grodzinsky, 2000). Most importantly, linguistic tools appear critical for the analysis of brain/language relations.

4. RECONSTRUCTING THE GRAMMATICAL BRAIN: NEUROIMAGING

This story has an important lesson to functional neuroimaging. Early studies that used this experimental methodology grappled with many hard questions, one of which had to do with the choice of experimental materials, determined, to a large extent, by the experimenter's theoretical tastes. One would have expected neuropsychological data to have played a central role in this new effort; in practice, functional imaging of language witnessed an attempt to start almost from scratch. Caught by the excitement that swept the field when neuroimaging techniques were introduced, many investigators have largely tended to dismiss aphasia data, rather than seek cross-methodological convergence. Some important mistakes were repeated as a result. Preliminary studies conducted contrastive investigations of *activities and modalities*. The first ones (e.g., Petersen *et al*., 1990) investigated the production versus comprehension of various linguistic stimuli in PET and then fMRI; and while a distinction was made between overt and covert sentence production, the nature of stimuli – their structure – remained unanalyzed and unspecified. No wonder, then, that anatomical overlap among studies was very limited: verb production versus comprehension, for instance, activated the cerebellum and culliculi for Petersen *et al*., whereas in more recent studies it was localized in the left posterior temporal lobe and the anterior insula bilaterally.

    Early studies were also concerned with *cross language comparisons*, with language once again taken as one unanalyzed whole, leading to great variation in stimuli (and subsequent anatomical variation). Thus Mazoyer *et al*. (1993) conducted a PET investigation of the functional anatomy of sentence comprehension in a known (in fact, native) versus unknown language (French vs. Tamil); other authors looked at PET

activations during the comprehension of active declarative sentence in spoken language, as compared to similar stimuli in sign language, finding multiple activations in the left frontal lobe, as well as in the temporal lobe bilaterally. Still others compared the fMRI activation during the comprehension of native (Japanese) versus second (English) and an unknown (Hungarian) language in the same speakers. Here, some aspects of the frontal cortex was activated bilaterally, whereas Broca's and Wernicke's regions, as well as some neighboring ones, were activated only on the left side. Similarly, the BOLD response in a comparative fMRI study of English sentences versus sentences in Mandarin Chinese resulted in bilateral activations in the inferior prefrontal cortex (BA 44, 45, 47), bilateral middle prefrontal cortex (BA 9, 8 6) and secondarily in the left temporal region (BA 22, 21, 38), the left angular gyrus (BA 39), and bilateral activations in the anterior supplementary motor area (BA 8), the superior parietal region (BA 7), and in some occipital regions. So, while most of these studies demonstrated activation in the left Broca's area and around left Wernicke's area, scattered activations in many more regions – in both the left and the right hemispheres – were also recorded, thus blurring the picture, and making it much less stable than we would like it to be.

Yet, when previous neuropsychological data and linguistic considerations are taken into account, it quite possible that activities or languages may not be the correct units of analysis for a precise characterization of brain/language relations. One possible reason for the lack of anatomical congruence among past studies, then, is that they made incorrect choices of analytic units, and as a consequence, they simply did not use appropriately minimal contrasts in their comparisons. From this perspective, a sentence in sign language is an incorrect control for a condition that contains English sentences – it may be as inappropriate a control as a Mozart symphony would be for a test for visual object recognition. Psychologists have realized that, and as a next step, set themselves to study finer-grained distinctions. Again, following the neuropsychological tradition, some have attempted to test distinctions among *levels of linguistic description*. Yet here, too, localization has been somewhat disappointing.

Friederici and her colleagues conducted a series of studies which also contrasted syntactic with semantic variables, and sought neural correlates for it, as monitored through MEG. In one study they tried to localize syntactic processes through the measurement of magnetic response during auditory exposure to "syntactically correct" and "syntactically incorrect" sentences. They found that "early syntactic parsing processes" activated temporal regions, possibly the planum polare, as well as by fronto-lateral regions. They further comment that "the contribution of the left temporal regions to the early syntactic processes seems to be larger than that of the left fronto-lateral regions." Friederici (2000) summarizes the results from PET and fMRI studies, "the posterior region of the left superior temporal gyrus and the adjacent planum temporale is specifically involved in auditory language comprehension"; there is also "an involvement of left inferior frontal regions in phonetic processing," and for syntax there is "maximal activation in the left third frontal convolution…but additional activation in the left Wernicke's area as well as some activation in the homotopic areas in the right hemisphere". Another experiment this group conducted sought to dissociate the phonological, semantic and syntactic subsystems. They presented active declarative sentences along with sentences with the same syntactic "frame" but with nonsense words, and with unstructured word lists and non-word lists. When sentences with real and

pseudo words were compared to word and non-word lists (as a reflection of syntax), certain bilateral temporal, parietal, frontal and subcortical areas were activated.

In a similar vein, Dapretto & Bookheimer (1999) tried to dissociate syntax from semantics through fMRI. They asked subjects to make same/different judgments on sentence pairs of 2 types: one involving the same sentence structure but with one different word; another involving same meaning but different sentence structure (active versus passive). For both the semantic and syntactic comparisons, they report Broca's region and its vicinity (BA 44, 45, 47) and the superior and middle temporal gyri (BA 42, 22, 21) bilaterally as the main activated area in the comparison, with some more activation on the left for the syntactic comparison (BA 44).

The reader may have noticed that here, too, the anatomical overlap between studies is not very promising. Again, Broca's and Wernicke's regions are activated, providing support to the view – originating in Broca's and Wernicke's writings – that these regions are crucial parts of the language faculty. Yet this is not enough: other regions are activated in a non-overlapping manner, and we must try and understand what this may mean. Three interpretations are imaginable: either language is widely distributed in the brain, and moreover, linguistic representations are unstable, varying from one individual to the next in a manner that affects findings, or the available imaging technology is unreliable, or experiments do not test what they purport to test.

My own tendency is optimistic, leaning toward the third possibility: while there is clearly individual variation in the precise size, location and structure of the language areas (cf. Amunts *et* al., 1999 for compelling cytoarchitectonic evidence), brains appear to be relatively stable in what they represent. A large amount of functional variation and spreading, I would argue, is a consequence of the great variation among experiments at this point, caused mainly by an insufficiently refined view of linguistic structure. The fact that "syntax", "phonology" and the like are undifferentiated is likely an important reason for the wide range of anatomical loci imputed to sentence processing. A linguistic perspective – especially one that seeks to account not just for the functional imaging data, but also the rich body of knowledge that comes from lesions studies – might make matters more uniform.

An attempt to be more detailed psycholinguistically has been made by Caplan, Just, and Carpenter. These groups have attempted to view language processing in the brain from the point of view of the putative processing difficulty of different sentence types. To take some salient examples, Just *et al*. (1996) looked at the comprehension of 3 sentence types in fMRI, and found that they all activated left and right Broca's and Wernicke's regions, yet the magnitude of the effect grew with processing difficulty; using PET, Stromswold *et al*. (1996) showed differential activation in left Broca's region for differentially difficult relative clauses; and study by the same group conducted a PET study with similar materials, yet with a slightly different task, and found activations The centromedian nucleus of the left thalamus, the medial frontal gyrus, Broca's area, and the posterior cingulate gyrus. Differential processing difficulty, used as a marker that delineates the language faculty, again results in poor overlap. The similarity in the questions posed by most of these studies, suggests that discrepancies in anatomical findings may either be due to different imaging devices, or choice of tasks and materials. Still, experience with the linguistic interpretation of lesion data leaves one with a gnawing sense that systematic linguistic description of functional imaging results – and

subsequent planning of linguistically motivated experiments – is somewhat lacking. In the case of complexity, the blurred picture may well be due to the fact the linguistic complexity is not a well defined notion, and its varying construal affects the nature of experimental materials and the analysis and interpretation of results. It is perhaps advisable to go back to aphasia, and try and find some hints there. The strong link between grammatical transformations and the language areas may be a good place to start, if we seek to tease particular components *within* the grammar apart from others.

When transformations are tested in fMRI, a fairly clear picture emerges, in which this rule type is neurologically dissociable: The left Broca's region (BA 44, 45) and to a lesser extent, both Heschl's gyri, are most strongly involved in transformational analysis. Ben Shachar *et al*. (2001) have conducted this experiment in Hebrew, searching for a T(ransformational)-effect. They used minimal pairs of sentence that are identical in every respect, except that one set contained a transformation (3a) and another did not (3b):
(3)   a. I helped the nurse [that John saw __ in the living room]
         b. I told   John   [that the nurse slept in the living room]
A T-effect was found in left Broca's area (BA 44/45): A higher BOLD signal was detected for +transformation sentences relative to –transformation sentences. These results suggest a critical role for Broca's region in the analysis of transformations in the healthy brain, and converge on the available lesion data. Moreover, an ROI approach detected activations in the posterior inferior frontal gyrus and the anterior insula, the posterior superior temporal sulcus and Heschl's complex. Thus the core computational resource for Movement structures is in areas 44,45. Auxiliary computations occur at temporal areas bilaterally.

5. DISCUSSION
We have reviewed results that point to the neurological distinctness and locus of the transformational component of syntax. They also suggest that at least some of the results obtained in the fMRI and PET syntactic complexity experiments could be recast in transformational terms, which may lead to a radical reduction in the amount of variation, and to convergence of cross-linguistic and cross-methodological data from lesions studies, as well as from PET and fMRI.

So what is the image of the linguistic brain? We are just beginning to reconstruct it. Whether the somewhat variable anatomy will ever permit precise localization is still an open question; and gross localization, after all, is just a small step toward understanding. Yet our best bet, it seems, is to take linguistic rules as the basic unit of functional analysis of the intricate relationship between language and the brain.

REFERENCES
*Amunts, K., A. Schleicher, U. Bürgel, H. Mohlberg, H.B.M. Uylings & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology, 412*, 319-341.
Ben-Shachar M., Hendler T., Kahn I., Ben-Bashat D., and Grodzinsky Y. (2001). *Grammatical transformations activate Broca's region – an fMRI study*. Presented at the Cognitive Neuroscience Society, New York.
Blumstein, S.E. 1994. The neurobiology of the sound structure of language. In M. Gazzaniga, ed., *Handbook of Cognitive Neuroscience*. Cambridge, MA: MIT Press.

Dapretto, M. and Bookheimer, S.Y. (1999). Form and Content: Dissociating Syntax and Semantics in Sentence Comprehension. *Neuron, 24(2)*:427-32.

Friederici, A. (2000) The neural dynamics of language comprehension. In Alec Marantz, Yasushi Miyashita, and Wayne O'Neil, eds., *Image, Language, Brain*. MIT Press.

Friedmann, N. (1998). *Functional Categories in Agrammatism*. Doctoral dissertation, Tel Aviv University.

Friedmann, N. & Y. Grodzinsky, 2000. Neurolinguistic evidence for split inflection. In M. A. Friedemann & L. Rizzi, eds., *The acquisition of syntax*. London: Blackwell.

*Geschwind, N. (1979). Specializations of the human brain. *Scientific American*, September.

Goodglass H. & J. Berko, (1960). Agrammatism and inflectional morphology in English. *Journal of Speech and Hearing Research,* 257-267.

*Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *Behavioral and Brain Sciences, 23*, 1-71.

Just, M.A., Carpetner, P.A., Keller, T.A, Eddy, W.F. & Thulborn, K.R. (1996). Brain activation modulated by sentence comprehension. *Science, 274*, 114-116.

Mazoyer, B.M., S. Dehaene, N. Tzourio, V. Frak, N. Murayama, L. Cohen, O. Levrier, G. Salamon, A. Syrota and J. Mehler. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5*, 467-497.

*Ojemann, G. A. (1991) Cortical organization of language. *The Journal of Neuroscience*, *11*, 2281–2287.

Petersen, S.E., Fox, P.T., Snyder, A.Z., and Raichle, M.E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, *249*, 1041-1044.

Phillips, Colin , Tom Pellathy, Alec Marantz,Elron Yellin, Ken Wexler, Martha McGinnis, David Poeppel, Tim Roberts. (2001). Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study**.** *Journal of Cognitive Neuroscience*, in press.

Posner, M.I., and Pavese, A. (1998). Anatomy of word and sentence meaning. *Proceedings of the National Academy of Sciences*, *95*, 899-905.

Stromswold, K. D. Caplan, Alpert, N. and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain & Language, 52,* 452-473.

*Zurif, E.B. (1980). Language mechanisms: a neuropsychological perspective. *American Scientist*, May.

*Zurif, E.B. (1995). Brain regions of relevance to syntactic processing. in L. Gleitman and M. Liberman, eds., *An invitation to Cognitive Science*, Vol. I,. Cambridge, MA: MIT Press [2nd edition].

Zurif, E. B. & A. Caramazza. (1976). Linguistic structures in aphasia: Studies in syntax and semantics. In H. Whitaker and H. H. Whitaker, eds., S*tudies in neurolinguistics,* Vol. 2. New York: Academic Press.