Toward an evolutionary linguistic theory

MSc Thesis (*Afstudeerscriptie*)

written by

David Fiske
(born January 30, 1982 in Bryn Mawr, Pennsylvania, U.S.A.)

under the supervision of **Prof. Dr. Martin Stokhof**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
June 21, 2010	Prof. Dr. Jeroen Groenendijk (chair)
	Prof. Dr. Martin Stokhof
	Dr. Maria Aloni

Abstract

This work concentrates broadly on linguistic theory and the methodology thereof, criticizes the role intuition plays, and suggests a gradual replacement of intuition-based data with neurological data to remedy the problem. It argues that linguists can exploit the evolutionary design of the brain to solve interface issues and that, as technology becomes more sophisticated, neurological data will yield more fine-grained data that will inform more detailed theories of syntax, semantics, etc. This hypothesis is demonstrated through a detailed review of neurological data which attempts to categorize some phenomenon (e.g. the Gricean maxims, indirect speech, etc.) as either semantic or pragmatic in nature. Specifically, it is argued that Theory of Mind, a cognitive ability distinct from the Language Faculty, and the right hemisphere of the brain are essential to the felicitous use of pragmatic speech. Following from this, it is argued that a two-channel model of communication, where speakers communicate meaning-altering information not just linguistically, but also through Theory of Mind processes, is better equipped to characterize the relevant data than current one-channel models; the two channel model proposed is a minor alteration of Inquisitive Pragmatics (Groenendijk and Roelofsen 2009). Some other implications which are discussed include the assumption of nativism (which is defended with a poverty of stimulus argument in the style of (Pullum and Scholz 2002)), and the universality of the Gricean Maxims (which is defended through an appeal to cheater detection (Cosmides 1985) and the logic of evolution). This work adheres to the principles of evolutionary psychology in argumentation and draws upon findings from said approach to explain selected empirical observations and implications that arise throughout the argument. Thus, this work is best characterized as a philosophical work of evolutionary psychology and hopes to be seen as a first, modest step toward integrating evolutionary theory with linguistic theory.

Contents

Living in Amsterdam for two years has been a very interesting experience and I owe a tremendous amount of thanks to the following people for helping me to maintain what little sanity I have left: Hanne Berg (and her strange brainwaves), Egil Asprem, Margaux Smets, Christian Geist, Inés Crespo, Hélène Tourigny, Spencer Johnston, Cathy Ankers, Ramunas K. Fishermang, Sara Gothard, Nick Paskert, Kristina Haynie Smith, Roän Smith, Myra Kunas, Ira Brooker, Kate Weber, Jes Burns, Hendrike Peetz, and Thomas Göbel Peetz.

Special thanks go to Martin Stokhof (who once referred to the poverty of stimulus argument as "highly detested, highly detested . . . oh! I mean *con*tested? *Con*tested!, *Con*tested.) who was an invaluable guide through this, my first thesis. In retrospect, I do not think I could have asked for a better advisor. He was generous with his time and looked at everything with an unfailingly sharp eye. The argument presented here would have been much weaker had it not benefited from his insightful criticism.

Thanks also to Jeroen Groenendijk for personal support and for sitting on my defense committee, Maria Aloni for sitting on my defense committee and advice over the last two years, Stefan Frank for personal support, and Michiel van Lambalgen for wonderful editorial assistance concerning a different project. I really appreciate it!

Thanks to Annemie Ploeger for advising me on some of the aspects of evolutionary psychology addressed here and pointing out some references that will help me in my new life as an evolutionary psychologist.

Thanks also belong to Tanja Kassenaar for helping me to beat a path through the jungle of UvA bureaucracy when necessary and making practical arrangements for my defense.

Additional thanks go to Stéphane Airiau and (again) to Thomas, Hanne, and Inés for technical assistance and to my committee members for tolerating a muddled submission process.

I gratefully thank and acknowledge the Evert Willem Beth Foundation for financial support provided while at the ILLC.

Apologies to Jes and Kate, Ryan and Heather, and Thomas and Hendrike whose weddings I had to miss and congratulations to Paula and Patrick, and Ira and Myra who had children (Griselda and Selby, respectively) while I was a student here and also to Roän and Kristina who are due in September!

I would not have ended up here in Amsterdam in the first place if it weren't for discovering linguistics at UMass, Amherst. I learned much from the department as a whole but, in particular, I owe special thanks to Barbara Partee, Lyn Frazier, and Chris Potts for taking a chance on a lowly legal assistant and getting him back into academia where he belongs.

The greatest thanks go to my parents, Robert and Adrienne, who were a constant source of support in every imaginable way. This thesis is dedicated to them.

Lastly, thanks also belong to the friends and strangers not named above with whom I have engaged in a 1000 forgotten conversations that have silently shaped this work into what it is. All errors are theirs.

"If you call someone who knows two languages "bilingual" and someone who knows three languages "trilingual," what do you call someone who knows only one language?

— an American." - John Perry

Chapter 1

Motivating the bio-modeling approach

"In psychology, the study of narrative processes has primarily fallen within the domain of cognitive psychology, although the broad appeal of this topic is reflected in the large number of researchers labouring in a variety of other disciplines. Despite the distributed nature of research in this area, however, the synthesis of ideas from different domains is disappointingly rare." – Raymond A. Mar 2004

"The research reported in this book grew out of the disillusionment that accompanies a graduate student's transition from consumer to producer of ideas." – Steven Pinker 1984

1.1 Introduction

In this, the year following the world-wide party (Shapin 2010) that accompanied the 200th birthday of Charles Darwin and the 150th anniversary of the publication of *On the Origin of Species*, it seems that one cannot help but encounter evolutionary thinking around every corner.

From the major motion picture biopic "Creation" to the Rambert Dance Company's *The Comedy of Change* to the evolution-themed hip-hop stylings of, not just Baba Brinkman, but also Greydon Square ("The Walking Stephen Hawking[1]") the 'Darwin brand' and evolution-conscious thinking have seeped into some unlikely cracks of both the arts and sciences (Shapin 2010; Judson 2010). In the field of engineering, for example, the benefits of 'bio-inspiration' seem so obvious, it is difficult to believe that it took intelligent people so long to think of it (a hallmark of all good ideas).

[1]From the song Rational Response: "First of all, microevolution been proven / In the same way that sound waves let you listen to music / and macroevolution is the system of change / at or above the level of the species inflicted with change / that's microevolution on a grand scale and we can't even say a word / because our life span's too short to observe."

One design challenge in the field of robotics was how to make a robot that can move with ease over granular surfaces like sand. Solving this problem would help to avoid massively expensive tragedies in interplanetary exploration like the one recently suffered with NASA's Spirit, a Mars rover which has become hopelessly trapped in a bed of iron sulphate (III) (Li, Umbanhowar, Komsuoglu, Kodltschek, and Goldman 2009; Editors 2010). Engineers looked to the design of an animal that has no problem walking over granular surfaces, the crab, and reverse-engineered the (evolutionary) design of its legs in order to create SandBot, a robot that moves with ease over granular surfaces and which is likely to be sent to Mars sometime in the future.

Evolutionary-inspired thinking has also led to the discovery of a (patented) self-cleaning dry adhesive (technology derived from the toes of geckos (Autumn, Liang, Hsieh, Zesch, Wai, Kenny, and ad Robert J Full 2000; Autumn, Sitti, Liang, Peattle, Hansen, Sponberg, Kenny, Fearing, Isrealachvili, and Full 2002; Peattie and Full 2007)). Moreover, search and rescue robots capable of walking up the sides of buildings are well under way, based on the same technology. With these examples we see profitable, novel innovation grounded in the principles of reverse-engineering evolutionary design in order to inform deliberate, innovative design.

This work, however, is not about engineering, per se; it is about linguistics. Though the subject matters are vastly different, the challenges both linguists and engineers face are similar in many respects. Both, for example, face the challenge of inventing (or designing, or engineering) a system that will reliably behave in a specified way under certain conditions. For the engineer, it may be designing a robot that can reliably walk across sand or wiring a house for electricity, for the linguist it may be creating a model that can accurately model the semantics of events (Bennett and Partee 2004; Krifka 1989; Landman 2008; Landman and Rothstein 2009; Fiske 2010) or the (ir)regularity of quirky case (Zaenen, Mailing, and Thráinsson 1985).

Here, I would like to suggest that linguists borrow the general strategy adumbrated above: look to evolution to help solve complex design problems, to reverse-engineer evolutionary design in order to inform deliberate theoretical design. But what is the evolutionary design of language?

Darwin famously drew inspiration for his theory of evolution from linguistics; perhaps it is time linguists let 'Darwin' repay the favor and allow some principles of evolutionary theory - and how that has progressed - to inform linguistic theory.

The current state of linguistics is similar to the state biology was in before The Modern Synthesis (Huxley 2010). Just as there was "genetics, developmental physiology, ecology, systematics, paleontology, cytology, [and] mathematical analysis" (Huxley 2010), we now have syntax, morphology, pragmatics, phonology, and semantics. The job of linguists is to describe how language works (just as the job of biologists is to describe how life works) but before attempting such an explanation, one must first answer the question: what is language? (note that, analogously, the question: "What is life?" only becomes tractable with an understanding of mendelian genetics (Dawkins 2006)). Just as the biologists had done, linguists broke their question into the smaller questions of syntax, morphology, semantics, etc. in an attempt to make the question more approachable. And just as in the case of biology (before The Modern Synthesis), the fields 'have been taken up in turn and worked on in comparative isolation' (Huxley 2010).

It was mendelian genetics that synthesized the different fields in biology, but what force will unite the sub-branches of linguistics? From Optimality Theoretic phonology to transformational grammar or Minimalism to Montague Grammar, the diversity and disparity of frameworks makes the question an even harder one to answer. I should note here that many linguists would probably criticize me for asking the question in the first place. There have been attempts to, for example, use Optimality Theory to do not just phonology but also syntax; it could very well be that Optimality Theory will synthesize all of linguistic theory, time will tell. Furthermore, just because no force or theory is in sight does not mean that one will not come in to view in the future.

However, I wish to suggest here is that there is a unifying force in sight and it is known to us, namely, the brain. To some the suggestion will seem so obvious and commonplace as to be unremarkable and to others (to my personal surprise) the suggestion will seem very controversial. I want to elaborate, briefly (I will go into more detail in other sections) on a couple of short-term motivating factors (as opposed to the long term goal of synthesis within modern linguistics) for taking seriously the notion of incorporating neurological data.

The first is that it is a third source of evidence which is often ignored. Despite a wealth of psycholinguistic data, these data almost never have a theoretical impact (as mentioned in Wasow and Arnold 2005 though this fact is widely acknowledged throughout much of the community). The evidence that *is* used comes largely from two different sources: (A) intuitions and (B) corpora. If one is Chomskyan - that is, they consider linguistics to be a branch of psychology - it is rather confusing that intuitions play the role they do since the importance of intuitions has been dismissed by psychologists even before Chomsky. Anderson (2005) writes (regarding the state of psychology at the beginning of the twentieth century), "It was becoming clear [from the work of Wilhelm Wundt and other continental psychologists-*DF*] that introspection did not give one a clear window into the workings of the mind" (Anderson 2005) and goes on to refer to introspectionism as "a naive belief in the power of self-observation" (Anderson 2005). I will elaborate on this point in more detail (in section 2.7) but one reason for considering psycholinguistic evidence (other than the fact that it exists and is potentially useful) is the fact that intuitions are not adequate to the task (a century-old finding which flies in the face of Chomskyan linguists who consider what they do to be of cognitive relevance). With respect to (B) data from corpora is more reliable. As Chomsky writes "the behavior of the speaker, listener, and learner of language constitutes, of course, the actual data for any study of language," (Chomsky 1967) but as he notes earlier in his review of B.F. Skinner's *Verbal Behavior* and elsewhere ((Chomsky 2004b)) the study of corporeal data alone will not get you very far if your goal is to say something about the system(s) that generate language.

The second is that, whether you are interested more in a processing model or some other kind of model of language, these data are still valuable and potentially relevant. I have already hinted at what this is by analogy with the Sandbot designers[2]. What I mean to suggest is that linguists may benefit from (partially) resigning themselves to Orgel's Second rule[3] and entertain the notion of incorporating what direct evidence we have of the natural, neural model of language with the more common theoretically-derived ones. Now, looking to the architecture of the brain does not necessarily entail that the model will be a processing model. For example, in the modest step I am taking here in trying to say something of theoretical importance by looking at neurological data, I will focus largely on a broad issue: the division between semantics and pragmatics, without focusing on any theoretical particulars.

Given the current state of technology, most neurological data would likely only be able help with 'macro questions' [4] in linguistics such as "interface issues", or issues that do not fit squarely into one of the subdivisions listed above (e.g. is the proper treatment of performatives a problem for semanticists or pragmaticians (Boër and Lycan 1980; Hengeveld and Mackenzie 2008)? Is quantifier scope a problem for syntacticians or semanticists (Szabolci 2001)? Are clitics a problem for syntacticians or phonologists (Anderson 2005)?) (also see (Ramchand and Reiss 2007)).

Another collateral benefit is that this approach (as I will demonstrate in the body of this work) can also help us with understanding a distinction between the Faculty of Language in the Broad sense and the Faculty of Language in the Narrow sense (discussed more below). That is, aside from deciding on whose shoulders explanations of arbitrary phenomenon belong, we can also begin to make headway on answering questions like: which properties of the language faculty are unique to language (Hauser, Chomsky, and Fitch 2002) and where does language end and other cognitive abilities begin (Chomsky 2005)[5]?

This idea, however, is a bit complex as it ties the physical (brain) to the theoretical (linguistics) so I want to unpack it in a few steps. In what follows, I will, at various points, refer you to sections in the second chapter. The present discussion assumes a degree of familiarity with the relevant concepts but if you are not sufficiently familiar with a given field, topic, or idea then you may read the section to which I refer you as an extended, detailed footnote. The sections are meant to add depth to the somewhat shallow, but complete, outline presented in the following few pages. Therefore, you can read them as footnotes, if you so like, or you can read them independently as the second chapter of this work which simply elaborates (with greater depth, more evidence, and minimal redundancy) on topics sketched below.

---

[2] Or even earlier when Christopher Wren studied the anatomy of a horse's eye in his attempts at building a better telescope.

[3] "Evolution is cleverer than you are."

[4] Since details that are more specific than localization of function or modularity are more or less beyond our grasp given the current state of technology. But see section 5.6 for a glimpse into the future.

[5] Much of this thesis is focused on this latter question.

Before continuing, let me forecast my thesis. Unfamiliar terms, the concepts they represent, and the relations I propose they stand in with respect to one another will be discussed below. My thesis, the problem I want to make some headway to demonstrate that this 'bio-modeling' method is viable, is this. I propose that Theory of Mind is part of the Faculty of Language in the Broad Sense and can be considered, in part, a pragmatic module and, additionally, that the right hemisphere plays a crucial role in the processing of pragmatic language.

1.2 Natural models

Above, I mentioned the benefits of reverse-engineering the brain in order to inform a deliberate (forward, theoretical) design. Actually, the notion of reverse-engineering is not new to the discussion of the mind/brain. For example, Pinker (1997) writes "The complex structure of the mind . . . [is an] idea [that] can be captured in a sentence: the mind is a system of organs of computation, designed by natural selection to solve the kinds of problems our ancestors faced in their foraging way of life, in particular, understanding and outmaneuvering objects, animals, plants, and other people." He continues, "On this view, psychology is engineering in reverse"(Pinker 1997). The view that Pinker refers to is evolutionary psychology and it is the investigatory framework adopted here because it is the most appropriate available approach (for hopefully obvious reasons; see section 2.2). But to what degree can the structure of the brain inform the structure of linguistic theory?

Given the high degree of structure already present in the brain, there already exists a linguistic theory, of sorts, in the brain - the task is to try to get at it. Luckily, we can exploit certain structural features and characteristics of the brain to get started. The brain is structured in such a way that it gives rise to the property *localization of function* (Gray 2007). Localization of function refers to the fact that certain parts of the brain are dedicated to specific tasks, including (different parts of) language, for example syntax and semantics[6]. Evidence for this property of the brain comes from two different sources: people with brain damage and (e.g) fMRI data. Brain damage to a specific area causes a loss of a certain capability which implies causation; certain parts of the brain being engaged when a participant is presented with materials carefully designed to engage a certain process implies correlation. The convergence of both kinds of evidence is taken to be adequate to make one's case that a certain part of the brain is responsible for a certain function (see section 2.3).

[6]Different aspects of mental experience which have been shown to have neural correlates are vision (Hubel and Wiesel 1977), the ability to detect cheaters (Stone, Cosmides, Tooby, Kroll, and Knight 2002), language (Broca 1965; Wernicke 1874; Wernicke 1885; Wernicke 1906; Geschwind 1970; Sahin, Pinker, Cash, Schomer, and Halgren 2009), the ability to recognize that other people have internal mental states ("Theory of Mind") (Baron-Cohen, Ring, Moriarty, Shmitz, Costa, and Ell 1994; Carrington and Bailey 2009), and there are many more.

Though not entirely historically accurate, we can safely think of the localization of function feature giving rise to the concept of *modularity*, or *domain-specificity* (Cosmides and Tooby 1994). Since different parts of the brain are specialized to perform different tasks, we can think of each task that the brain is specialized to perform as being (potentially semi-) independent of other cognitive capacities. So, for example, after incurring brain damage to a specific part of the brain, it is possible to lose the ability to speak, but not be impaired in any other way[7].

Unfortunately, locating which part of the brain is responsible for which function tells us nothing about the function itself[8]. We can identify the module, but how the module functions is a black box. Modeling the black box is the work of linguists (cf. sections 2.7 and 5.6) since the tools with which we observe the brain are too coarse so say anything of fine-grained theoretical importance. However, since different parts of the brain are responsible for different parts of language (e.g. syntax, semantics, etc.), this method can help determine which phenomena in language 'belong' to which discipline. It is not my intention to address the specifics of a theory of, say, semantics or of syntax here (though I will occasionally make references to specific theories). My primary purpose is to merely make some claims about divisions of labor - the data each theory needs to account for and the extent to which one can treat the other like a wastebasket[9] (Bar-Hillel 1971).

Another benefit which follows naturally, as mentioned above, is that we may gain some degree of insight into a theoretical distinction central to discussions regarding the evolution of language - the distinction between the Faculty of Language in the Broad sense ("FLB") and in the Narrow sense ("FLN") (Hauser, Chomsky, and Fitch 2002). The FLN is comprised of those cognitive abilities which are unique to language and the FLB consists of those cognitive functions that are used in language but also other domains. An example of a cognitive capacity that is part of the FLB is vision.

As a brief illustration of this fact, McGurk and MacDonald (1976) famously showed that what you see influences what you hear. In an experiment, they filmed a woman saying the syllable [ba] over and over again but replaced the audio with a recording of [ga]; "adults," they write, "reported hearing [da]" (McGurk and MacDonald 1976). But clearly vision is used for non-linguistic tasks like navigating a forest; this property of being used in linguistic and non-linguistic tasks alike is what makes it part of the FLB. (Furthermore, vision is obviously part of the FLB in signers who sign their language since vision is the modality in which their language is implemented).

In order to demonstrate the viability of this 'bio-modeling' approach, I will address a current debate: where to draw the theoretical line between semantics and pragmatics (see sections 1.3 and 2.4). I hope to demonstrate that some

[7]In fact, this kind of phenomenon has been documented, it is called Broca's aphasia.

[8]Thanks to Annemie Ploeger for illuminating this point for me.

[9]Treating a field as a 'wastebasket' is when one faces a problem that they are not interested in or is too complex for them to deal with so they say it is not a problem for their field but a different one.

progress can be made in resolving some of these issues by grounding argumentation in the functional, evolutionary design of the brain.

It is worth, briefly, making explicit the parallel between this approach and the approach that the SandBot engineers (and others) have taken. Here, I am trying to reverse-engineer an aspect of evolutionary design in order to (forward-) engineer a better (theoretical) model. We need a principled way in which to reverse-engineer the brain; evolutionary psychology is the best available method. (The forward-engineering, will be where hard work and cleverness of linguists come in, cf. last paragraph of 2.6 and section 5.6.). Furthermore, evolutionary psychology will become important again towards the end in order to cash out some implications that arise during the course of this investigation (Chapter 5).

1.3 Conclusion

Up until now I have primarily been discussing methodological considerations and have briefly mentioned that I would test this method by attempting to shed some light on the distinction between semantics and pragmatics without really saying what these concepts are. Though discussed in a bit more detail in section 2.4, we can presently think of the semantic meaning of a sentence, phrase, utterance, etc. as the meaning that comes from the words alone whereas the pragmatic meaning is what comes from its use in a particular context.

As stated above, the thesis defended here (as a demonstration of the viability of this 'biomodelling' method) is that Theory of Mind is part of the FLB and can be considered, in part, a pragmatic module and, additionally, that the right hemisphere plays a crucial role in the processing of pragmatic language.

There is a hemispheric division of labor that is in-keeping with the notion of modularization. Specifically, that the right hemisphere processes novel behavior and that the left hemisphere processes structured, patterned behavior (section 2.5). However, the other notion relevant to my thesis, Theory of Mind ("ToM"), does not follow neatly from a discussion of design, modularity and hemispheric specialization, so I would like to explicitly introduce it now.

ToM is the ability to attribute mental states to others and yourself as well as make predictions regarding one's behavior based on those attributions (section 2.6).

The claim that ToM is part of the FLB has been made before (Hauser, Chomsky, and Fitch 2002) but it has not been made in reference to human data (Hauser, Chomsky, and Fitch cite primate data). That being said, similar claims have been made using different terminology (Ferstl, Neumann, Bodger, and von Cramon 2007) and there are several studies showing that ToM is responsible for certain linguistic tasks, primarily narrative comprehension (Mar 2004; Mason and Just 2009; Fletcher, Happé, Firth, Baker, Dolan, Frackowiak, and Frith 1995; Martin and McDonald 2003) but also word learning (Bloom 2002). Sabbagh (1999) has argued that ToM is necessary for proper felicitous (in Gricean terms) conversation. Sperber and Wilson (2003) have argued that a submodule of ToM has become a pragmatic module (though they offer no evidence, they simply reason that it must be so). My goal here is to collect and examine this evidence in order to see how it converges with each other and how it compares to a different source of pragmatic language - the right hemisphere.

The claim that the right hemisphere ("RH") plays *some* role in language processing is not new either (e.g. (Beeman and Chiarello 1998)). Furthermore, the claim that the right hemisphere is responsible for some aspects of pragmatic language is not new (Taylor and Taylor 1990). They point out that damage to the right hemisphere results in an inability to determine what parts of a "story are plausible in context . . . tend to accept literal meanings of metaphoric statements without finding them funny . . . [and] may inject themselves into the plot [of a fictional story] or argue with the story's premises."(Taylor and Taylor 1990)

What I hope to do here is expand on the claim that the right hemisphere is responsible for the pragmatic use of language and to explain, in the light of evolution, why it is unsurprising that the right hemisphere plays the role that it does. Also, I want to compare the ways in which ToM is responsible for pragmatic processes like story comprehension and the ways in which the right hemisphere is responsible for pragmatic language like understanding metaphor and aspects of discourse comprehension. (As far as I am aware this comparison has never been made.)

In section 2, I will, as promised, elaborate on the discussion above for the primary purpose of adding some depth to the idea just sketched. Chapters 3 and 4 review the evidence that ToM (Chapter 3) and the right hemisphere (chapter 4) play crucial roles in the processing of pragmatic language. Chapter 5 addresses some implications of this approach in general, which includes some lightly detailed speculation about a theory of syntax which (with work not done here) may be derived using this approach as well as a poverty of the stimulus argument. Chapter 6 concludes the argument. A word is warranted here.

Though it is not directly relevant, this work assumes a nativist position on language and supports this assumption with a poverty of the stimulus argument. If you are a nativist, then please read on. If you are an empiricist who is bothered by the (I should say, weak) assumption of nativism, you may wish to peek ahead at section 5.2 for a defense of this position. If you are an agnostic, innocent, civilian, or someone who simply has no idea what I am talking about then please keep reading as it is not likely that section 5.2, nor the debate at large, will make much sense without the background provided between this sentence and the beginning of that section.

Chapter 2

Elaborations

“A long tradition of scholars have looked for unique mental abilities in humans. Three allegedly unique abilities are grammatical language; theory of mind, or social cognition the capacity to understand another’s mental state and to take it into account; and certain forms of intelligence, such as intuition.” – Bryan Kolb and Ian Q. Whishaw 2009

2.1 Introduction

This work is, in a sense, two things at once. First, as already discussed, it is a work of linguistics. Second, however, it is a kind of philosophical work of evolutionary psychology. It is the first step toward a top-down effort to integrate evolutionary and theoretical linguistic thinking. The benefits of this are two-fold. The first, as already discussed, is that it may allow linguists to use evolutionary design as a cheat-sheet when attempting to solve problems or design tasks. Second, it brings linguistics closer to its original (and I would argue preferable) conceptual foundation as a theory that attempts to create an accurate, formal model of what is happening in the brain.

Why the brain? I adopt the perspective of “good old-fashioned cognitive science” in that my aim is “to discover how actual systems [of the brain] work” and I am only concerned with simulations that contribute to this goal (Chomsky 1997).

This view is nicely summed up by Chomsky (1997[1]): “The legitimacy of the study of the states of the brain implicated in language use can hardly be denied. We can ask whether particular theories of these states and their functioning are true or false, but the injunction that some strange category of researchers called “linguists” must keep to a different enterprise is not easy to comprehend”

[1]Though the cite is from 1997, earlier in the article he writes that the point of view he expresses in the 1997 paper is “from a point of view close to that of my paper on language at the 1956 Cambridge conference.” (Chomsky 1997) Therefore, it is an accurate reflection of the beginnings of cognitive science.

(Chomsky 1997). This is what he calls the "naturalistic approach of mental linguistics" (Chomsky 1997).

The point of view that Chomsky is arguing against is what he calls the "non-naturalist" position which holds that "there is some object, a language, independent of the speakers and their brains; the speaker may have some confused picture of this object, but a qualified linguist ought to be able to find a better one" (Chomsky 1997). This is juxtaposed to the naturalist position which "undertakes a study of a real object in the natural world, the brain and its function"(Chomsky 1997).

While Chomsky argues against the non-natural position, I contend that an argument against either is misguided and that the two perspectives can peacefully coexist. There is confusion which arises, I think, out of a conflation of vocabulary. It will be good to briefly, disentangle some notions before continuing any further.

It is possible to characterize Language in many different ways because the phenomenon has been discussed for over a millennium (see (Allan 2004)). There are natural languages such as Dutch, Romanian, English, Lenape, or Mokoko. There are artificial, man-made languages, such as Logic, C++, Klingon, Elvish, and Esperanto (though these last three are, clearly, much less useful than the first two). Here I am concerned only with natural languages.

Aside from different types of languages (discussed above) there are different levels of language. I will, following Carnie (2002), borrow a notational strategy to differentiate between the levels of discussion presented here. He writes, "Language (written with a capital L) is the part of the mind or brain that allows you to speak, whereas language (with a lower case l) is an instantiation of this ability (like French or English)." (Carnie 2002)

This brings me to the conflation of ideas that, I think, is evident in the Chomsky quote above. The naturalist position holds that a person knows their *Language* (e.g. *grammar and vocabulary*) in the way that Chomsky describes (i.e. an individual's brain states), but they do not know their *language* in the way he describes. What's more, this is not a controversial claim. For example, most[2] people do not know every word of their language. So, there is a sense in which one's language exists abstractly in the community at large (in the set of brains of speakers of the same language). The fact of (subtle) dialectical constructions that exist in one but not another grammar also complicates the picture by complicating the notion of what constitutes a language in the first place. (The Language/grammar Chomsky refers to is an idiolect. If we read Chomsky literally, we may never reach a theory of Universal Grammar. The idea of a language is an abstract community-dependent notion to begin with).

So, language is a social construct that (abstractly) exists in and belongs to the community, but the flip side of that coin is that the speakers of the community must have knowledge of the language of that community (their Language, their individual and unique grammar and vocabulary). In order for a language

[2] I qualify this because there may be so-called autistic savants, lexicophiles, hobbyists, etc. who do know every word of their language. However, the average person does not.

to exist in a community, grammars must also exist in the brains of each individual member of the community (and each single person's mental representation of their Language is only a subset of that language).

Undoubtedly there are other ways of looking at l/Language but here I approach it as a physical ability grounded in the brain. This is not to say that other ways of studying l/Language are not valid, many of them they are, but my purpose here is to stand in advocacy of the idea that the organization of Language in the brain (I will explain what I mean by this below) can inform linguistic theory with respect to making certain theoretical decisions. Under the view pursued in this work, Language comes from the brain.

Below, I will begin with a brief description of the principles we will take when making the first step towards reverse-engineering the brain and then go on to elaborate on some of the points I made above with an eye towards explaining how we can exploit certain evolutionary design features to help us in theory-construction and end with an argument as to why this method is preferable over the largely intuition-oriented method that dominates much of linguistics today (Chomsky 2004a).

2.2 Evolutionary psychology

Of course, as already stated, we need a principled methodology for reverse-engineering the brain. Rather than re-inventing the wheel, we can borrow a approach that is already on the books: evolutionary psychology. While not itself a work of evolutionary psychology, this work adheres to the principles and tenets and takes on board the assumptions of evolutionary psychology. So a review of evolutionary psychology is needed.

Evolutionary psychology is an approach to psychology that looks at psychological mechanisms that have evolved to carry out specific tasks that were adaptive to the environment in which they evolved; for modern humans most of that evolutionary time was spent as hunter-gatherers in the Pleistocene roughly up until the time agriculture was invented 10,000 years ago (Cosmides and Tooby 1992). Evolutionary psychologists regard the brain as a natural computer which has evolved to efficiently solve the problems that faced our hunter-gather ancestors and task themselves with reverse-engineering that computer, the brain.

Evolutionary psychologists look at "three complementary levels of explanation" (Cosmides and Tooby nd) in order to explain behavior in evolutionary terms.

"Theories of adaptive problems," they write, "can guide the search for the cognitive programs that solve them; knowing what cognitive programs exist can, in turn, guide the search for their neural basis" (Ibid.). It should be noted that at this point, most of the work is on the top two levels of explanation. This is for the simple fact that it is difficult to reliably establish the neurophysiological basis of cognitive programs. Still, I will, whenever possible, discuss whatever neurophysiological data is available in order to suggest a picture which can, in principle, be experimentally falsified (though carrying out the experiments is

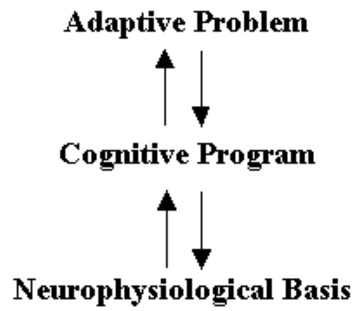

Figure 2.1: Three complementary levels of explanation in evolutionary psychology. Inferences (represented by the arrows) can be made from one level to another.

(Cosmides and Tooby nd)

beyond the scope of the present work as it would involve access to an fMRI lab or certain clinical populations to which it is not easy to gain access as a logician).

While evolutionary psychology is the broad framework for the present discussion much of the argument presented here revolves around psychological experiments and relies on what we know about the structure of the brain; therefore a review of the methods of investigation are in order. In the following section, I will review some basic investigatory techniques, criteria that data must meet in order to draw a reasonable conclusion, as well as discussion some basics of brain structure.

2.3 Investigatory principles

Above I made the (not uncommon) comparison of the brain to a computer. We can think of the circuits of that computer, the neural substrates, as being dedicated to one particular cognitive process. The sum (or set) of all neurons dedicated to a particular function is called a *module* and this general property of the brain is referred to as *localization of function* (or alternatively *modularity*)

I have already mentioned that evidence for localization of function comes from brain imaging studies and studies of brain-damaged people. There are different types of brain imaging techniques such as (functional) Magnetic Resonance Imaging ("(f)MRI") technology and variants (e.g. Positron Emission Tomography "PET"), Magnetoencephalography ("MEG"), which measure the blood flow to a certain part of the brain. After neurons have fired in the brain (or have transmitted an electrical charge), blood reliably flows to the area of recent electrical activity (a phenomenon known as *hemodynamic lag*). The above named technologies measure hemodynamic lag. Before this technology existed the best way to measure brain activity was Electroencephalography ("EEG"), which measures electrical activity at the scalp and is still used today.

There is also a rare and highly accurate (though highly invasive) method, intra-cranial Electrophysiology ("ICE"), where "electrodes are [placed] directly in contact with the brain. These electrodes are left in place for some days while the patients spend time in the hospital (the insertion sites are all covered and bandaged)[3]." The electrodes measure at the level of the neuron and offer the highest resolution of any currently available method.

In the case of people who have suffered some kind of trauma and incurred a lesion, loss of a certain kind of function - say the ability to recognize faces[4] - is associated with damage to a certain area of the brain. These lesion studies are particularly interesting because they demonstrate the necessity of a brain region to perform a specific task. However, lesion studies are also imprecise because the lesion often impairs several different Brodmann areas[5] and has ill-defined boundaries both in terms of breadth and depth. Furthermore, it is difficult at this stage in brain research to understand what damaging one area does to connected areas in terms of computational or(/vs.) metabolic activity (Rorden and Karnath 2004). This type of study offers certainty but little clarity. It provides evidence for causation but not for correlation. Therefore, a lesion study can only serve as ancillary evidence.

The second class of studies (save EEG), allow researchers to get a clear picture of the brain in terms of localization of function. However, it is difficult to concretely determine which areas are responsible for which aspects of the task at hand. These types of studies can only prove an association. They offer a clear picture of correlation but no certainty of causation.

Converging evidence, however, taken from both methods provides strong reasons to believe that a particular brain region is responsible for a certain function. Imaging studies provide the picture we want (but see below for a discussion of just how clear) and corroborating lesion study evidence gives us the certainty we are looking for in terms of causation. (See (Rorden and Karnath 2004) for a more in-depth review that draws the same conclusion.)

Imaging is the more exciting of the two methods because, for the first time, it gives a picture of how the brain functions. However, one must remember that CT scanning, the first kind of imaging technology with which one could take a useful picture of the brain, was first implemented only on October 1, 1971 and it was not until 1990 that the first fMRI scan was reported (Ogawa, Lee, Nayak, and Glynn 1990; Filler 2009). Since many of the imaging studies we look at will be fMRI and PET studies, and at the time of this writing the first implementation of that technology was only 20 years ago, we must ask: how clear is the picture we are getting with this technology?

[3] Website of Ned T. Sahin, nedsahin.com

[4] Clinically called prosopagnosia. It is generally thought that the fusiform gyrus is responsible for this ability. (Renzi, Perani, Carlesimo, Silvers, and Fazio 1994),(Kanwisher, McDermott, and Chun 1997) See (Wilmer, Germine, Chabris, Williams, Loken, Nakayama, and Duchaine 2010) for evidence that "Human face recognition is specific and highly heritable."

[5] Brodmann 1909 delineated 52 distinct cytoachitectonic areas. These are commonly employed to discuss the topology of the brain.

Little is known about the brain relative to its complexity and neuroimaging technology, as just discussed, is in its infancy. It is also, in a sense, imprecise. As Saxe, Carrey, and Kanwisher (2004) point out "the location of activation in the brain is often specified by general region (e.g. the occipital pole, or the temporo-parietal junction), or by the gyrus or sulcus where the activation is found (e.g. the fusiform gyrus, or the intraparietal sulcus)" but each one of these areas "spans ten or more centimeters of cortex"(Saxe, Carey, and Kanwisher 2004). Furthermore, there are imprecisions at the voxel[6]-level, each voxel representing "hundreds of thousands of neurons"(Saxe, Carey, and Kanwisher 2004) and, of course, there is individual variation; no two brains are the same.

This, however, is not to say that more precise technology exists only in the distant future. The first CT scan was taken in 1971, but less than less than thirty years later, in 1999, the first direct brain-machine interface was demonstrated (Lebedev and Nicolelis 2006). Thus, it is rational to assume that the technology necessary to obtain more fine-grained images of the brain, bottlenecks aside, is soon approaching.

All of that being said, there exists data and if one has access to empirically collected data it is important to take that data and construct a theory out of it, not only because it is fun, but because it helps direct the course of future data collection by creating testable predictions.

2.4 Semantics and pragmatics

The bulk of this thesis is about what linguistic phenomena count as semantic and which count as pragmatic with respect to theoretical models. Before setting out on this journey, it would be a good idea to have some notion of what counts as "semantic phenomena" and what counts as "pragmatic phenomena." However, this is actually a deceptively tricky proposition because what I have proposed is to describe that which I also seek to discover empirically, which is simply not possible. Despite this fact I will do the impossible for the sake of clarity of exposition and the fact that these theoretical notions already exist going into the project and they can serve as useful guidelines as long as one is mindful that these definitions may change as a result of empirical data. Moving on then, let's equate *pragmatics* with the ability to use the words of a language to get your point across even though not every sentence may be perfectly grammatical and *semantics* with syntactically and semantically well-formed sentences of a language. To elaborate on this as much as possible while simultaneously attempting to remain as theory-neutral as possible, I will discuss the terms "semantics" and "pragmatics" by addressing their original (modern) conception as well as giving illustrative examples.

Morris (1938) first writes, "the study of [the relations of signs to the objects to which the signs are applicable] will be called semantics . . . the study of [the relation of signs to interpreters] will be called pragmatics" (Morris 1938). Of course, Morris is using the word "sign" in the Saussurian sense to refer to words.

[6] A voxel, or volumetiric pixel, is a three-dimensional pixel.

This definition obviously is not the whole story, however, as some signs/words (e.g. *the*, *however*, *and*, *before*, etc.) do not refer to any object.

In my attempt to remain theory-neutral (to not put too fine a point on a definition of semantics or pragmatics), I will sketch some examples and point to, what many would deem, the semantic and pragmatic aspects of the situations in order to give you a 'feel' for the difference between these two dimensions of meaning. However, in general, I will take the semantic meaning to be that which is *literally* true and the pragmatic to be that which is *not literally* true, but which is still communicated.

Let's begin with a classic example: "*Can you pass the salt*?" If one were to 'answer semantically' then the answer should be *"yes"* or *"no"*. On the other hand if one were to 'answer pragmatically' then a verbal response may not even be necessary as the passing of the salt would suffice. In principle, any question that begins with *"Can you . . ."* can be properly answered by *"yes"* or *"no"*. The form of "*Can you perform open heart surgery*?" is the same as the salt example which is the same as "*Can you please leave now*?" But the salt and leaving examples are clearly requests for action whereas the heart surgery one is not (let's hope!). So, what does "*Can you pass the salt*?" mean? How do you *know*?

Another example is the non-literal nature of metaphor. For example, when I say that I have 'a lot on my mind,' I am not talking about my hat - I am saying something about my mood and state of mind.

Conflating a literal interpretation with a non-literal interpretation is what licensed the comedian Mitch Hedburg to say "You know, I'm sick of following my dreams, man. I'm just going to ask where they're going and hook up with 'em later."

In the first sentence he speaks non-literally; his dreams are not physical objects in the world and he is not running after them. What he means is he is tired of working hard to achieve a desired outcome. However, in the second sentence, he interprets the non-literal expression literally and then trades on the unexpected literal interpretation to derive a punch line.

Yet another example is sarcasm. There are several different conceptions of sarcasm, but the one I want to focus on is the idea of saying the opposite of what one means. The problem with giving an example of sarcasm on paper is that the difference between a literally true statement and a sarcastic one is usually conveyed through tone of voice.

A good illustration, however, can be seen in the following newspaper exerpt[7]:

> DAYTON, OH—Confusion and awkwardness are still being felt after a boisterous guffaw intended to convey sarcastic dismissal of a comment was incorrectly interpreted as a positive response by the very person who was being mocked Tuesday. "My laughter led the speaker to infer that I was expressing approval and delight at his statement, and he therefore came away believing that I was supporting said statement when I was in fact ridiculing it," Clayton

[7]Taken from the humor 'newspaper' The Onion.

> Obermeyer said of his remarks to coworker Derek Havers. “However, despite my heartfelt desire to puncture the arrogance of the speaker, an unfortunate opposite reaction seems to have taken hold.” Obermeyer said he was really looking forward to attempting to explain the nature of sarcasm to Havers, because he was just so sure that the absolute genius would completely grasp every last word without the slightest little bit of trouble whatsoever. (Editors 2006)

The fact that it was misinterpreted not only goes to show that sarcasm is a dangerous weapon to wield given that it can be misinterpreted but also speaks to the difficulty of writing about it due to the crucial role of prosodic factors. However, the final sentence of the ‘report’ is a very clear demonstration (in light of the preceding conversation) of how the semantic or literal meaning of a sentence can, in fact, be virtually opposite of what it means in a given context. Obviously ‘Mr. Obermeyer’ does not think ‘Mr. Havers’ is an absolute genius who would completely grasp every last word without the slightest little bit of trouble whatsoever.

So, again, the problem facing the linguist is: what does the sentence *mean*? Is the meaning of the sentence a product of the words themselves or of the speaker’s intention in using them? How do you know?

Despite my attempts to motivate the idea that this is an interesting and contested research area, making this distinction may still seem like a trivial problem. In fact, a series of books dedicated to the problem (Current Research in the Semantics/Pragmatics Interface, CRiSPI, which, at the time of writing, has 20 volumes in the series, each about a different topic), a number of individual books ((Szabó 2005; Turner 1999; von Heusinger and Turner 2005)), too many papers to even mention, and an entire journal (*Semantics and Pragmatics*) have emerged. Thus, the controversies surrounding the topic have generated a significant amount of ink.

This particular problem is also an interesting and appropriate for the approach described here for several reasons. First, it deals with linguistics at the macro level. This is appropriate because of the current state of brain-imaging technology (cf. the last section). Brain research on language is not yet at the point where we can see brain function or structure clearly enough to develop, say, a theory of syntax (but see section 5.6 for some tantalizing new research which may re-open a slightly altered version of an old - and perhaps prematurely resolved- debate).

Second, it is a notoriously complex distinction to draw because there is no consensus over how semantics and pragmatics are defined (see (Bach 1999)). In (at least partially) relativizing linguistic theory to the brain, distinctions can more easily be made between various processes - semantic, pragmatic, syntactic - because they are spatio-temporally separated.

From the discussion above, one might wonder whether different kinds of functions have been localized to the two different hemispheres of the brain. Perhaps you have heard the old cliché about there being left-brained people and right-brained people? In fact, there is evidence for this kind of distribution and it is another 'architectural detail' that we can exploit to help us draw theoretical distinctions between semantics and pragmatics.

2.5 Hemispheric diagnostics

If you have heard the old cliché about left and right brained people then you know that left-brained people are supposed to be better at math and science, are well organized, and have their feet firmly planted on the ground while right-brained people, on the other hand, are supposed to be more artistic and creative, are disorganized, and have their heads in the clouds. If you believed it, I have some bad news: there is, in fact, not much to this.

For example, a poet may be considered a 'right-brained' person but they spend much of their time using language; language areas are primarily located on the left side of the brain. Furthermore, many people might consider Albert Einstein to be the archetype of a left-brained person though he famously said "Imagination is more important that knowledge" and gained much of his stupendous mathematical insight through the use of thought experiments (which would typically be thought of as right-brained activity).

Actually, the reality that the cliché is based on turns out to be even more interesting than the cliché itself. The modern understanding of hemispheric differences is that the right hemisphere is responsible for the detection of novelty in the environment whereas the left hemisphere is responsible for the processing of patterned, stereotyped or sequenced behaviors. In fact, this trait does not appear to even be unique to humans.

Based on evidence from behavior and observation from a wide range of vertebrates (from toads to chimpanzees)[8], MacNeilage and colleagues theorize that "the left hemisphere of the brain was originally specialized for the control of well-established patterns of behavior under ordinary and familiar circumstances . . . the right hemisphere, the primary seat of emotional arousal, was first

[8]Here is a sampling of quotes which demonstrate the breadth of the phenomenon. (Remember that the left hemisphere controls the right side and vice-versa): 1. "Fishes, reptiles and toads . . . tend to strike at prey on their right side under the guidance of their right eye and left hemisphere . . . [for] chickens, pigeons, quails and stilts[,] the right eye is the primary guide for various kinds of food pecking and prey capture"
2. "Phillip J. Clapham . . . and his colleagues discovered that 60 out of 75 [humpback] whales had abrasions on the right jaw; the other 15 had abrasions only on the left jaw" which is evidence that whales are predominantly "right-jawed"
3. ". . . common marmosets open the right side of their mouths wider than the left side when making friendly calls to other marmosets. People also generally open the right side of their mouths to a greater extent than the left when they speak - the result of greater activation of the right side of he face by the left hemisphere"
4. "Not only do chimpanzees tend to be right-handed when they manipulate objects, but they also favor the right hand for communicative gestures. Gorillas, too, tend to incorporate the right hand into complex communications that also involve the head and mouth. Adrien Meguerditchian and Jacques Vauclair, both at the University of Provence in France, have even observed a right-handed bias for one manual communication (patting the ground) in baboons."
5. "Toads, chameleon, chicks, and baboons are more likely to attack members of their own species to their left than to their right" (MacNeilage 2009).

specialized for detecting and responding to unexpected behavior in the environment"(MacNeilage 2009).

This lateralization even appears to be the case for some high-order cognition in other animal species. They write, "In birds, for instance, studies have shown that the left hemisphere controls singing. In sea lions, dogs and monkeys, the left hemisphere controls the perception of calls by other members of the same species" (MacNeilage 2009).

In fact, it is interesting to note that the traditional language centers, Broca's area, Wernicke's area, and the arcuate fasciculus, are all on the left side. Undoubtedly, the left hemisphere has proven to be more active in language-oriented tasks and damage to the left side of the brain has more dramatic consequences to language than damage to the right. (See sample dialogue of someone with Broca's aphasia on page 51-52 compared to the other examples of RH damage throughout chapter 4.)

In keeping with this picture, MacNeilage et al. (2009) write, "Michael D. Fox and his colleagues at Washington University in St. Louis conclude that humans posses an "attentional system" in the right hemisphere that is particularly sensitive to unexpected and "behaviourally relevant stimuli," or in other words, the kind of stimuli that say, in effect, Danger ahead! . . . even right handed people respond more quickly with their left hand (right hemisphere) than with their right hand"(MacNeilage 2009).

They conclude that in all species, apparently including humans,

> To detect novelty, the organism must attend to features that mark an experience as unique. Spatial perception calls for virtually that same kind of "nose for novelty" because almost any standpoint an animal adopts results in a new configuration of stimuli. That is the function of the right hemisphere. In contrast, to categorize an experience, the organism must recognize which of its features are recurring, while ignoring or discarding its unique or idiosyncratic ones. The result is selective attention, one of the brain's most important capabilities. That is the function of the left hemisphere. (MacNeilage 2009)

Language comprehension, more times than not, involves a combination of the very familiar and the completely unexpected. What is completely familiar is the syntax of the language and most, if not all, of the words. What is unexpected is how those words will be used, in what combination and how literally they will be used. In fact, linguists often claim that most sentences have never been spoken before. Language-in-use, on this view, is always novel.

So, then what of the implication that the right hemisphere is involved with speech in light of this new theory? (Remember, I mentioned that traditional language centers are on the left). Actually, evidence has been accumulating which attests that the right hemisphere does play some role in language processing.

As Lindell writes, there is "a growing body of research demonstrating that, far from being nonverbal, the right hemisphere has significant language processing strength"(Lindell 2006). She goes on to write, "the RH lacks the capacity to generate productive language in over 95 percent of the population (Loring et al.1990), thus the research discussed in the review focuses on the language capabilities of this "typical" RH" (Lindell 2006). This brings up a minor, but technically required point. Humans, and our brains, are biological entities and therefore are subject to individual variation. This thesis is concerned with the 95+ percent - the average brain.

The picture that is emerging in the literature is this: damage to the right side of the brain should affect pragmatic functions of language and not, say, syntax or semantics and damage to the left side of the brain should affect syntax and semantics and not pragmatics[9]. Of course, this is a weak hypothesis since it has been known since Broca (1861) that damage to the left side of the brain impairs syntax but as we will see, damage to the right side of the brain seems to affect language at the pragmatic, and not the semantic (or syntactic), level.

We have just seen that the notion of hemispheric diagnostic and the association between the left hemisphere and structure-dependent elements of language (like semantics) and the right-hemisphere and novelty-dependent elements of language (like pragmatics) follows directly from what we know about the architecture of the brain. Another important notion, which does not follow directly from the architecture of the brain (but obeys a different sort of logic) is the role that Theory of Mind plays. Here I will expound a bit more on ToM: its nature, it's similarity to Language, and to evidence that strongly suggests ToM (like Language) is an evolved (innate) capacity.

2.6 Theory of Mind

Abu-Akel writes that ToM is "[t]he capacity to represent one's own or another's mental states such as intentions, beliefs, wants, desires, and knowledge." (Abu-Akel 2003) The term "Theory of Mind" comes from Premack and Woodruff (1978) who ask if the chimpanzee has this ability (if not it seems likely that this is a species-defining characteristic). They write:

> In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others (either conspecifics or other species as well). A system of inferences of this kind is properly viewed as a theory, first, because such states are not directly observable, and second, because the system can be used to make predictions, specifically about the behavior of other organisms. (Premack and Woodruff 1978)

Baron-Cohen et al. (1985) began a fire storm of research when he asked the same question not of the chimpanzee but of the autistic child. Since then a mountain of evidence (e.g. Baron-Cohen 1985-2010) has accumulated in support of the idea that some kind of ToM deficit is responsible for some of the behavior

[9] So, in MacNeilagian terms semantics is the study of recurring (predictable) features whereas pragmatics is the study of recurring (predictable) features+novelty.

of people with autism. It has also been shown that this ability, like language, has a predictable neural basis ((Baron-Cohen, Ring, Moriarty, Shmitz, Costa, and Ell 1994; Abu-Akel 2003; Carrington and Bailey 2009)) in both normal ("NT") people and autistic ("ASD") people. I have postulated that ToM is partially a pragmatic module[10]. Here, I want to present some preliminary evidence that the ToM and the Language faculty are related modules.

Actually, the similarities between the two faculties run deeper than just having predictable neural bases. Stone et al. (1998) note six reasons to consider the case that ToM is modular (Leslie 1987; Leslie 1994; Leslie and Roth 1993) in the same way that language is commonly thought to be. They enumerate the following,[11]

1. Theory of mind can be selectively impaired in the developmental disorder of autism, while other aspects of cognition are relatively spared (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, 1989b; Baron-Cohen, 1995).

2. Theory of mind can be selectively spared while other cognitive functions are impaired, as in Down's syndrome and Williams syndrome (Karmiloff-Smith, Klima, Bellugi, Grant, & Baron-Cohen, 1995).

3. Use of theory of mind is also rapid,

4. *and* automatic, requiring no effortful attention (Heider & Simmel, 1944),

5. Theory of Mind is universal, as far as is known (Avis and Harris 1991).

6. Finally, theory of mind has a particular stereotyped developmental sequence. (Stone, Baron-Cohen, and Knight 1998)

Regarding (1) and (2), double dissociation is widely thought to be evidence for, not only modularity, but for a genetic dissociation (Pinker 1999) (though this view is not entirely uncontroversial; see (fen Hsu and Karmiloff-Smith 2008)).

Fletcher et al. (1995) also note that the ability to mentalize, like language, is independent of intelligence ((Fletcher, Happé, Firth, Baker, Dolan, Frackowiak, and Frith 1995) and see (Yamada 1990))

As a brief aside, it is quite interesting that language disorders should so consistently (though not totally) be seen in ASD people, especially given the fact that autism "is heterogeneous both in its phenotypic expression and its etiology." (Tager-Flusberg and Joseph 2003) and that the FOXP2 gene (a gene associated with language) "lies in one of the most consistently linked autism chromosomal regions of interest" (Wassink, Piven, Vieland, Pietila, Goedken, Folstein, and Sheffield 2002). Of further interest is the fact that, as far as we know, this is coincidental. Wassink et al. (2002) (Wassink, Piven, Vieland, Pietila, Goedken, Folstein, and Sheffield 2002), Newbury et al. (2002) (Newbury, Bonora, Lamb, Fischer, Lai, Baird, Jannoun, Slonims, Stott, Merricks, Bolton,

[10]See (Baron-Cohen 1988) for the first similar claim.

[11]Technically they only enumerate (1)-(5).

Bailey, Monaco, and the International Molecular Genertic Consortium 2002) and Meaburn et al. (2002) (Meaburn, Dale, Craig, and Plomin 2002) all show, there is no link between the language disorders associated with autism and the FOXP2 gene. This fact is of mild interest since FOXP2 is the only currently known language-related gene and pragmatic abnormalities seem unassociated with it. Though one cannot draw any conclusions from such a small kernel of data, it is of mild interest that pragmatic abnormalities seem unassociated with the only known language gene and associated with a region of interest in which abnormalities cause pragmatic (and ToM) deficits. End aside.

I will discuss (3) and (4) below as they are relatively self-explanatory but become more interesting when considering the larger picture.

Regarding (5), the universality of a phenomenon has "only a few general explanations" (Brown 1999). One is that the phenomenon is both useful and culturally ancient so it would have had time to spread around the globe; cooking food with fire is an example of this. The other is that a phenomenon has a (partly) genetic basis and the phenomenon is universal because it is a species-specific trait. Given the identification of genetic factors involved in both Language and ToM[12], the evidence so far suggests that the latter species-specific explanation is the appropriate one for Language and ToM. (Also see (Dediu and Ladd 2007) for some provocative research towards the same end.)

Further evidence of a genetic basis is seen in (6). Baron-Cohen (1997) enumerates some examples of behavior that requires ToM by comparing the development of NT children and ASD children. For example, NT children "[a]t 14 months old [will] actively monitor where someone else is looking, by turning to look in the same direction as another person when that person turns to look at something" (known as joint attention) (Baron-Cohen 1997).

Furthermore, NT children will look to what someone is pointing at rather than, for example, showing no interest or looking at the pointing hand itself. NT children, at 14 months, will also retrieve objects being pointed to. All of these behaviors are examples of joint-attention and display evidence of a "meeting of minds" (Baron-Cohen 1997). At 18 months (and beyond into preschool) ASD children do not exhibit evidence of engaging is joint-attention with others. Furthermore, ASD children do not monitor the attitude of peers by evaluating facial expressions (as NT children do) and do not engage in pretend play (as NT children do).

Beyond this is the well-established fact ((Doherty 2008; Baron-Cohen, Tager-Flusberg, and Cohen 2000)) that between three and four years of age children gain the capacity to understand that other people can have false beliefs (see fn. 13, p.28). This is another developmental regularity.

[12]Though no 'autism gene' has been found, researchers have begun zeroing in on genetic factors - specifically gene (SLC6AH)- gene (ITGB3) interaction (Ma, Rabionet, Konidari, Jawoski, Wright, Abrason, Gilbert, Cuccaro, Pericak-Vance, and Martin 2010) and genes that may increase the susceptibility of the individual to become autistic , namely D6S283 (Philippe, Martinez, Guilloud-Bataille, Gilberg, Røastam, Sponheim, Coleman, Zapella, Aschauer, van Maldergem, Penet, Feingold, Brice, Leboyer, and the Paris Autism Research International Sibpair Study 1999).

Language, too, has a specified pattern of development shown in the following table:

age	Accomplishment
4 months	Babbling
12-15 months	One word
14-20 months	Two words
24 months	Multi-word telegraphic
2.5-3 years	Inflections
3 years	Mostly complete grammar
4-7 years	Irregular forms

The universality of differences in joint attention, predictable ToM and Language development taken together strongly suggest that the ability to acquire language and ToM have a specific genetic basis.

Evolutionary psychologists note that "construct validity is never fully established with a single set of observations or correlations. Instead, researchers build a nomological network of evidence that is continually constructed and pruned as new evidence comes along" (Schmitt and Pilcher 2004). They list the following eight Modes of evidence with example types of evidence:

1. Theoretical "Evidence"
   (a) Evolutionary Biology Theories*
   (b) Theoretical Selection Pressures*
   (c) Cost-benefit Analyses
   (d) Game-theory Simulations
   (e) Computer Modeling and AI
2. Psychological Evidence
   (a) Behavioral Tests and Surveys*
   (b) Cognitive reasoning specificity
   (c) Emotional System Specificity*
   (d) Developmental Specificity*
   (e) Sub-Discipline of Psychology*
3. Medical Evidence
   (a) Fertility and fecundity Studies
   (b) Physical Health and Mortality
   (c) Mental Health and Happiness*
   (d) Psychiatric Disorders*
   (e) Nutrition/Exercise Applications

4. Physiological Evidence
   (a) Morphology and Fertility
   (b) Neuroanatomical Structures*
   (c) Neurotransmitters & Hormonal Substrates
   (d) Brain and behavior Research
5. Genetic Evidence
   (a) Behavioural or Population genetics*
   (b) Molecular genetics
   (c) Experimental gene mapping
   (d) Manipulation Studies & Gene Replacements
6. Phylogentic Evidence
   (a) Animal Ethology
   (b) Comparative Psychology*
   (c) Primatology
   (d) Physical Anthropology
   (e) Paleontology
7. Hunter-gatherer evidence
   (a) Cultural Anthropology*
   (b) Ethnographic Studies
   (c) Human Ethology
   (d) Human Behavioral ecology
   (e) Human Sociobiology*
8. Cross-Cultural Evidence
   (a) Ethnological Comparisons
   (b) Human Universals*
   (c) Ecology-dependent Variability*
   (d) Facultative Conditional Adaptations

Note that Stone et al.'s (1998) (1)-(6) touch on at least one element of all the modes (indicated by a *).

(1) and (2) (together representing the property of double dissociation) satisfy 1a and 1b because ToM and Language are beneficial adaptations (Pinker and Bloom 1990), 3c and 3d because a ToM deficit (autism) represents an instance of mental health disorder (to varying degrees, as well as potentially harming the fitness of the individual), and 4b since there is neuroanatomical evidence for the

localization of both Language and ToM. One's native language is determined by the environment, which satisfies 8c.

(3) and (4) speak towards 6b. Most researchers agree that ToM (of even *near* human magnitude) and Language are species-specific traits. Most researchers agree that Language is species-specific and that some primates, like chimpanzees, may have a rudimentary ToM.

(5) touches on 5a (as briefly discussed already), 8b because Language and ToM are human universals as well as 7a and 7e, generally.

(6) speaks towards 2d, developmental specificity touched on by the stereotyped development of both Language and ToM, and (hopefully obvious by now) 2e. Furthermore, false-belief tasks[13] are instances of 2a and and low empathy in ASD people (Lawson, Baron-Cohen, and Wheelwright 2004; Golan, Baron-Cohen, Wheelwright, and Hill 2006; Chakabarti, Bullmore, and Baron-Cohen 2006; Lombardo, Barnes, Wheelwright, and Baron-Cohen 2007) 2c.

While further evidence that ToM and Language are intimately connected will be presented in Chapter 3, I consider this discussion sufficient to proceed under the presupposition that both Language and ToM are evolved capacities.

If evolutionary psychology is the principle we take as for reverse-engineering, then what is our principled way of moving forward once we have our evolutionary-derived insight? It will either be totally novel or will be (or resemble) some already existent theory of linguistics (and, in fact, I suspect the latter is more likely to be the case).

You may think that there is no point in doing all of this if we are just going to end up where we were before, but in fact this approach has the potential to yield powerful, cognitively plausible insights into which types of theories are correct (e.g. Lexical Functional Grammar, Montague Grammar, Head-Driven Phrase Structure Grammar, Optimality Theory, Tree Adjoining Grammar, Relational Grammar, Transformational Grammar, etc.). Currently, much of linguistic theory is intuition-based. It is worth looking at the historical roots of this situation in order to understand why it is the case that intuition plays the role that it does in linguistics as well as to understand why we should want to move away from this methodology.

2.7 Intuition schmintuition

Chomsky (1959) writes "Insofar as independent neurophysiological evidence is not available, it is obvious that inferences concerning the structure of the organism are based on observation of behavior and outside events"(Chomsky 1967). He takes the lack of independent neurological data to be justification for basing linguistic research on the performance and intuitions of speakers. However, his

[13] A typical false belief task consists of the following: A character puts an object in a certain place (e.g. chocolate in a cupboard), then the character leaves the room and a second character moves the the item to new location. The first character re-enters the room and then the participant who has been watching this vignette is asked where the first character will look for the object. Children under the age of 3-4 and most autistic people will answer that the first character ill look in the new location rather than the old.

statement, while true in 1959, is simply no longer true in 2010. There *is* independent neurophysiological evidence available and while the lenses that we use to look at it vary in clarity, it is a good idea to begin to compare evidence based on outside events to the neurophysiological data that is available to see how the two match up. What this will result in, ultimately, is a model of language that (more) closely resembles what is happening in the brain, a kind of 'bio-model' of language.

To be clear, all cognitive processes are simply neurons firing in patterns. However, in examining the organization of language in the brain and studying how (when, where, why, etc.) these neurons fire we can extrapolate that data into a model which - while perhaps somewhat metaphorical - will reflect and help us to gain some insight into the cognitive reality of language. Let biological design inform theoretical design.

A side benefit to this way of doing things is that it would 'fact-check' current linguistic theory. I will describe this in more detail, but before motivating this I want to state something up front: I do not propose to start from scratch with Linguistic theory. Sampson (2001) speculates, "No doubt there are many cases where linguistic intuition and linguistic reality correspond. . . [b]ut the only way to check that is to describe languages on the basis of observable reality, and see how well speaker's intuitions match the description" (Sampson 2001). I second that. Intuitions are a useful guiding force but cannot serve as the bedrock of explanation.

Much of linguistic theory relies on the use of intuitions. I wish to point out five problems with a(n over) reliance on intuitions: (1) justification for reliance on intuition is no longer entirely valid (2) reliance on intuition may be circular (3) experimental evidence shows that intuitions cannot be entirely trusted and can facilitate the overlooking of certain potentially-interesting facts and patterns (4) non-linguists do not know their language as well as is commonly assumed (5) the intuitions of linguists tend to converge differently than the intuitions of non-linguists. I will address these respectively below.

2.7.1 Justification for reliance on intuitions no longer entirely valid

As already mentioned, there is now independent neurological data available with which to construct a theory. The need to rely on intuition has its (unfortunate) roots in the dawn of modern linguistics. Chomsky (1965) writes,

> Clearly, the actual data of linguistic performance will provide much evidence for determining the correctness of hypotheses about underlying linguistic structure, along with introspective reports (by the native speaker, or the linguist who has learned the language). This is the position that is universally adopted in practice, although there are methodological discussions that seem to imply a reluctance to use observed performance or introspective reports as evidence for some underlying reality. In brief, it is unfortunately the case that

> no adequate formalizable techniques are known for obtaining reliable information concerning the facts of linguistic structure (nor is this particularly surprising). There are, in other words, very few reliable experimental or data-processing procedures for obtaining significant information concerning the linguistic intuition of the native speaker. (Chomsky 1965) (emphasis mine)

Elsewhere, Chomsky describes his displeasure with the role of intuitions. In 1958 at the Third Texas Conference, he is described as saying that "he 'dislike[s] reliance on intuition as much as anyone, but if we are in such a bad state that it is only intuition that we are using, then', he feels, 'we should admit it'" (Matthews 1993). This reliance on intuitions received some early criticism. (As a matter of fact, most of the questioning of the role of intuition within the generative tradition seems to be restricted to its early years whereas modern criticism comes from outside the generative bubble.)

In a somewhat puzzling criticism of ordinary language philosophy, Fodor and Katz (1964) write, "All we have is over worked intuition. It is in this way that the ordinary-language philosopher's reliance upon linguistic intuition may have been responsible for his failure to develop an adequate theory of language." Then, in a footnote to that sentence, they (partially) write,

> One of the main dangers encountered in the construction of the rules of a linguistic theory is that they may be formulated so as to be workable only when an appeal to linguistic intuition is made. This means that in order for the rules to serve their intended purpose it is necessary that a fluent speaker exercise his linguistic abilities to guide their application. This, then, constitutes a vicious circularity: the rules are supposed to reconstruct the fluent speaker's abilities, yet they are unable to perform this function unless the speaker uses these abilities to apply them. As much of the abilities of the speaker as are required for the application of the rules, so much at least the rules fail to reconstruct. This, however, is not meant to imply that the appeal to linguistic intuition plays no role in theory construction in the study of language." (Fodor and Katz 1964)

They then go on to defend the role of intuitions in language concluding that "the appeal to linguistic intuition is question begging when intuitions replace well-defined theoretical constructs in an articulated system of description, or when intuitions are permitted to determine the application of rules. Intuition in its proper role is indispensable to the study of language, but misused, it vitiates such a study" (Fodor and Katz 1964). Two points seem obvious to me. The first is that they have not broken free of the circularity they articulate so well. (See below for discussion.) The claim that linguistic intuition should not replace well-defined theoretical constructs means nothing when the constructs are themselves built from intuition. (See section 2.7.3 for an example.) Second: what prevents one from using intuition to construct a counter-example that would invalidate a rule? They give primacy to the rule over intuition without providing any clear standard for when it is appropriate to deem a rule incorrect or in need of adjustment (this could easily stem from one's intuition about a cleverly chosen example).

2.7.2 Possible circularity

Further reasons for dissatisfaction which relate to circularity of argumentation are found in Searle (1969):

> I have no operational criteria for synonymy, ambiguity, nounhood, meaningfulness, or sentencehood. Furthermore, any criterion for any one of these concepts has to be consistent with my (our) knowledge or must be abandoned as inadequate. The starting point, then, for this study is that one knows such facts about language independently of any ability to provide criteria of the preferred kinds for such knowledge. Any appeal to a criterion presupposes the adequacy of the criterion and that adequacy can only be established by testing the criterion against examples such as these. The point is not that the claims made in linguistic characterizations cannot be justified in the absence of the preferred kinds of criteria, but rather that any proposed criterion cannot be justified in the absence of antecedent knowledge expressed by linguistic characterizations.

Yngve (1996) summarizes Searle's point: "In other words, one knows the answer beforehand independently of any criteria" (Yngve 1996). Yngve (1996) takes up this same point and hammers Chomskyan linguistics: "the philosophical methods of definition and explication of intuitive concepts have no place in science, and . . . we cannot build a proper science on an assumption that creates its own object of study" (Yngve 1996).

I am rather sympathetic, perhaps paradoxically, to both Searle/Yngve and Chomsky. We are, in some ways, much better off than Chomsky was when he wrote the passages above insofar as we have more than intuitions now, we have neurological data.

However, in a different sense, we are not much better off because the quality of that data is such that using it to inform a theory of language to the desired levels of precision is not entirely possible. What I propose is that we not ignore intuitions completely but rather begin to compare intuition-based theory with what neurological data is available in order to see how they match up.

Thus, I am more sympathetic to William Labov, who writes, "it must be acknowledged that linguistic analysis will always rely to a large extent on elicited judgments, the intuitions of the native speakers." (Labov 1996). So acknowledged[14] but see the next three sections for further evidence about the quality of those intuitions (and who holds them).

[14] Though I take exception to the degree linguistics will have to rely on intuitions.

2.7.3 Experimental evidence casts doubt on the quality of intuitions

Wasow and Arnold (2005) provide experimental evidence that shows intuitions about structural features of sentences are not directly nor consistently available to intuition. Specifically, they find experimental evidence against the claim "that only NPs dominating S can occur in shifted position" in heavy NP shift (Ross 1967) (but see (Wasow and Arnold 2005) for further discussion).

They define two levels of intuition: primary and secondary. Primary intuitions "are simply introspective judgments of a given linguistic expression's well-formed or of its meaning." Secondary intuitions are "intuitions about why a given expression is (or is not) well-formed or has the meaning it has" (Wasow and Arnold 2005). They present a case study based on an early claim of Chomsky's, namely that

> While . . . both [*the detective brought in the suspect*] and [*the detective brought the suspect in*] are grammatical, in general the separability of the preposition is determined by the complexity of the NP object. Thus we could scarcely have . . . *the detective brought the man who was accused of having stolen the automobile in.*
>
> It is interesting to note that it is apparently not the length in words of the object that determines the naturalness of the transformation, but rather, in some sense, its complexity. Thus 'they brought all the leaders of the riot in' seems more natural than 'they brought the man I saw in.' The latter, though shorter, is more complex . . . ((Chomsky 1975), (Wasow and Arnold 2005))

They note that these two paragraphs are interesting because they contain both primary (the first paragraph) and secondary (the second paragraph) intuitions. They tested Chomsky's secondary intuition by creating quadruples of sentences that are (1) of equal length but differing complexity (in Chomsky's terms) and (2) heavy NP shift .

1. Differing Complexity
   (a) The children took everything we said in.
   (b) The children took in everything we said
   (c) The children took all our instructions in.
   (d) The children took in all our instructions.
2. Heavy NP shift
   (a) The company sends what Americans don't buy to subsidiaries in other countries.
   (b) The company sends subsidiaries in other countries what Americans don't buy.

(c) The company sends any domestically unpopular products to subsidiaries in other countries.

(d) The company sends subsidiaries in other countries any domestically unpopular products.

Participants had to rank sentences on a scale from 1 ("completely unacceptable") to 4 ("fully acceptable") The results are as follows:

Construction	Acceptability score
Simple (took . . . in)	2.8 (avg.)
Simple (took in)	3.4 (avg.)
Complex (took . . . in)	1.8 (avg.)
Complex (took in)	3.3 (avg.)

They write that these results provide "some support" to Chomsky's intuitions however, they found that "about one-sixth of the time participants judged [split complex NP sentences like (1c)] to be no worse than awkward," pointing to the variability of intuitions which they describe as "the rule, not the exception."

The results for the heavy NP shift questions (like (2b) above) were a bit clearer. Participants significantly preferred the non-shifted construction in the simple cases (like (2a)) and the shifted construction in the complex NP cases (like (1b)). They further note that scores across the board were quite high and that "the 3.0 mean score for shifted examples with simple NPs undermines the claim made by Ross (1967, Rule 3.26) that only NPs dominating S can occur in the shifted position. This illustrates nicely the need to test secondary intuitions empirically."(Wasow and Arnold 2005)

This demonstrates that (a) a single person's intuitions cannot reflect the grammar (or the *l*anguage following from the discussion in 2.1) and (b) that intuition-based theoretical constructs (like Ross') can be disproved by experiments. Therefore, it would behoove linguists to use intuitions as valuable guides toward conducting experiments which can then confirm, refute, or inform hypotheses.

2.7.4 Non-linguists don't know their language as well as is commonly assumed

But Labov 1995 (also see (Labov 1975),(Sampson 1987), (Sampson 2001)) also calls for a greater emphasis on data by demonstrating that, apart from having a dubious scientific status, intuitions are also fallible. That is, speakers do not necessarily know their own language as well as is commonly believed. For example, Sampson (2001) quotes Labov (1975) regarding the dialectical (Midwestern United States) 'positive anymore'."

> Faced with a sentence like *John is smoking a lot anymore* they [speakers of the dialect -*DF*] said they had never heard it before, did not recognize it as English, thought it might mean 'not smoking',

> and showed the same sign of bewilderment that we get from . . . speakers outside the dialect area. This describes the behavior of Jack Greenberg, a 58-year-old builder raised in west Philadelphia. His introspective reactions were so convincing that we felt at first that we had to accept them as valid descriptions of his grammar. Yet two weeks later, he was overheard to say to a plumber, '*Do you know what's a lousy show anymore? Johnny Carson.*' ((Labov 1975), s omitted; (Sampson 2001))

Others have also noted that linguistic research should not solely rely on intuition. Chomsky (2004) writes, "It just seems absurd to restrict linguistics to the study of introspective judgments, as is very commonly done" (Chomsky 2004a). Intuitions need to be shored up with empirical data whenever possible. So, by 'fact-check' I mean to compare intuition-based theory with observable reality, following arguments that linguistics cannot or should not rely solely on intuitions since, if even scientifically valid, they can be faulty.

2.7.5 The intuitions of linguists and non-linguists are different

"Ah!" says the skeptic, "But linguists create linguistic theory and they are not lay people (like Mr. Greenberg, in the prior section), they are experts. Therefore, their intuitions are better and deserve more trust." Perhaps. But, Spencer (1972) took "150 exemplar sentences from 6 linguists' articles," presented them to 43 people with no linguistic training and found that the 43 people agreed on the acceptability or unacceptability of 80 percent of the sentences but "subjects shared intuitions with linguists in only half of the exemplars" (Spencer 1972). Furthermore, Sampson (Sampson 1987), (Sampson 2001) invokes Snow and Meijer (1977) who show that the intuitions of linguists tend to converge more often than those of lay people. Therefore, a consensus among linguists may be out of step with language in the wild.

The problem, then, becomes obvious. Linguistic theory is largely based on the intuitions of linguists whose intuitions coincide (more or less) about the language of their community but not with the language of community *at large* - the language they purport to be describing. (Also see (Wasow and Arnold 2005) for criticism that linguists often cite psycholinguistic evidence that favors their theory but ignore that which does not.)

2.8 Conclusion

Put plainly, I have outlined a research approach that boils down to this: do "good old-fashioned cognitive science." Since the idea was conceived of but never implemented (due to technological restrictions), no methodology exists for doing good old fashioned cognitive science. Here I have tried to sketch what that methodology could be and why.

I have proposed that we reverse-engineer (the natural model of) language in the brain in order to understand what our (artificial, somewhat metaphorical) model should look like. I have proposed that the best approach for reverse-engineering is evolutionary psychology. What's more, there is no inherent problem with the structure of the brain that invalidates this proposal, indeed the structure of the brain lends itself to this approach; the only impetus is the limit of technology.

A further benefit is that this approach can be progressively blended into current approaches by backing up or disproving intuition-based reasoning which, I have argued (echoing many others), is not as reliable as some have claimed it to be in the past.

Above I said that I wanted to compare evidence of ToM being responsible with pragmatic language to evidence suggesting that the right hemisphere plays some role in language. Now I want to draw a distinction between two different kinds of pragmatic language. That is, a theory of pragmatics will have to do at least two things (1) account for intrasentential phenomenon (i.e. sentence integration) and (2) pragmatic instantiations of language. In the next two chapters I will argue that ToM accounts for (1) and the right hemisphere accounts for (2).

Chapter 3

Theory of Mind as a partially pragmatic module

"It would be surprising if the intentional content of thought and language were independent phenomena, and so it is natural to try to reduce one to another to find some common explanation for both." – Ned Block 1995

3.1 Introduction

ToM is a fairly new concept (Premack and Woodruff 1978) and evidence of its neurological basis is even newer (Baron-Cohen, Ring, Moriarty, Shmitz, Costa, and Ell 1994). As a result, there is very little that is definitive or clear about what parts of the brain contribute to one's ToM, what specific capacities they endow normally developing adults with, or even what exactly happens when these brain regions do not develop normally - for example, in people with ASD or individuals with schizophrenia (Abu-Akel 2003; Brüne 2005).

Despite the small state and relatively young nature of the body of research on ToM, what is known is substantial and interesting enough to warrant examination. Before beginning an exploration of the neuroimaging literature, it may be useful to review what we might expect the ToM mechanism or module to look like in the brain.

Stone et al. (1998) write of ToM that "such a complex, multi-component, cognitive ability is unlikely to be localized in a small region of the cortex" (Stone, Baron-Cohen, and Knight 1998) and propose that the neural correlates of ToM are likely to be something more along the lines of a "circuit." However, there may be some danger in this view as well.

Steven Pinker (in an interview with the website BigThink), in response to the (rhetorical) question "How is the mind organized into components?", responds that no one knows if the components are "discrete slabs of real estate[1]" or "completely inter-dispersed like the hard disk on your computer when it's fragmented[2]" or "something in between." This is simply to point out that these are open questions, but the current state of neuroimaging research with respect to ToM suggests that it may be a good candidate for the 'in between' analysis of brain organization.

As already discussed in Chapter 2 (section 2.6) there is good reason to believe that the ability to acquire language(s) and ToM are evolved capacities with a genetic basis. (Since genes build the body, including the brain, something having a genetic basis confers the property of innateness.) However there is even further evidence that the two abilities are closely related in other ways.

There is an ever-growing body of evidence that the two faculties co-develop and that the maturation of one may be dependent on the other.

Morgan and Kegl (2006) present compelling evidence that deaf children (signers of Nicaraguan Sign Language) who acquire language late incur ToM deficits. "Access to sign language by 10 years of age," they write, "with possible advantages in language fluency was a strong predictor of performance on both false belief task[3] and mental state narrative task"(Morgan and Kegl 2006). They also make reference to a number of other studies which come to similar conclusions based on deaf signers and the age at which they were first exposed to language (Peterson 2004; Courtin and Melot 2005; Harris, de Rosnay, and Pons 2005).

Further unique evidence from Nicaragua also suggests interconnectedness between the two capacities. Understanding a few aspects of the history of Nicaraguan Sign Language ("NSL") is crucial to understanding its importance. But before beginning, it is worth noting that linguists and psychologists do not consider there to be any important differences between signed and spoken languages. For example, it is well-understood that sign language even has a phonological component (Sandler ip).

Before 1977 deaf children in Nicaragua were largely isolated from one another because there was no educational system (or center for) deaf people. As a result of this lack of interaction, no sign language emerged and people developed "home signs" which are communication systems developed so one can communicate with one's family; home signing systems show no universal characteristics and vary greatly from one person to the next and therefore are not considered languages in the technical sense (Senghas, Kita, and Özyürek 2004).

This changed in 1977 when an "elementary school for special education" (Senghas, Kita, and Özyürek 2004) opened which gave deaf children a place to interact socially and linguistically. The teaching was done in Spanish, but the children soon developed a gestural communication system which evolved into an early sign language. Senghas, et al. (2004) write "through continued use, both

[1]Though he also says that he thinks this is kind of unlikely.

[2]He continues, "So the different parts that belong to one system are scattered all over the place and work because of their intricate connections, but we'll never be able to see them as blobs on a brain scan."

[3]See fn. 13, p.28) for a description of a false belief task.

in and out of school, the growing language has been passed down and relearned naturally every year since, as each new wave of children entered the community" (Senghas, Kita, and Özyürek 2004).

Pyers and Senghas (2009) conducted a 2-year study on speakers of NSL and found "language learning, over and above social experience, drives the development of a mature theory of mind" (Pyers and Senghas 2009). Their study followed signers of the first and second cohorts. (Signers of NSL are typically divided into three cohorts: the first are those who were exposed to NSL before 1984, the second between 1984-1993, and the third from 1993 to present.)

Both cohorts have similar social experience - school, comparable social networks, "access to a modern urban culture, including public transportation and television" [Ibid]. The only significant difference between the groups in terms of nonlinguistic factors is the first cohort, being older, has had a decade more experience. The only difference between them linguistically is that the grammar of the second cohort is more complex than that of the first cohort. Note also that the second cohort was exposed to language at an earlier age than those in the first cohort.

They compared speakers from both cohorts in 2001 and then again in 2003. In 2001, they administered standard false belief tasks (in video or cartoon form; see page 28, fn. 13) and tallied the number of references to mental state terms ("terms of belief and knowledge (e.g. *think*, *know*)" but not *desire*) and found that "the second cohort had evidently developed mental-state vocabulary that the first cohort lacked" (Pyers and Senghas 2009). Furthermore, they found that first cohort signers who produced no mental state verbs failed false belief tasks.

Interestingly, when they went back and re-ran the same experiments in 2003 they found that the first cohort signers had acquired mental state vocabulary (through greater interaction with second cohort signers) and improved their ability on false belief tasks to the point that they no longer differed significantly from the second cohort. Thus it seems that exposure to language (within the critical window) can influence ToM development, though the extent to which this is the case is uncertain (they note language is not the only factor nourishing ToM development - social interaction may also help).

Jill de Villiers (2000, 2007) has suggested that "the interface between language and Theory of Mind is bidirectional" based on evidence from developmental psychology and linguistics (e.g. that the epistemic "must[4]" does not appear in the speech of normally developing children until 3.5-4.5 years of age in several different languages, see (de Villiers 2007) for further discussion).

Paul Bloom (2002) has provided convincing evidence that ToM is employed in the acquisition of nouns ("Names for things"). He cites experimental evidence which points to two key factors in lexical learning: "bias against lexical overlap" and the "whole object bias." The bias against lexical overlap stems from (1) an expectation that speakers will be maximally informative and (2) a feature stemming from ToM (Clark 1997) namely mutual exclusivity. Mutual

[4]Obviously, I mean the epistemic "must" in English and its equivalent in other languages.

exclusivity refers to an expectation speakers have "to use familiar words to refer to familiar things"(Bloom 2002).

There is strong evidence that the whole object bias simply exists (see Bloom 2002) and therefore does not need to be defended which is not to say that it does not need to be explained. The explanation Bloom provides is that "children attend to objects when learning new words because objects are salient individuals and that children attend to objects when learning new words because they believe that adults intend to refer to objects" (Bloom 2002). This is just to say that the relationship between Language and ToM appears not to be something that comes online, so to speak, but may be present at birth and learning more about the relationship between these two faculties may help gain further insight into language acquisition.

Thus, there is already sufficient reason to believe that the Language Faculty ("LF") and ToM are connected in some way based on empirical observation. Below I will compare three studies that, when taken together, single out at least two neural correlates that are interesting candidates for a pragmatic processor.

Remember that to make my case convincing I have to provide for correlation as well as causation. While not wholly definitive, I hope to paint a picture that strongly suggests a relationship between the two faculties.

3.2 Evidence for correlation

"[I]t is likely that the conceptions of TOM and pragmatic understanding are highly related and possibly inextricably entwined, rendering it unlikely that a clear causal direction exists between the two." – Martin and McDonald 2003

In a review of eleven different fMRI studies conducted into story-telling, Mar (2004) writes "Story comprehension appears to entail a network of frontal, temporal, and cingulated areas that support working-memory and theory-of-mind processes" (Mar 2004). More specifically, he writes, "[a]ctivations from available studies of narrative processing appear to cluster in five regions . . ." (Mar 2004) which are listed in the table below.

region	Brodmann's areas	Activation pattern
Medial prefrontal cortex	8,10,11	bilateral
Lateral prefrontal cortex	44, 45, 47	right-hemispheric
Temporoparietal region	21,22/39, 40	bilateral
Anterior temporal region	38	bilateral
Posterior cingulated cortex	23/31, 31	bilateral

Note that the anterior temporal region listed above also includes the temporal poles. Take Mar's (2004) results compared to Schmithorst et al. (2006). Schmithorst et al. (2006) is a large-scale (313 person) fMRI study of 5- to 18-year-olds in which they were presented with "simple stories without an element of "theory of mind" [which] were quite short, obviating the necessity of higher-order integrative processes (between groups of events) in order to make the difficulty level suitable for young children" (Schmithorst, Holland, and Plante 2006). Thus, the study is large which is beneficial because of the "excellent signal to noise ratio" (Schmithorst, Holland, and Plante 2006).

An advantage of the Schmithorst et al. (2006) study is the methodology they employed. Mar reviews fMRI studies which "are relatively good in terms of their ability to localize neural activity," writes Anderson, "but poor in their ability to trace out the time course of that activity" (Anderson 1984).

However Schmithorst et al. (2006) performed an independent component analysis ("ICA", discussed below) in which temporal distinctions can be made. Thus, we can not only compare the anatomical regions that are activated, but also in what sense that region may be related to pragmatic processing (as opposed to other processes relative to phonology, syntax, semantics, theory of mind) since one would predict a pragmatic process as happening later temporally (as linguistic information accumulates[5].

In the Schmithorst et al. (2006) study, subjects were read short[6] stories that required no mentalizing[7] and asked to answer questions about the story that they heard and then performed a group ICA of the data. "The advantage of ICA," they write, "a data-driven approach, lies in its ability to mine more information from the data than is possible via conventional hypothesis-driven analyses" (Schmithorst, Holland, and Plante 2006). In a sense, then, an ICA analysis is beneficial because it helps us remove any prior biases we may have before analyzing the data.

ICA is a method of statistical analysis which solves the "sources of separation problem" (Comon 1994) inherent in much data analysis. "Sources of separation" refers to the fact that "most measured quantities are actually mixtures of other quantities" (Stone 2005). For example, fMRI scans measure both spatial and temporal aspects of activation. "If the "activations" do not have a systematic overlap in time and/or space then the distributions can be considered independent components" (Calhoun, Adali, Hansen, Larsen, and Pekar 2003). Calhoun et al. (2003) write, "ICA can reveal inter-subject and inter-event differences in the temporal dynamics [of brain function]. A strength of ICA is its ability to reveal dynamics for which a temporal model is not available. Spatial ICA also works well for fMRI as it is often the case that one is interested in spatially distributed brain networks" (Calhoun, Adali, Hansen, Larsen, and Pekar 2003).

ICA cannot "reveal the precise cognitive correlates of the components found" (Schmithorst, Holland, and Plante 2006) and thus can only serve to corroborate extant hypothesis or motivate new ones when then need to be independently confirmed. Therefore, the results of this study are appropriate to compare to Mar's study. Schmithorst et al. (2006) identified six distinct areas. They are as follows (the areas are presented from the earliest activated area at the top to the latests activated area at the bottom):

[5]This fact is observed by Schmithorst et al. (2006) as is mentioned in the following.
[6]9, 10, or 11 sentences long.
[7]*Mentalizing* is the verb for 'using one's Theory of Mind.'

Area of activation	Brodmann's area	Activation pattern
Primary auditory cortex	41, 42	Bilateral
Superior temporal gyrus	22	Bilateral
Inf. par. lob. and inf. frntl. gyrus	21, 46, 44/45, 40, 8	Left-Lateralized
Hippocampus		Bilateral
Superior temporal gyrus	22	Bilateral
ang. gyrus and pre/post. cing.	39, 7/31	Bilateral

See footnotes:[8], [9], [10]. As a brief aside before discussing the data, it is worth noting that we can be reasonably sure that the data is not 'noise' (that is due to non-experimental conditions). This is because, as already mentioned, the large sampling size provides an excellent 'signal' to noise ratio.

Furthermore, it has been suggested that "physiologic noise is the dominant factor in fMRI studies. In the ICA model these "noises" are often not explicitly modeled, but rather manifested as separate components" (Calhoun, Adali, Hansen, Larsen, and Pekar 2003). Since signals that control physiological processes (breathing, heart beating, etc.) come from "ventricles (fluid-filled regions of the brain) and areas with large blood vessels present, respectively" (Calhoun, Adali, Hansen, Larsen, and Pekar 2003) we can disregard components related to these areas. However, this was not necessary here as none of these components corresponding with these areas were measured. End aside.

The results in the beginning are not terribly surprising. Activation in the auditory cortex is expected (though it is curious why it is absent from Mar's (2004) findings). Activation of the superior temporal gyrus (BA 22) is also expected as this corresponds to Wernicke's area[11] (on the left, the activation on the right will be discussed below). The left-lateralized activation of the inferior parietal lobule and inferior frontal gyrus (network) is not surprising as this corresponds to Broca's area. Activation of the hippocampus, as Schmithorst et al. (2006) note, is also not surprising given that it is "critical to human memory" (Anderson 1984).

What is interesting is the late bilateral activation of the posterior superior temporal gyrus ("pSTG") and angular gyrus. Not only that they are active late in processing, but that the activation is bilateral. They write,

> [T]he existence of a component with right hemisphere laterality ([the pSTG]) supports the intriguing possibility that there may in fact exist a right-hemispheric dominant network for higher-order language processing involved with speech and/or narrative comprehension as hypothesized previously (Meyer et al., 2000). (Schmithorst, Holland, and Plante 2006).

[8] Inferior parietal lobule and inferior frontal gyrus (network)

[9] The most posterior part of of the superior temporal gyrus with respect to the second activation of this area

[10] Angular gyrus and the precuneus/posterior cingulate

[11] Interestingly, this component showed the least inter-subject variability and displayed the highest correlation coefficient of any of the components in the experiment.

They further point out that "the delayed rise in hemodynamic response supports the interpretation that the posterior STG may be involved in higher-order integrative processes, as the process of integrating the content of the sentences into the larger narrative will not begin until a certain period of time after the beginning of the story" and further conclude that this region is responsible for "higher-order integrative processes" (Schmithorst, Holland, and Plante 2006).

These two studies taken together show that the pSTG is activated in narrative comprehension whether or not there is a ToM component. But can it be activated by non-verbal tests that are designed to elicit ToM activity? Yes. Schultz et al. (2004) showed participants video of geometrical shapes one of which was 'chasing' the other. The chaser used one of two strategies: chase the object or jump in front of it - as though it had predicted its path - and catch it when the object arrives at the same location. They found that "[bilateral] activation in the [posterior] superior temporal gyrus . . . increases when a simple object appears to chase another object by understanding the target's goal and predicting its movement instead of simply following it" (Schultz, Imamizu, Kawato, and Frith 2004).

This result indicates that the pSTG is not simply activated by language (ToM component or not) but also by language-free ToM tasks. The posterior part of the superior temporal gyrus is activated in linguistic tasks with no ToM component, ToM component with no language and, and linguistic tasks with a ToM component. One may take this as evidence of a domain-general component of the brain, one that can be utilized for linguistic or ToM tasks, however, recall the 'conversation' with Pinker, above. It could just as easily be the case that ToM neurons and Language neurons are so intertwined (or "fragmented") in this area that we simply cannot tell when one set is firing and the other is not so either task gives the appearance of using the same gross areas of neocortex.

Both Schmithorst et al. (2006) and Mar (2004) also show that the posterior cingulated cortex is activated. Unfortunately, as the role the posterior cingulated cortex plays "remains fairly poorly understood" (Adolphs and Spezio 2009) it is difficult to draw any concrete conclusions. Saxe and Powell (2006) found that the posterior cingulated cortex was "recruited selectively when subjects read stories about a protagonist's thoughts or beliefs, but not when they read about subjective, internal physical feelings or other socially relevant information, such as appearance and personality attributes" (Saxe and Powell 2006). The Schmithorst et al. (2006) result, measuring stories with no ToM component directly contradicts this.

One hypothesis could be that this region is activated by narration, but that would not explain the fact that it becomes significantly less active during one type of story (about thoughts) and not others (about appearance or bodily sensation).

While the exact role of the posterior cingulated cortex remains fairly poorly understood, it is a well established fact that it is part of the limbic system and plays a role in the processing of emotions. Actually, its significance will become a bit clearer in Chapter 5 when I discuss the relationship between language and cheater-detection. However, for now I will leave the topic.

The correlation of activity in pSTG, angular gyri and possibly the posterior cingulated cortex for language tasks that do involve ToM and that do not involve ToM suggests that this part of the ToM processor may play a role in language processing. However, as correlation does not imply causation it must be shown that damage to the identified areas would cause some kind of impairment or problem with pragmatic uses of language.

3.3 Evidence for causation

The evidence for localization gives rise to the feature of predictable breakdown. This is to say that if you correlate a function to a specific brain region, the function should be lost if the area is damaged; if an area predicted to be responsible for a certain function is damaged and the function is not lost then the theory is likely incorrect. Therefore, it is predicted that damage to the pSTG, angular gyri or the cingulated gyrus would cause some form of pragmatic impairment.

Actually, Klein (2000), reporting on Amyes and Nielsen (1955), notes that "Damage to the cingulate gyrus causes a person to stop talking and walking" (Klein 2000). This kind of evidence - damage to a particular area causing an inability to communicate at all - would be rather striking evidence in support of the thesis presented here, especially in light of the previous discussion. However, this result is rather global (since walking is affected) and not specific-to-language as predicted.

Surian et al. (1996) found that children with autism (widely believed to be, in part, a ToM disorder) "were impaired at detecting pragmatic violations, compared to both normal children and children with specific language impairment matched on linguistic development" (Surian, Baron-Cohen, and der Lily 1996). They tested the three groups by playing pre-recorded conversations on tapes which were then acted out by puppets. At the end of the recording one character would ask a question and the other two puppets (there were three voices/puppets in total) would give different answers one being canonical and another violating a Gricean Maxim. Then the researcher would ask the participant to point to the puppet that said "something funny or silly."

Of the eight ASD participants, only three passed the test and these same three individuals also passed a ToM test showing that it is not autism, per se, that affects pragmatic competence but rather ToM. Interestingly, the NT children and children with SLI "only found it difficult to evaluate the violations of the two Maxims of Quantity" (Surian, Baron-Cohen, and der Lily 1996). This suggests a connection between ToM (whether ASD or not) and the ability to identify felicitous conversation. That being said, autism is a complex and poorly understood condition and one should be wary of drawing too strong a conclusion (e.g. extrapolate to NT people) from it. In keeping with the approach described above, ideally we would want evidence from NT people who have suffered brain damage.

As already mentioned, I cannot say much about the role of the cingulated cortex. I can, however, say something about the pSTG and the angular gyrus.

Though the evidence is neither definitive nor conclusive it is worth looking at, briefly, in order to see that what is known is suggestive.

3.3.1 Posterior superior temporal gyrus

Damage to the pSTG can cause conduit aphasia (also called conduction aphasia)(Hickok, Erhard, Kassubek, Helms-Tillery, Naeve-Velguth, Strupp, Strick, and Ugurbil 2000; Quigg and Fountain 1999), and abnormalities with this region are correlated with schizophrenia (Jacobsen, Giedd, Castellanos, Vaituzis, Hamburger, Kumra, and Rapoport 1988; Nestor, Shenton, McCarley, Haimson, Smith, O'Donnell, Kimble, Kikinis, and Jolesz 1993; McCarley, Saltsbury, Hirayasu, Yurgelun-Todd, Tohen, Zarate, Kikinis, Jolesz, and Shenton 2002; Menon, Barta, Aylward, Richards, Vaugn, Tien, Harris, and Pearlson 1995), and dyslexia (Shaywitz, Shaywitz, Pugh, Fulbright, Constable, Mencl, Shankweiler, Liberman, Skudlarski, Fletcher, Katz, Marchione, Lacadie, Gatenby, and Gore 1998; Eden and Zeffiro 1998). I will briefly discus the first two of these respectively.

"Conduction aphasia," writes Hickok et al. (2002) "is an acquired disorder of language characterized by good auditory comprehension, fluent speech production, relatively poor speech repetition, frequent phomenic errors in production, and naming difficulties" (Hickok, Erhard, Kassubek, Helms-Tillery, Naeve-Velguth, Strupp, Strick, and Ugurbil 2000). It is usually caused by a lesion damaging the arcuate fasciculus which connects Broca's area to Wernicke's area, but damage to the pSTG seems to result in the same language profile as conduction aphasia. Hickok et al. (2002) report an object-naming PET study that was performed to test whether the pSTG was "activated in a speech production task even when there is no verbal auditory input" (Hickok, Erhard, Kassubek, Helms-Tillery, Naeve-Velguth, Strupp, Strick, and Ugurbil 2000). They put headphones on the participants and played white noise so that they could not hear their own voice. They found that there was, in fact, activation in the left pSTG when participants 'said' something without actually producing any noise (they thought the word and mouthed it but no sound left their mouth). The authors conclude that this area of the brain is involved with phonological output.

Furthermore, Quigg and Fountain (1999) elicited this same kind of aphasia via electrical stimulus to the left pSTG adding additional support to the notion that damage to this area results in an inability of contextualize linguistic input. They found that "(1) recitation was normal; (2) naming "thumb," "index finger," and "tie" was normal; (3) command following -"show me your thumb"- was normal; and (4) repetition "Bill Clinton", "Commonwealth of Virginia" was reproducibly impaired" (Quigg and Fountain 1999).

It is something of a mystery why only the ability to repeat sentences should be impaired while the ability to have a conversation is not (Goodglass 1992). This is likely due to the fact several different neural substrates are responsible for different aspects of the complex task of linguistic interaction.

Goodglass (1992) reports that it has been suggested that "conduction aphasia is a disorder of sequential programming of linguistic elements at the level of the phoneme"(Goodglass 1992). Another symptom of conduction aphasia is speech "marred by more or less frequent errors in the selection and sequencing of phonemes and syllables; these may be omitted, substituted, or transposed, creating "literal paraphasias"" (Goodglass 1992).

The only other evidence with respect to the pSTG for causation comes from schizophrenic individuals. While schizophrenia itself is orthogonal to my main thesis, it is related because as mentioned above ToM deficits are thought to contribute to the underlying cause of schizophrenia (Abu-Akel 2003; Brüne 2005). Therefore, a brief discussion is warranted.

Menon et al. (1995) found that "a significant correlation was detected between delusion scores in schizophrenics and the total volume of the left dominant posterior STG" (Menon, Barta, Aylward, Richards, Vaugn, Tien, Harris, and Pearlson 1995) which give us some reason to think that this area is responsible for higher-level language cognition but other data (Barta, Pearlson, Powers, and Stephanie Richards 1990) complicate the picture and it is difficult to draw solid conclusions[12].

Frith (1992) reports on several studies (Andreasen 1985; Frith and Allen 1988; McGrath 1991) that support the idea that "only the highest levels of language processes are impaired in schizophrenia . . . In schizophrenic speech some or all of [Grice's] maxims are broken" (Frith 1992). He further points out that these "abnormalities are expressive rather than receptive" (Frith 1992). This is interesting not only because it speaks to the depth of the deficit but also because it mirrors classical (Broca's) aphasia patterns in terms of production impairment with unimpaired comprehension (see next section 4.1 for example).

In sum, this area is implicated in speech production (specifically phonological output) and the ability to repeat what one hears[13], and, based on evidence from schizophrenics who have reduced brain mass in the pSTG, higher-level language cognition, including the Gricean Maxims. While this last kind of evidence seems the strongest, one must be a bit careful about drawing too strong a conclusion from data regarding poorly understood disorders. This evidence taken in juxtaposition with the less suggestive but more reliable PET and electrical stimulation evidence along with the evidence for correlation provides reason to be cautiously optimistic that this part of the brain is involved in the pragmatic

[12]The Scale for the Positive Assessment of Symptoms (SAPS) (Andreasen 1984). A positive symptom of schizophrenia is a symptom that is 'abnormal by its presence' (as opposed to a negative symptom which is present in normal people but not in those with schizophrenia. One example of a positive symptom is auditory hallucinations which are commonly described as ""discussing my actions", "talking to me", "repeating my thoughts"" (Frith 1992). Thus this may serve as an indication between the interface of ToM (as evidenced by the postulation of an intentional mind behind the voices (they speak to and comment about the person, it is not as though they hear something equivalent to a radio being left on) and higher-order language functions (the sentences form a coherent gestalt). However, other studies (Barta, Pearlson, Powers, and Stephanie Richards 1990) have provided evidence that it is the anterior and not the posterior part of the superior temporal gyrus that is responsible for these hallucinations.

[13]Or *Interchanability*, one of the thirteen essential design features of language (Hockett 1960)

use of language.

3.3.2 Angular gyri

V.S. Ramachandran (2004) has noted the role of this area in synesthesia, a neurological condition in which cross-modal interaction is present in certain domains. An example is consistently seeing (black) letters in (particular, non-black) colors.

He explains this in terms of real estate: "the angular gyrus . . . is strategically located at the cross roads between the parietal lobe (concerned with touch and proprioception), the temporal lobe (concerned with hearing [and, as we have seen language-*DF*] and the occipital lobe (concerned with vision)" (Ramachandran 2004).

This area is at the juncture of vision (which I have already established as part of the FLB), hearing (obviously part of the FLB), and proprioception (we may allow gesture and deictic pointing as part of the FLB). This is an area where integration takes place and, given the findings by Schmithorst et al. (2006), may be an area of linguistic and non-linguistic (e.g. contextual) integration. Ramachandran also suggests that this area is crucial to the understanding of metaphors, though each hemisphere[14] may also be responsible for different kinds of metaphors.

Specifically, he suggests that the left angular gyrus is responsible for cross-modal metaphors ("loud shirt", "sharp cheese") and the right angular gyrus is responsible for spatial metaphors ("stepped down" from his post). He notes, that this hypothesis needs to be tested but also notes that he "tested five patients with left angular gyrus lesions and found that they were abysmal at interpreting proverbs and metaphors" (Ramachandran 2004).

This is not only interesting because this part of the brain has been shown to have a relationship with metaphor (and thus language) but also because language is heavily metaphorical and we might, therefore, expect a brain region involved in the processing of metaphor to be involved in pragmatic processing. For example Pinker (2007b) makes a very strong case that "metaphors are omnipresent in language" (Pinker 2007b) by demonstrating that there is a theory of space in the prepositions, a theory matter in our nouns, a theory of time in tense systems and a theory of causality in verbs.

To convince you of this, I will, all too briefly, partially reiterate Pinker's case for space, time, and matter (leaving out causality, but see Pinker 2007b).

In the broadest stroke, Pinker argues that language metaphorically schematizes space, time, and matter and this schematization gives rise to a dependence on metaphor.

With respect to prepositions he argues in the following vein.

> The schematic modeling of shapes is the kind of geometry that defines most spacial terms in English and other languages. A preposition, for example, locates a figure relative to a reference object, and

[14]Note the activation in the Schmithorst et al. study is bilateral

in doing so has to specify something about the shape of the figure and something about the shape of the reference object. The most common kind of preposition, like in, on, near, and at says nothing at all about the figure being located, treating it as a 0-D point or lump. . . anything at all can be in or on something, whether it is a pebble, a pencil, or a pad, and it doesn't matter which way it is pointing. The reference object, in contrast, has to have a certain geometry for a preposition to apply. In, for example, requires a 2-D or 3-D cavity. Along needs a 1-D primary axis: a bug can walk along a pencil, but not a CD, though it can walk along the edge of the CD. Through demands a 2-D opening or an aggregate, as when a fish swims through water or a bear runs through the woods. Inside demands an enclosure, usually 3-D. (Pinker 2007b)

This demonstrates that when speaking about objects in relation to one and other we often times discuss them metaphorically by altering their dimensionality. He also argues that time is schematized into a ribbon consisting of The Past (ranging from a few seconds ago to the birth of the universe) The Specious Present (roughly three seconds long) and The Future (a few second from now into eternity). He also points out an underlying conceptual structure between space and time in language:

Tenses (locations in time) work like prepositions and other spatial terms (locations in space). A tense locates a situation only relative to a reference point (the moment of speaking or a reference event), rather than in fixed coordinates such as the clock and the calendar. It cares about direction (before or after), but ignores absolute distance (days, hours, seconds). (Pinker 2007b)

Tense, he explains, is less complicated than prepositions because time is only one dimensional where as space is, perceptually speaking, three-dimensional; therefore, there are fewer combinatorial possibilities which yields lesser complexity. He describes how we talk about nouns in terms of an "intuitive materials-science" (Pinker 2007b). He discusses the distinction between mass and count nouns and notes that mass nouns "act like the plurals of count nouns . . . [t]he overlap in grammar reflects a similarity in the way we conceive of substances . . . " (Pinker 2007b). He concludes that "Our ability to think about things and stuff is surely rooted in our perception of lumps [roughly count nouns -*DF*] and guck [roughly mass nouns-*DF*] in the physical world, we easily extend it to the world of ideas" (Pinker 2007b). He sketches the underlying architecture thusly,

Just as in the realm of matter we beheld bounded objects (*cup*) and unbounded substances (*plastic*), in the realm of time we see bounded accomplishments (*draw a circle*) and unbounded activities (*jog*). Just as we met substance words that named homogeneous aggregates (*mud*) and plurals that name aggregates made of individuals (*pebbles*), we now meet durative verbs that name a homogeneous

> action (like *slide*) and iterative verbs that name a series of actions (like *pound*, *beat*, and *rock*). (Pinker 2007b).

So why lumps and guck and not some other distinction? "Perhaps if humans could see the crystals, fibers, cells, and atoms making up matter, we would never have developed a count-mass distinction in the first place," (Pinker 2007b) reminding us of Richard Dawkins' concept of Middle World which holds that "Our brains [have] evolved to help us survive within the scale and orders of magnitude within which we exist" (Twist 2005).

3.4 Conclusion

There is strong evidence for correlation but the evidence for causation is murky and diffuse (though provocative). Potts (2010) writes, "Pragmatics is central to the theory of linguistic meaning because, to paraphrase Levinson (2000) (Levinson 2000), the encoded content of the sentences we utter is only the barest sketch of what we actually communicate with those utterances." That is because we are not only communicating in the digital medium of language but also on the non-linguistic channel, ToM. He continues,

> Utterance interpretation involves complex interactions among (i) semantic content, (ii) the context of the utterance, and (iii) general pragmatic pressures (of which Grice's maxims are one conception). The starting point for a formal pragmatics is the observation that speakers agree to a remarkable extent on the interpretations of the utterances they hear, suggesting that there are deep regularities across speakers, utterance contexts and sentence types in how (i)-(iii) interact. (Potts 2010)

I hope to have shed a little bit of light on why speakers should agree on how (i)-(iii) interact. (i) is due to the digital nature of language and the Principle of Compositionality which takes the whole of the meaning to be the sum of it's parts (or the 'literal meaning' as discussed in the previous chapters). (iii) likely is a result of ToM processes. ToM processes are made use of in the following ways: (a) being able to understand people in a context, (b) being able to understand their intentional states, and (c) integrating these facts with the atmosphere (formal, friendly, informal). In fact, when one examines the maxims (see section 4.3 for a list) in this light, the intentionality in all of them become stark. (ii), above, will be the be addressed in the next chapter.

The above discussion was motivated by (a) the close relationship between ToM and the LF(B/N) and (b) by the fact that compositional linguistic theory more or less was limited to sentence-level discussion (an irony in some ways given the fact that Transformations, an early driving force in syntactic theory - which largely diverted research away from pragmatics- was originally developed by Harris (1952) (Harris 1952) for discourse analysis). The next section is motivated by instantiations of pragmatic uses of language.

Chapter 4

The right hemisphere

"[The] avoidance of the metaphorical, this penchant for the concrete particular, stands in striking contrast to the speech of [RHD] patients, which is in fact characterized by certain strange features: a flat monotonous tone, and a tendency to joke, though with a cutting edge, or to make improbable puns. An individual with right-hemisphere disease will typically call his physician "Doc" and make light of the paralyzed arm: "Oh, I don't need them gills to get around, anyway." Asked for the name of the institution he is in, he may proclaim, "Mount Cyanide Hospital." Although such retorts might superficially seem to indicate a superior sense of humor or irony, they actually signal the patient's unrealistic assessment of his situation - in minimizing the true extent of the injury, or embracing what may appear to be a form of "gallows humor," he demonstrates his unawareness of the full seriousness of his plight. . . . And in his spontaneous speech, the patient again seems to neglect the emotional quality of the present situation or the extent of his impairment in favour of dispensing wisecracks - these, in turn, seem to reflect the operation of a "joke machine" rather than the holistic integration of feelings, situational cues, and interpersonal relations appropriate to a serious interview. Such an individual can provide somewhat amusing "one-liners," but hardly seems likely to produce a balanced, developed work of art. However, to the extent that the patient was merely called upon to adapt or reformulate materials he had penned before (e.g. legal briefs, cooking recipes, letters to friends, bureaucratic memoranda), he might succeed reasonably well." – Howard Gardner 1976

"A version of Weizenbaum's ELIZA . . . was classified as human [in a Turing Test-*DF*] by five of ten judges . . . the winning program's topic was "whimsical conversation," a "smart -aleck" way of interacting that all the judges fell in love with immediately, and one that would have had the same effect, even without set topics." – Ned Block 1995

4.1 Introduction

In the Introduction, I stated that the right hemisphere seems to be specialized to process novel stimuli and the left hemisphere seems to be responsible for processing structured or recurring data. Here I wish to explore the idea that it may be profitable to think of the right hemisphere as the 'pragmatic' hemisphere, or a 'novelty detector' and the left hemisphere as the hemisphere of structure. Traditional language areas responsible for syntax (Broca's area), and semantics (Wernicke's area) are left-lateralized, so, beginning naively, this idea makes sense.

Thus far, the claim for hemispheric specialization has been made solely on the basis of language data but there are reasons independent of language to believe this is more broadly true. Dehaene et al. (1999) "taught," bilingual participants, "a set of exact or approximate sums of two-digit numbers in one of their two languages. In the exact addition condition, subjects selected the correct sum from two numerically close numbers. In the approximate addition condition, they were asked to estimate the result and select the closest number" (Dehaene, Spelke, Pinel, Stanescu, and Tsivkin 1999).

They found that subjects performed better on the exact addition problems when presented in the training language. This "provided evidence that the arithmetic knowledge acquired during training with exact problems was stored in a language-specific format and showed language-switching cost due to the required internal translation of the arithmetic problem. For approximate addition, performance generalized without cost to novel problems in the same range of magnitude, providing evidence that new knowledge was stored using a [more general -*DF*] number magnitude format." (Dehaene, Spelke, Pinel, Stanescu, and Tsivkin 1999)

What's more these general results held in a subsequent experiment that involved functions more complex than addition such as base 6 and base 8 addition, cube roots and logarithms in base two. When they scanned the brains of the participants they found greater bilateral parietal lobe activation for approximation than exact calculation[1]. It is important to note that most of the activated areas fall outside of traditional language areas. Scans taken while performing exact calculations, on the other hand, "revealed a large and strictly left-lateralized activation in the left inferior lobe[2]" (Dehaene, Spelke, Pinel, Stanescu, and Tsivkin 1999). This demonstrates that exact, predictable mathematical thinking is lateralized to the left whereas approximate, novel estimation is lateralized to the right.

As a matter of fact, the first kind of evidence for the left hemisphere having this property (and the first evidence for localization of function) was discovered in 1861 by Paul Broca. Broca, presents a case history and autopsy of a man named Leborgne who had been checked into a mental institution because he had lost the ability to speak (except for the word "tan," which became his nickname, and the phrase "Sacre nom de dieu!," which he only said when upset). Broca

[1] Banks of the left and right intraparietal sulci, extending anteriorly to the depth of the postcentral sulcus and laterally intro the inferior parietal lobule . . . right precuneus, left and right precentral sulci, left dorsolateral prefrontal cortex, left superior prefrontal gyrus, left cerebellum, and left and right thalami

[2] Smaller activation was also found in the left cingulated gyrus, left precuneus, right parieto-occipital sulcus, left and right angular gyri, and right middle temporal gyrus.

writes, “[w]hen he arrived at Bicêtre he had already been unable to speak for two or three months. He was then quite healthy and intelligent and differed from a normal person only in his loss of articulate language” (Broca 1965). Broca, after having autopsied Leborgne’s brain, found that at this beginning stage in his illness there was likely selective damage to the third frontal convolution in the frontal lobe in the left hemisphere; this area is more commonly known as Broca’s area and it is commonly accepted that this area is responsible for syntax and fine motor control of the jaw while speaking (for example, patients can retain the ability to sing songs; see (Wilson, Parsons, and Reutens 2006); also see (Horwitz, Amunts, Bhattacharyya, Patkin, Jeffries, Zilles, and Braun 2003)).

Above I mentioned that we should treat the left hemisphere as responsible the rules of language, for example syntax. In order to fully appreciate the kind of linguistic deficits we will be looking at, it will be useful to explicitly look at damage to a traditional language center for the sake of comparison. In order to do this, I will review, as an example, what happens when the part of the brain responsible for syntax and motor control of the jaw, Broca’s area, is damaged.

What follows is an example of a transcript between a researcher and a patient with Broca’s aphasia. It is an excerpt from Gardner 1977; he interviews a man named Mr. Ford - Mr. Ford has what we might think of as typical verbal behaviour for as person with this type of aphasia. Gardner writes,

> I asked Mr. Ford about his work before he entered the hospital. “I’m a sig . . . no . . . man . . . uh, well, . . . again.” These words were emitted slowly, and with great effort. The sounds were not clearly articulated; each syllable was uttered harshly, explosively, in a throaty voice. With practice it was possible to understand him, but at first I encountered considerable difficulty in this.
> “Let me help you,” I interjected. “You were a signal . . .”
> “A sig-nal man . . . right,” Ford completed my phrase triumphantly.
> “Were you in the Coast guard?”
> “No, er, yes, yes . . . Massachu . . . chusetts . . . Coastguard . . . years.” He raised his hands twice, indicating the number “nineteen.”
> “Oh, you were in the Coast Guard for nineteen years.”
> “Oh . . . boy . . . right . . . right,” he replied.
> “Why are you in the hospital, Mr. Ford?”
> Ford looked at me a bit strangely, as if to say, Isn’t it patently obvious? He pointed to his paralyzed arm and said, “Arm no good,” then to his mouth and said “Speech . . . can’t say . . .talk, you see.”
> “What happened to make you lose your speech?”
> “Head, fell, Jesus Christ, me no good, str, str . . . oh Jesus . . . stroke.”
> “I see. Could you tell me, Mr. Ford, what you’ve been doing in the hospital?”
> “Yes, sure. Me go, er, uh, P.T. nine o’cot, speech . . . two times .

> . . read . . . wr . . . ripe, er, rike, er, write . . . practice . . . get-ting better."
> "And have you been going home on weekends?"
> "Why, yes . . . Thursday, er, er, er, no, er, Friday . . . . Bar-ba-ra . . . wife . . . and, oh, car . . . drive . . . purnpike . . . you know . . . rest and tee-vee."
> "Are you able to understand everything on television?"
> "Oh, yes, yes . . . well . . . al-most." Ford grinned a bit. (Gardner 1976)

It is obvious that Mr. Ford's comprehension is unimpaired. Furthermore, his ability to accurately gesture and even to joke remain, presumably, as before, but his ability to easily articulate grammatical language is impaired. In fact, this is typical of Broca's aphasiacs who can still use "gesture, facial expression, body language, affect, and intonation" (Myers 1978) to communicate.

One may be tempted (as I initially was) to think of this as a loss of recursion and perhaps conclude that recursion is a part of the FLN. In fact, Broca's aphasia is commonly associated with a loss of the ability to write which further strengthens the case since recursion is necessary for writing as it is for speech.

Goodglass and Hunter (1970) compared the speech to the writing of one Broca's patient to one Wernicke's patient and found that "both patients produced longer grammatical runs in speech than in writing", "the Broca's aphasiac makes many more [errors] in writing than in speech (40 vs. 14), a difference significant beyond p=.001", and note that their study supports "the view that the formulation of written language is, at least in part, the formulation of spoken language converted to graphic form" (Goodglass and Hunter 1970). While this study is too small to draw any concrete conclusions, it is representative of the research done since their initial investigation into writing abilities of aphasiacs. This demonstrates that it is not an issue of control of the jaw that is at issue, but the recursive property of language.

In fact, Greenfield (1991) suggests that Broca's area may be a "supramodal hierarchical processor" not only responsible for recursion in language but the ability to structure information in complex ways. She reports Grossman (1980) who gave Broca's aphasiacs two different tree structures to copy using tongue depressors. One was symmetrical, the other was not. Grossman found that Broca's patients could not build a structure even when an exemplar was available if it "exhibited two or more subcomplexes vertically subordinate to a unifying structure" (Grossman 1980) (see Greenfield for a far more in-depth discussion and further evidence).

Despite this, one cannot draw the conclusion that that recursion, per se, is lost. Corballis (2007) argues that language is not the only mental faculty that exhibits recursion, pointing at ""theory of mind", mental time travel, manufacture, the concept of self and possibly even religion" as being recursive (Corballis 2007). The question is not settled. One interesting possibility that I will not explore here is that Broca's area is responsible for certain types or degrees-of-complexity of recursion (see (Fitch 2010) for example).

In the next section, I will quickly review the largely uncontested fact that the RH is responsible for *some* Language functions before moving on to look at some (more interesting) lesion studies that suggest that damage to the right hemisphere selectively impairs pragmatic functions while leaving in tact the syntactic and semantic systems at large.

The picture I will detail will be similar to that of Myers (1978) who compares left hemisphere damaged patients with right hemisphere damaged patients and notes that their linguistic abilities are very different. We have just seen an example of what damage to the left lateralized language center can be like (indeed, there are other types of aphasia that are different, but no less dramatic).

But, on the other hand, Myers notices that right hemisphere patients appear "peculiarly unconcerned about the impact of his message, insensitive to his situation, or to the environment." (Myers 1978) He references Gardner who writes,

> The more one observes the right-hemisphere patient's comprehension of questions (both literal and metaphorical) and his own spontaneous language, the clearer his deficits in linguistic competence become. He resembles a kind of language machine, a talking computer that decodes literally what is said, and gives the most immediate (but not necessarily the implicitly called for) response, a rote rejoinder, insensitive to the ideas behind the question or the implications of the questioner (Gardner 1977).

This demonstrates that strict linguistic ability and communicative ability disassociate from one another on the basis of hemisphere. I will review evidence that damage to the right hemisphere can impair communication skills, generally, but more specifically the Gricean Maxims, non-literal speech, sarcasm, indirect speech, focus, and metaphor. However, I wish to briefly touch on some evidence for correlation.

4.2 Evidence for correlation

I will not spend much time discussing the evidence for the correlation between the right hemisphere and language as it is a commonly accepted fact that the two are correlated. Beeman (2005) writes

> Hundreds of neuroimaging studies reveal strong neural activity in the LH during language tasks; but many of these studies also observe weak signal in homologous (anatomically equivalent) areas of the RH. Moreover, a growing number of studies report greater RH than LH brain activity while subjects perform higher-level language tasks, such as comprehending metaphors, getting jokes, deriving themes, and drawing inferences, generating the best endings to sentences[,] mentally repairing grammatical errors, detecting story inconsistencies, and determining narrative event sequences. This is

not to say that the RH performs all of the above functions, but it appears to contribute to them, at least in some circumstances.

One striking piece of evidence comes from Van Lancker and Cummings (1999) who report an interview of a patient named *E.C.* Norman Geschwind interviewed him several months after his entire left hemisphere was removed (due to a growing tumor). E.C. was still able to, with great difficulty, spontaneously say some individual words (e.g. *one*, *three*, *no place*) and fillers (e.g. *um*, *oh boy*), but he could not name common objects (safety pin, measuring tape, watch, clock). The most prevalent speech was swearing, featuring seven productions of "Goddamnit!", one of "God!" and one instance of "shit" in the 5 min session (Lancker and Cummings 1999). This is an unequivocal demonstration that the RH has *some* linguistic capabilities.

Getting back to the Beeman (2005) report, he suggests that one view on LH/RH interaction, the one adopted here, "suggests that RH processing is integrative, whereas LH semantic processing is predictive narrowing the scope of expectations of upcoming input" (Beeman 2005). Actually, this may be an instance where a little linguistic terminology (and associated distinctions) can help.

Beeman 2005 argues that the RH is encodes 'course *semantic* information' and writes "I propose that the evidence points to RH processing being more coarsely tuned that LH processing, because a greater spread of inputs and outputs in RH semantic areas produces more diffuse semantic activation, compared with homologous LH semantic areas" (Beeman 2005). Of the phenomenon he describes as examples of 'course semantic coding' are the following: the understanding of metaphors, "indirect (i.e. mediated) semantic relations" (Beeman 2005), the ends of jokes, unusual interpretations, drawing "connective inferences" from sentences (or cohesion in (Ellman and Tait 2000)'s terminology, or understanding the theme of a discourse or story). After adumbrating some of this evidence (examples of which, by the way, will be discussed in the next section) he writes, "Other aspects of discourse processing may depend on what linguists refer to as pragmatics rather than semantic processing . . ." (Beeman 2005) but many of those phenomena *are* what many linguists would consider to be pragmatic and not semantic.

Fine distinctions from linguistics and philosophy of language can be used to help better understand the scope of a persons communicative deficits and likewise making psychologists aware of these distinctions (and potential psychological correlates) can serve as diagnostics which can, in turn, better help them understand the nature of the human mind. Rather than hand-wave at the distinctions of another, closely related field, the fields should be working together in a cooperative effort.

4.3 Evidence for causation

Evidence from lesion studies touches on a number of capabilities that are commonly thought to be pragmatic.

4.3.1 Gricean maxims

In 1975 the philosopher H. Paul Grice proposed the Cooperative Principle which was meant to explain the expectations people have when they communicate. He also proposed four maxims and some respective submaxims as a means to elucidate this principle. Since that time they have proved to be enormously influential. They are listed below:

The Co-operative Principle: Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

1. Maxim of Quality
   (a) Make your contribution such as informative as is required (for the current purposes of the exchange).
   (b) Do not make your contribution more informative than is required.
2. Maxim of Quantity: Try to make your contribution one that is true.
   (a) Do not say what you believe to be false
   (b) Do not say that for which you lack adequate evidence.
3. Maxim of Relation: Be relevant.
4. Maxim of Manner: Be perspicuous.
   (a) Avoid obscurity of expression
   (b) Avoid ambiguity
   (c) Be brief (avoid unnecessary prolixity)
   (d) Be orderly. (Grice 1989)

These maxims are one of the cornerstones of modern research in pragmatics.

Sabbagh (1999) implicitly blends pragmatic ability and ToM when he argues, based on evidence from people with right hemisphere damage ("RHD")[3], that the understanding of someone else's Communicative Intentions (CIs) is crucial for the 'smooth flow of everyday conversation.' "Grice noted," Sabbagh writes, "that a prerequisite for identifying the needs of the conversation is the consideration of an interlocutor's CIs," (Sabbagh 1999) noting that in order to "be relevant" one must first understand what the hearer would consider to be relevant.

I would add that in order to understand what "the accepted purpose or direction of the talk exchange" is, one needs to understand the communicative intentions of the speaker, as this is usually not explicitly discussed at the beginning of a conversation before a linguistic context is established. So, one needs

[3] And converging evidence from the communicative abilities of individuals with autism.

to understand a speaker's CIs not only as the conversation takes place but also to in order to begin a conversation in the first place.

He argues that participants tailor their conversational contribution to the person or persons being spoken to and that the ability to tailor a conversation as well as the ability to recognize the tailoring efforts of others is impaired in people with RHD (Gardner 1976; Myers 1978; Gardner, Brownell, Wapner, and Michelow 2004). Sabbagh makes an interesting connection between Grice (1989) and Myers (1979). He writes that Myers (1979) "characterized RHD patients productive speech as "copious and inappropriate, confabulatory, irrelevant, literal, and occasionally bizarre"" (Sabbagh 1999). He continues, "Interestingly, this list seems to relate directly to an inability to abide by [Grice's] communicative maxims . . . concerning making one's conversational contributions relevant, brief, and truthful (Sabbagh 1999).

Moreover, Joanette et al. (1986) task RHD patients with a story narration task. They are given an eight-frame cartoon that had a "setting followed by a complication, and a final resolution[4]" representing a typical narrative structure (Joanette, Goulet, Ska, and Nespoulous 1986) they found that the narrative of RHD and controls were "not significantly different in terms of the number of words and of [units of communication]" but that the "narrative discourse of [RHD] subjects not only contained overall less verbalized information than that of [control] subjects, but moreover, this information may be described as being less complexly organized, overall" (Joanette, Goulet, Ska, and Nespoulous 1986).

They determined the 'amount of verbalized information' by subjecting the narratives of the participants (n=56) to a "propositional analysis" (see below) in which two components, *referential* and *modalizing*, were separated. In accordance with the previous literature, they define "the referential component of discourse as the one containing specific reference to persons, objects, and ideas. . . [t]he modalizing component of discourse is defined as the part of discourse which corresponds to the locutor's personal attitude toward what he is saying" (Joanette, Goulet, Ska, and Nespoulous 1986). For ease of comparison, they discarded the modalizing component and analyzed the referential component.

The propositional analysis they use is adapted from Kintsch and van Dijk (1978) who describe a text comprehension and production model that "operates at the level of assumed underlying semantic structures, which they characterize in terms of propositions," (Kintsch and vanDijk 1978) where a proposition is a predicate "and usually one or more "arguments"" (Joanette, Goulet, Ska, and Nespoulous 1986) (see Appendix for an example of a propositional analysis). They determined that, despite the fact that both groups use the same amount

[4] "First, there is a setting, in which a tired cowboy comes into a small town, gets off his horse and rests, sitting on a bench holding his horses bridle (Frames 1, 2, and 3). Second, there is a complication, in which a tricky young boy passes by with a small wooden horse, cuts off (with scissors) the horses bridle, and having replaced the horse with his wooden horse, runs away with the real horse (Frames 4,5, 6, and 7). Third, a resolution occurs, in which the awakening cowboy finds himself holding a small wooden horse; he is quite surprised to see that his horse has changed (Frame 8)" (Joanette, Goulet, Ska, and Nespoulous 1986).

of words and units of communication, the RHD population expressed fewer propositions, less 'verbalized information.'

We could think of this as evidence that these RHD patients were the same as controls with respect to Quantity and also Relevance; presumably, the fact that word frequencies were not significantly different but the amount of verbalized information was less indicates a fair amount of (needless) repetition or 'noise' either of which would obscure the narrative. What's more since the task was to describe what happens in a series of pictures and RHD patients state fewer propositions about the picture, their conversational contributions en masse are less relevant than that of controls.

Quantity is reduced in another way. The cartoons presented are very simple so there are only so many ways of summarizing what happens in the cartoon. Thus, there is an "expected version of the cowboy story on the basis of the author's narratives;" (Joanette, Goulet, Ska, and Nespoulous 1986) they call the discrete units of the expected version "core propositions." Any core proposition that was not uttered by at least 20 percent of either population was omitted. This left the authors with 32 core propositions. They found that the RHD group contained fewer core propositions (mean of the controls: 16.4 mean of RHD: 12.0). Thus, the informativity of the RHD narrative is reduced.

Hough (1990) performed a study of the ability of RHD patients to understand the main theme of a narrative. "A theme," she writes, "enables a person to extract meaning from individual sentences and integrate that meaning into the context supplied by other sentences in the narrative"(Hough 2003). Participants had either left or[5] right lateralized brain damage or were controls. She presented the participants with short (6th grade-level) stories in which the theme was organized normally (that is, the order of events in the story are written in order) or in which the theme was presented at the end of the story (the story begins *in media res*).

She ran a pilot study in which she presented neurologically normal participants (n=15) with one of the stories in question and asked them to identify the 'central theme' of the discourse. The theme that was identified by a 2/3 majority or better was taken to be the central theme. So the central theme was independently established before the experiment.

Half of the experimental questions asked subjects to verbally report what the central theme of the discourse was and the other half participants were presented with four pictures and asked to choose which one represented the central theme of the discourse. The pictures fell into four categories: "correct response; an unrelated but plausible incorrect response; an incorrect but related response which had been mentioned in the narrative; and a nonsensical response which was totally unrelated to the narrative context" (Hough 2003).

She found that "subjects with right hemisphere brain-damage performed significantly poorer when theme presentation was delayed [in media res - *DF*] compared to its normal organization. . . . [and that] [f]or the multiple choice condition, the subjects with right hemisphere brain-damage performed signifi-

[5]Exclusive reading.

cantly worse than the other two groups"(Hough 2003).

She found that "only the subjects with right hemisphere brain-damage identified significantly fewer central themes when theme presentation was delayed" (Hough 2003). She further found that, regardless of theme condition, right hemisphere brain-damaged subjects produced "significantly more embellishments and confabulations than the left hemisphere and non-brain damaged subjects . . ." (Hough 2003). This is interesting because producing embellishments and confabulations is in violation of Quality as well as Relevance, respectively. So, in addition to Sabbagh's finding that RHD can damage the ability of a person to accurately detect one's CIs, we see specific-case-evidence for maxim-specific impairments: Quantity (Joanette et al. 1986) as well and Quality and Relevance (Hough 1990).

Actually, perhaps I have been a bit hasty here; one might wonder whether confabulations count as a violation of Quality. The fundamental question is: is a false utterance a violation of Quality if the speaker believes it to be true? Grice says the following about Quality:

> Quality: Try to make your contribution one that is true 1. Do not say what you believe to be false 2. Do not say that for which you lack adequate evidence.
>
> I expect your contributions to be genuine and not spurious. If I need sugar as an ingredient in the cake you are assisting me to make, I do not expect you to hand me salt; if I need a spoon, I do not expect a trick spoon made of rubber. (Grice 1989)

There does seem to be a strong epistemological undertone to Grice's conception Quality. But what of confabulations in the form of direct contradictions? What is the epistemological state of the participants involved? Unfortunately questions at this level were not addressed in the paper. Thus, we will have to take this issue as a good motivator for future work and reserve strictly speaking judgment on whether these confabulations qualify as a violation of Quality. That being said, there is still further evidence that RH damage has an affect of Pragmatic performance (including Quality).

Surian and Siegal (2001) compared capabilities of LHD patients to RHD patients with respect to (1) ToM and (2) the ability to judge the pragmatic appropriateness of a response. The author's are explicit in that improper responses violate Gricean Maxims. To test ToM, subjects were read two stories with a ToM component, one involving a character holding a false belief and one in which there were no false beliefs. At the end of each story, the subjects were asked one of two types of question, either an Implicit Test Question or an Explicit Test Question. An example of an implicit question is, "Where will Paul look for his dog?" An example of an explicit question is, "Where will Paul first look for his dog?" In addition to having the story read to them, visual aids were present to serve as a memory aid.

In the pragmatic task, subjects were presented (one at a time) with a series (n=26) of brief, transcribed conversations. At the end of each conversation, one character in the story asks the other character in the story a question and the subject is to choose which answer is appropriate. Two options are given, one that is appropriate and one that violates on of Grice's Maxims. Below are examples:

1. First Maxim of Quantity ("Be informative")
L.: "What did you get for your birthday?"
*T.: "present."
J.: "A bike."

2. Second Maxim of Quantity ("Avoid redundant information")
L.: "What did you have for breakfast?"
T.: "I had cornflakes, and then a boiled egg and toast."
*J.: "A hard boiled egg cooked in hot water in a sauce pan."

3. First Maxim of Quality ("Be truthful")
L.: "Have you seen my dog?"
T.: "Yes, he is in the kitchen."
*J.: "Yes, he is in the clouds."

4. Maxim of Relation ("Be relevant")
L.: "What games do you know?"
T.: "I know how to play football."
*J.: "I know your name."

5. Maxim of [Manner] ("Be polite")
L.: "Do you want to play with me?"
*T.: "No, you are too stupid."
J.: "No, I am too tired." (Surian and Siegal 2001)

They found no significant difference between the implicit and explicit questions in the ToM tests but did discover a "reduced sensitivity to pragmatic violations" (Surian and Siegal 2001).

This study is an excellent example, because it explicitly tests violations of the maxims. Surian and Siegal (2001) conclude "While their sensitivity to conversational constraints was not severely impaired, they nevertheless showed a significantly poorer performance than normal controls" (Surian and Siegal 2001).

They note that Relevance, Manner, and Quality were rated as more severe than violations of either of the submaxims of Quantity and point out that "This pattern of results is consistent with Grice's (1975) suggestion that violations of certain maxims (e.g. the second Maxim of Quantity) would be less damaging to human communication than the violations of other maxims (e.g. Quality or Relevance)" (Surian and Siegal 2001). This result also complements the result of Surian et al. 1996, discussed above (section 3.2).

This is further interesting because it has been noted in the past (Potts p.c.) that Quantity violations are quite common. It is beyond the scope of the present discussion, but, taken together, these facts may suggest a re-thinking of this Maxim in terms of formulation or necessity.

Furthermore, there are many other studies (Silverman 2008; Brownell, Gardner, Prather, and Martino 1995; Gardner, Brownell, Wapner, and Michelow 2004; Myers and Bookshire 1996; Joanette and Goulet 1990; Delis, Wapner, Gardner, and Moses 1983; Brownell, Potter, Bihrle, and Gardner 1984; Brownell, Potter, Bihrle, and Gardner 1986) which have shown that RHD patients have difficulty extracting a "main theme from discourse" (Sabbagh 1999). However, space considerations preclude me from discussing each in detail. It suffices, based on the above discussion and general support in the literature, to conclude that the right hemisphere is necessary for competent use of language.

4.3.2 Non-literal speech

Non-literal speech is generally considered to be a pragmatic issue. Since the speaker's intended meaning does not come directly from the words, something else (context, world-knowledge, etc.) must be contributing. The derivation of the meaning, then, relies on the interaction between the utterance meaning and the context/environment. This is generally taken to be a task for pragmatics.

Kaplan et al. (1990) tested the ability of RHD patients (against age matched controls) to use information about the relationship between two people in a conversation to discriminate between sarcasm and a white lie. They note that it has long been recognized that the point of sarcasm is to sharpen a criticism where as a white lie is something that you say to take the edge off of a failure. However, this is the only difference between these two kinds of remarks - both are literally untrue and do not reflect the real-world situation.

They played a series of short vignettes in which the relationship between two characters is described as either friendly or hostile. In the vignettes an action or a performance takes place and that action or performance is described by one of the characters in the vignette as either good or bad. The following matrix represents the six possible relationships:

Relationship	Good Performance	Bad Performance	Comment type
Friendly	Literally True	White Lie (literally false)	Positive
Hostile	*Hostile Lie	Literally True	Negative
Friendly	*Unmotivated Lie	Literally True	Negative
Hostile	Literally true	Sarcasm (literally false)	Positive

Those that are literally true, white lies, and sarcasm speak for themselves. The two entries marked with a * are logical possibilities but, practically speaking, infelicitous comments[6] which were not portrayed in the vignettes.

[6] Actually what I have marked as hostile lie could be just as easily described as 'bitter jealousy' or something like that but since neither of the *-marked entries concern us here; I

After each vignette the participants were asked four questions. The first two had to deal with factual content. The third asked for an assessment of the speaker's pragmatic intent (they were presented with five options: telling the truth, joking or making fun of him, trying to be sarcastic, saying something wrong by mistake, telling a lie on purpose) and the fourth asked the subject to assess the impact the statement had on the actor ("Did Al's statement make Bob feel better or worse?")

Both the RHD patients and the controls understood the literally true statements and the impact it had on the actor. However, they found that "[w]hen a speaker characterizes an actor's performance with a literally false statement [e.g. sarcasm or a white lie-*DF*]. . . RHD patients exhibit difficulty using contextual information to guide interpretations of these utterances" (Kaplan, Brownell, Jacobs, and Gardner 1990). That is, RHD patients had difficulty in distinguishing between a sarcastic remark and a white lie.

They go on to speculate that the poorer performance of RHD subjects may reflect their "inability or disinclination to understand mutual knowledge (i.e. the knowledge of what another person knows) and to impute mental states (i.e. intentions, beliefs, desires) onto the characters" (Kaplan, Brownell, Jacobs, and Gardner 1990). That is, they suggest that it may be due to a mild ToM deficit. They further fortify this conclusion with the observation that RHD subjects seemed to, more often than controls, use the literal meaning of the sentence as a guiding factor in deciding whether it made the actor feel better or not (positive words cause positive feelings, negative words cause negative feelings regardless of relationship type).

The central message of this study, then, is that RHD patients have trouble integrating contextual information that affects the meaning of a sentence in context. A subtle but important question that arises from this study is: is it the inability of the patients to identify internal states (as Kaplan et al. 1990 suggest) or is their inability to integrate that knowledge with the linguistic data? Unfortunately, this issue was not addressed in the article, but it remains an interesting question for future research.

4.3.3 Sarcasm

Sarcasm is a special kind of non-literal speech because it (or at least the brand I will examine here) usually takes the form of being literally false (or the "opposite" of what is the case). For this reason it, too, is generally considered to be a pragmatic phenomenon because the meaning of the utterance is not the sum of its parts.

The above study intimated that RHD people may have difficulty understanding sarcasm due to a ToM deficit. This notion is further supported by Channon et al. (2003) who tested 19 people who had suffered a closed head injury ("CHI") and found there was a correlation between CHI and the ability to interpret sarcastic remarks as sarcastic rather than genuine. "Moreover,"

will just move on

they write "[the patient's] difficulties in processing sarcasm were related to mentalising abilities, but not social knowledge"(Channon, Pellijeff, and Rule 2003). They were able to infer that the utterance was not meant literally, however they failed to apprehend the correct non-literal meaning of the utterance. In other words, they knew the utterance was not meant literally but the individual could not recognize the CIs of the speaker.

As a brief aside, one may wonder if the patient's concept of sarcasm has somehow been damaged and it is this (and not the ability to detect the novelty of the utterance) that has been impaired. "Surprisingly little . . . is understood about how we acquire and deploy concepts," write Kumaran et al. (2009). However, they go on to argue on the basis of experimental data that it appears to be the "hippocampus alone [which supports] the efficient transfer of knowledge to a perceptually novel setting" (Kumaran, Summerfield, Hassabis, and Maguire 2009) and the ventromedial Prefrontal Cortex ("vMPFC") "mediates the online integration and evaluation of associative information conveyed by the hippocampus" (Kumaran, Summerfield, Hassabis, and Maguire 2009).

Therefore, it appears as if the hippocampus and vMPFC may serve as a 'concept generating circuit' which is presumably neurologically separate from areas where interpretation occurs. This is to say, it is not contradictory to say that a person has a concept (generated by the hippocampus) but cannot recognize it due to brain damage (as evidenced by the subjects in the Channon et al. study). End aside.

Unfortunately, due to the varied nature of the legion sites, "Inferences about localisation of damage underpinning any deficits cannot of course be drawn from the present data. Frontal lobe dysfunction is assumed to be the main mechanism underlying poor performance, since damage to the orbital and lateral frontal lobes typically occurs after injuries of this nature, regardless of the particular site of impact (Ommaya, Grubb, and Naumann, 1971), although other factors including right hemisphere damage or diffuse axonal injury could of course play a part" (Channon, Pellijeff, and Rule 2003).

While one cannot be certain, one can be somewhat confident that this impairment is due to Medial Prefrontal Cortex or Orbito Frontal Cortex damage. Since damage to the frontal cortex is common after CHI and the to Medial Prefrontal Cortex or Orbito Frontal Cortex is consistently activated in the vast majority of ToM studies (Carrington and Bailey 2009), it seems likely that this part of the brain would incur some degree of damage.

In fact, a closer examination of Ommaya et al. (1971), cited by Channon et al. (2003), reveals that there is consistently damage to the Medial Prefrontal Cortex or Orbito Frontal Cortex damage in both adults and rhesus monkeys as a result of traumatic impact. This consistency is thought to have to do with the mechanics of how the brain sits in the skull, how it moves in the skull upon impact, and how the shape of certain plates in the skull put certain regions of the brain at higher risk for incurring damage during impact (Ommaya et al. 1971).

This study (as well as Sabbagh 1999 and Kaplan 1990) naturally leads to a caveat: ToM is distributed in both hemispheres. Given the fact that many of the RHD studies I will examine are not precise as to the location of the damage (other than specifying the hemisphere) the possibility remains open that some of these RHD would better be classified as ToM studies. The Joanette et al. (1986), Hough (1990), Surian and Siegal (2001) studies gives us some degree of confidence, however, that the two capacities may be responsible for different aspects of pragmatic language given that ToM deficits and RH damage result in different behavioral profiles. In the next three sections, about prosody, indirect speech, and metaphor, the role that ToM plays in concert with the RH becomes a bit harder to suss out, but it is still worth mentioning them briefly since they are relevant and act as suggestive evidence for the notion that both ToM and the RH are responsible for pragmatic uses of language (though what the relationship is between the two abilities (i.e. what one does and what the other does specifically) is still a bit mysterious.

4.3.4 Prosody

As we saw in section 2.4 , tone of voice can make all the difference when it comes to meaning. As we will see below, tone of voice can also impart emotional meaning (which is a slightly different kind of meaning than that which I have been discussing up to this point) as well contrastive focus.

Ley and Brydean (1982) conducted a study in which they played two different sentences of a similar grammatical structure at the same time, one read neutrally, one read in an emotional tone (happy, sad, angry). Participants were instructed to focus on what they heard in one ear, disregard what they heard in the other ear and were asked "to judge both the emotional tone and the content of the sentence they heard on the designated ear" (Ley and Brydean 1982). They, then, scored the errors that people made in identifying the content of the proposition or the emotional tone of the utterance. They found that "the vast majority of individual subjects showed both a right ear advantage for the content judgments and a left ear advantage for the emotion judgments . . . This outcome indicates that independent parallel processing occurs in the two hemispheres for the preferred components (emotional or verbal) of a composite stimulus" (Ley and Brydean 1982). Since information gathered from an ear is processed by the opposite hemisphere, this also shows that the left hemisphere processes content whereas the right hemisphere processes emotional tone.

Ross et al. (1988), by performing a Wada test, concludes that the right hemisphere is responsible for effective prosody (happy, sad, angry, surprise) thus supporting Ley and Brydean's (1982) findings. A Wada test, named for Juhn Atsushi Wada (but also known as intercarotid sodium amobarbital procedure, "ISAP") is a procedure where sodium amobarbital is injected into one of the major arteries in the neck (internal carotid arteries); it travels up and into the hemisphere of the side it was injected to, temporarily shutting down the functioning of that hemisphere allowing researchers to approximate how a person would function if one hemisphere was removed.

They measured the speech of participants according to twelve different measures[7] (including several relating to fundamental frequency F0[8]) and test for "how well each patient was able to use an individual acoustical parameter for overall affective signaling across the spectrum of affects represented by neutral, happy, sad, angry, and surprise" (Ross, Edmondson, Seibert, and Homan 1988). They found momentarily disabling the right hemisphere led to "prominent loss of affective prosody" (Ross, Edmondson, Seibert, and Homan 1988) and that this loss of affective prosody was tied to the F0 measures.

This finding is important as it confirms earlier reports (Ross, Edmondson, Seibert, and Homan 1988; Shapiro and Danly 1985) that affective prosody is a property of F0 rather than intensity of timing. They write,

> It may well be that the linguistic aspects of prosody in English, e.g. syntactic disambiguation, contrastive stressing to indicate the new and important information in an utterance, word class (con'vict vs 'convict) etc., are more dependent on non-F0 acoustical parameters that are modulated by the left hemisphere and, therefore, would not be attenuated during a right-sided [Wada test]. (Ross, Edmondson, Seibert, and Homan 1988) emphasis mine)

The suggestion that contrastive focus is not a pragmatic one is not necessarily a surprising one. Many different fields have laid claim to the phenomenon of focus from Pragmatics (Kadmon 2001) to Syntax (Chomsky 1971), (Chomsky 1976), (Jackendoff 1972) to Phonology (Selkirk 2002) to Semantics (Hajiĉová, Partee, and Sgall 1998) as well as interfaces of the various sub-disciplines, integrated syntax, semantics and phonology (Bruel 2004), syntax and phonology (Zubizarreta 1998), (Winkler 1997). The present study suggests that a characterization of contrastive focus is not the job of pragmaticians.

Given the observation by Givón 2001, the fact that focus is a structure-based phenomenon, may not be terribly surprising. He writes, "the notion *contrast* . . . rests upon the more fundamental cognitive dimensions of *informational predictability* and its converse *counterexpectancy*, both of which are in principle scalar" (Givón 2001).

For example, take the following:

1. Joe lent me the bike

2. It was *Joe* who lent me a bike (rather than Mary)

3. It was a *bike* that Joe lent me (rather than a car)

He writes the "unpredictability arises from expectations established in the *local* discourse context, given in [the parenthetical material above]. That is,

[7] Fundamental Frequency (F0) slope, Total Time (log10 (msec)), Percent Pause Time, Percent Voiceless, F0 Register (semitones), F0 Variation (semitones), F0 Attack (semitones/sec), Delta F0 (semitones), dB Register (dB), dB Variation (dB), dB Attack (dB/sec), Delta dB (dB).

[8] The longest wave length (or lowest sounding part) of one's speech.

the unpredictability arises from the speakers assessment of what the hearers expectations are in the particular context."(Givón 2001)

Since Ross et al.'s (1988) result only dealt with (and is thus limited to) production, it may also be the case that repetition of material in the (familiar) common ground reduces the novelty of the utterance (save the focus material). However, as Ross et al.'s (1988) study was concerned with affective prosody and not non-affective prosody (contrastive focus syntactic disambiguation, etc.). Therefore, it is difficult to draw any definite conclusion (empirical tests designed to bring out the relevant factors would need to be constructed and executed), but the result is not a shocker.

Incidentally, this is a good example of how linguistics as currently practiced can benefit from the approach advocated here. There was a controversy between fields as to what kind of phenomenon contrastive focus is (syntactic, phonological, or pragmatic) and, based on a study which disables the 'pragmatic hemisphere' - at least on the picture presented here - we can reasonably rule out the notion that a theory of pragmatics must account for this phenomenon.

4.3.5 Indirect Speech

Indirect speech is also typically considered to be a pragmatic phenomenon because the speaker's meaning is not the same as the sentence meaning. An example of this kind of speech was discussed in section 2.4 with examples like *Can you pass the salt*?

Hirst et al. (1984) compared the ability of people with brain damage to Broca's area to people with RHD in order to determine how dependent people Broca's damage are on the context in order to parse indirect requests. What they found, to their surprise, is that the individuals with RHD were impaired in their ability to determine when a question requires a literal answer (*Can you play tennis*? in the setting of two people sitting in a living room) or action (*Can you pass the salt*? when said at a dinning table). The RHD responded 70 percent of the time that action was appropriate in the instances where the literal answer is correct[9]. That is, these patients judged it appropriate when one of the people in the video stood up and started swinging a tennis racket when asked, *Can you play tennis?*, in the context described.

One cannot draw too concrete of a conclusion based on this, due to the fact that the anatomical areas of the brain are loosely defined in the study (but all were limited to either "the right-temporal or parietal lobes" (Hirst, LeDoux, and Stein 1984). This finding has also found subsequent support in other studies ((Foldi 1987; Hirst, LeDoux, and Stein 1984; Weylman, Brownell, Roman, and Gardner 1989)).

One may be surprised by this finding given the literal mindedness of RH patients that has been reported up until now. Why do these patients in this case chose the *non*-literal interpretation? As I have stated before, further testing

[9]Conversely the Broca's aphasiacs had difficulty interpreting literal questions but had no problem interpreting requests formed as questions

would be necessary to draw conclusions one could be confident about, but a good candidate explanation is that the *Can you . . .* constructions has become (or is becoming) grammaticalized for some (if not many or most) people.

Heine and Kuteva (2002) write that "grammaticalization involves four main interrelated mechanisms

1. desemanticization (or "semantic bleaching") - loss in meaning content,
2. extension (or context generalization) - use in new context
3. decategorization - loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms, and
4. erosion (or "phonological reduction") loss in phonetic substance. (Heine and Kuteva 2002)

Certainly the *Can you . . .* construction meets the first two requrements. An interesting experiment would be to see if the high rise terminal used to indicate questions in English is present in the 'pass-the-salt' kinds of questions that are common in English. If the high rise terminal was absent, it may be evidence of the final requirement which would serve as further evidence that this construction is (becoming) grammaticalized which would, in turn then help to explain the behavior of the RH damaged individuals.

4.3.6 Metaphor

Finally, metaphor may be the quintessential non-literal speech mechanism taken by many to be a pragmatic phenomenon (which is to say the meaning of a metaphorical statement is not directly derived from the meaning of the parts of the sentence). It is, therefore, a bit surprising that the data is less clear than one may naively expect it to be. The evidence for this section is not as clear-cut as the evidence provided above, however, it warrants mention because some work has been done and it would be interesting to look at it.

An editorial in the journal *Brain and Language* asks "Is metaphor special?" In fact, they point out that the idea that right hemisphere processes metaphor is so commonplace that they describe it as "the traditional view" (Editors 2007). However, they also show that the traditional view may not necessarily be the correct one.

Studies like Mashal and Faust (2008) serve as evidence for the idea that there is a right hemisphere "advantage for the comprehension of unfamiliar metaphoric expressions . . ." (Mashal and Faust 2008). They ran two experiments. In both, word pairs were presented visually to the subject's right visual field (processed in the left hemisphere) or left visual field (processed in the right visual field). In the first experiment, the pairs were either metaphoric expressions taken from poetry (*mercy blanket*) or unrelated words (*opinion sky*). In the second experiment, the word pairs were either literal statements (*loyal friend*) or unrelated words (the same list as in the first experiment). They found, in accordance with previous studies (Ferstl, Rinck, and von Cramon 2005; Xu, Kemeny, Park, Frattali, and Braun 2005), that the right hemisphere,

> is better, i.e. more accurate and faster, than the [left hemisphere] in successfully discriminating between unfamiliar meaningful and meaningless linguistic stimuli. These consistent findings [throughout the literature that they cite] support the claim that the [right hemisphere] may be critically involved in the integration of the individual meanings of two seemingly unrelated concepts into a meaningful metaphoric expression. (Mashal and Faust 2008)

However, this finding is not without some degree of controversy. In an event-related fMRI (ER-FMRI) study, Stringaris et al. (2005) tested the abilities of patients to make sense of "metaphoric (*Some surgeons are butchers*), literal (*Some surgeons are fathers*), or non-meaningful sentences (*Some surgeons are shelves*)" (Stringaris, Medford, Giampietro, Brammer, and David 2007). They found little activation anywhere in the right hemisphere.

The editorial concludes that "metaphor per se is not unique . . . instead [the brain] is sensitive to degrees of meaning salience, remoteness of semantic relationships, open-endedness, transparency of stimuli's meanings, and speakers' intention (regardless of contextual appropriateness)" (Editors 2007). Actually both of these conclusions are in-keeping with the thesis put forward here. In the Mashal and Faust (2008) study, the RH is better at deriving novel meaning from "unfamiliar meaningful and meaningless linguistic stimuli." The Stringaris et al. (2007) finding is also in-keeping with the hypothesis promulgated here: that the right hemisphere can be productively looked upon as the 'pragmatic hemisphere' which serves as a 'novelty detector' and works with the left hemisphere to integrate context and language. What's more, based on the discussion of the trans-species property of hemispheric specialization (section 2.5), this ability (roughly) is evolutionarily very ancient as reflected by the fact that it appears to be common to animals with whom the last common ancestor lived at least between 50 and 60 million years ago (in the case of whales; (Hof and van der Gruch 2007) (cf. footnote 8, p. 21) and possibly even older.

At this time it seems that the debate on metaphor cannot be resolved. The two studies briefly discussed above both test literal, metaphoric and nonsense meaning combinations and differ importantly only in terms of (1) length of proposition and (2) methodology (3) repetition of materials. Actually, reason (3) could explain why little right hemisphere activation was found, the predictability of the majority of the stimuli (*Some surgeons are . . .*) may have resulted in less right hemisphere activation.

4.4 Conclusion

I hope to have shown a second way in which looking to the brain has the potential to help settle theoretical questions. In the previous section I argued that ToM is a pragmatic module and in this section that the RH is, generally, a pragmatic processor in so far as it it is sensitive to the novel aspects of utterances.

In section 3.4 I quoted Potts (2010). I repeat that quote here for convenience.

> Utterance interpretation involves complex interactions among (i) semantic content, (ii) the context of the utterance, and (iii) general pragmatic pressures (of which Grice's maxims are one conception). The starting point for a formal pragmatics is the observation that speakers agree to a remarkable extent on the interpretations of the utterances they hear, suggesting that there are deep regularities across speakers, utterance contexts and sentence types in how (i)-(iii) interact." (Potts 2010)

I argued that (i) is the job of semanticists (as is classically understood) and that (iii) was due to ToM processing. (ii) then seems to be due to RH language processing. As just demonstrated the RH seems to be responsible for detecting and interpreting the novel properties and subtleties of utterances.

More broadly, an aim of this thesis was to compare the pragmatic abilities that have been attributed to ToM and also to the RH. It appears that, for the most part, ToM is responsible for the ability to communicate effectively whereas the RH is responsible for novel uses of language. Both ToM and RH seemed to play some role with respect to the Gricean Maxims. It appears to be the case that the ability to follow these rules is a product of ToM processing though damage to the RH can also impair them to a lesser extent. Of course, many of the RH studies did not anatomically locate the damaged site and therefore the possibility remains that it was, in fact, ToM areas located in the RH that were damaged.

One viable hypothesis is that ToM is necessary for all uses of language i.e. literal, non-literal) whereas the RH is specifically involved in novel (non-literal) uses and that there is (as one might expect) overlap in terms of Maxim obeying/flouting; damage to either can result in similar deficits. Another is that ToM is necessary to understand what another person means with their words and the RH is responsible for integrating that information with the understanding of the words in context. While at this stage it is not possible to tease this aspect of the puzzle apart, I consider the discovery of an interesting problem to be just as important as answering an old one or discovering a new fact. Thus, this level of explanation will have to wait for further research.

Chapter 5

Some implications and a speculation

"The capacity for language is without doubt the crowning achievement of evolution, an achievement that feeds on itself to produce ever more versatile and subtle rational systems, but still it can be overlooked as an adaptation which is subject to the same conditions of environmental utility as any other behavioral talent." – Daniel C. Dennett 1981

". . . real languages appear to be organized by Kantian abstract categories." – Steven Pinker 2007

5.1 Introduction

In some respects, I have made the point I wished to make here. A quick look at the brain shows that the brain supports a rather traditional model of the semantics/pragmatics distinction where literal, compositional, or grammaticalized utterances are processed on the left and novel, non-literal utterances involve processing in the right hemisphere.

What's more I have argued that that another useful distinction the FLN/FLB distinction can be thought of, at least partially, as a distinction between left lateralized language centers such as Broca's area and Theory of Mind which is distributed in both hemispheres.

There are some implications that follow from what I sketched above. Actually, there are too many implications to address, but I would like to touch on a few of them and also, as I mentioned above, speculate a little on how this approach may yield more fine-grained theories as technology improves.

Here I will discuss the following implications: The implicit assumption of nativism, the universality of the Gricean Maxims, and the status of the possibility of a formal model of pragmatics. Furthermore, I would, by examining a single, remarkable study, like to speculate about the future of this approach by commenting on the basis for a potential theory of syntax.

5.2 Nativism?

All along I have been implicitly assuming something for which there is ground for objection. The issue comes down to modularity for which there are two explanations. Karmiloff-Smith (1992) states the case nicely, "Prespecified Modules versus a Process of Modularization" (Karmiloff-Smith 1992). That is, are the modules formed by experience or genes? The answer to this question "is an empirical one" which only future research can resolve - and the jury is still out (Karmiloff-Smith 1992). While further research is needed to resolve the dispute one cannot, on the basis of available data, help but have an opinion on the matter. I am more convinced of the prespecified modules claim than of the other one (what causes the mapping of the same kinds of data to the same brain regions in the population, if not genes?).

Of course if the modularization theory is true then data is absolutely necessary for this process of forming modules. Therefore, for our purposes, it is safe to equate this with empiricism, the idea that all knowledge is learned through experience and exposure. This is contrasted with the view tacitly assumed here and throughout this work that humans come into the world with some degree of genetically encoded knowledge. Let's equate this with nativism.

And just to be clear I do not (along with all other credible people) believe that nature explains *everything* or nurture explains *everything*. Obviously the answer is that genes are necessary to make the data relevant (through some neural instantiation). As it has been put before, humans, and not cats, will learn a language if exposed to it presumably because of some innate difference (Pinker 2002). Presumably we cannot see ultraviolet light because we lack the genes that would grant us that capacity. The real question is: what is the mix? And although it is a premature question[1], we can glimpse, however shallowly, some sense of what the answers may turn out to be. One method of insight is simply reasoning with the available data and examining the implications.

One of the largest issues that empiricists and nativists debate is the (so-called) *argument from poverty of the stimulus*. The claim is that, children will say grammatical constructions, which are a part of their language, which they have never heard before. This is not the only issue debated by empiricist and nativists (far from it!) but it is a significant one. If the poverty of the stimulus argument is true, then children could not learn language due to exposure alone (or, so it is claimed).

While the argument from poverty of the stimulus was employed by linguistic nativists for many years as a key piece of evidence that children must be born with some innate knowledge of grammar, the claim has been shown to be without empirical support. Pullum and Scholz (2002) demonstrate that all arguments to date are either unsubstantiated (lack empirical evidence) or posses technical

[1] In order to answer this question, we would need to know the totality of what genes confer on to us. Since we do not know how to read most of the genetic code, we do not know what the full contribution of genes (nature). Therefore to argue that something is nurture and not nature is premature because there could be a gene for that specific ability (though there are evolutionary-driven constraints).

flaws that invalidate their respective conclusions. As a response to the surprising fact that linguists, for nearly 20 years, took a fictional fact as support of a scientific theory, they suggest a specification schemata ("APS") that any future argument from the poverty of the stimulus will have to pass in order to be considered for validation/acceptance by the community at large.

Satisfying this schema requires one to do the following:

> ACQUIRENDUM: specify a grammatical structure that children say and to which they have not had (significant) prior exposure
> LACUNA: specify a set of all constructions which contain the acquirendum.
> INDISPENSIBILITY ARGUMENT: give reason to think that access to the lacuna is indispensable for the acquisition of the acquirendum.
> INACCESSIBILITY EVIDENCE: give evidence that the child is not (significantly) exposed to the lacuna before uttering the acquirendum.
> ACQUISITION EVIDENCE: and finally provide evidence that the child has acquired the acquirendum.

While Pullum and Scholz's (2002) review of the literature genuinely appears to be extensive, I maintain that have prematurely dismissed a case study that provides good evidence for this argument - *Simon* (Pinker 2007a), (Newport 1999).

Pinker (2007a) writes

> Simon's parents did not acquire sign language until the late ages of fifteen and sixteen, and as a result they acquired it badly. In ASL, as in many languages, one can move a phrase to the front of a sentence and mark it with a prefix or suffix (in ASL, raised eyebrows and a lifted chin) to indicate that it is the topic of a sentence. . . But Simon's parents rarely used this construction and mangled it when they did. . . Astoundingly, though Simon saw no ASL but his parents' defective version, his own signing was far better ASL than theirs. He understood sentences with moved topic phrases without difficulty . . ." (Pinker 2007a)

There, of course, remains the question: "what does *rarely* mean?" The reason this matters is because the argument from poverty of the stimulus states that the child has *never* heard the construction. However, Pullum and Scholz (2002) note a circularity in this stipulation[2] and weaken the(ir version of the) poverty of the stimulus argument to following: 'People attain knowledge of the structure of their language for which a lack of evidence that is adequate to the task is

[2]Specifically it is this: If a person claims to know a sentence or expression in their language then how can they know it is a pert of their language? If the person cites evidence for it, then the evidence undermines the claim of nativism. If, on the other hand, the person claims that the knowledge that the sentence is part of the language is innate, then nativism has been assumed in an argument for nativism.

available in the data to which they are exposed as children' (Pullum and Scholz 2002).

According to Elissa L. Newport, one of the researchers studying Simon's linguistic behavior, Simon was capable of understanding topicalization structures that "the parents consistently misinterpreted"(Newport 1999). Simon's parents only used one topicalization structure (S,VO where the underlined constituent indicates topicalization) and when asked to interpret other topicalization structures of ASL (O, SV; VO, S), Simon's mother misinterpreted 8 out of 9 trials and Simon's father misinterpreted 7 out of 9 trials; Simon misinterpreted none. Seeing as Newport (1999) presents very convincing evidence that Simon's only possible source of ASL were his parents[3], it is the case that Simon is capable of understanding[4] structures of ASL that his parents cannot. Knowing what can and cannot be topicalized demonstrates a deep knowledge of the syntactic structure of a language; Simon has, obviously, acquired that knowledge from input which was insufficient to the task (since his only sources of input do not have such knowledge). The final test is to see if I can satisfy the APS.

> ACQUIRENDUM: ASL topicalization structures other than S,VO e.g. O, SV; VO, S
> LACUNA: The set of all sentences that contain the following topicalization structures: O, SV; VO, S
> INDISPENSIBILITY ARGUMENT: If learning were data-driven, Simon would need access to the sentences in the lacuna in order to learn that O, SV; VO, S are permissible topicalization structures in the target language.
> INACCESIBILITY ARGUMENT: Simon's parents never use topicalization structures of the form O, SV; VO, S and lack the ability to reliably interpret such structures according to Newport 1999
> ACQUISITION EVIDENCE: Simon can understand non-canonical syntactic structures of ASL (Newport 1999)

Therefore, I take it to be the case that there is sufficient reason to believe that certain aspects of language are not learned from direct experience. Put another way: linguistic nativism stands. Put another way, we learn language differently than we learn other things and this ability is an evolved capability.

Anyone familiar with the literature will realize that this form of nativism is weaker, less 'miraculous' perhaps, than the ones disproven by Pullum and Scholz (2002). Actually, I only assume two things: (1) that humans are born with innate predispositions and (2) the brain has the ability to 'fill in the gaps' given enough data. In fact, both of these assumptions are not as far-fetched as one might think.

(1) is not difficult to show. The tendency to walk on two legs is a good candidate for an innate disposition. However, the most dramatic evidence I

[3]This point is conceded by Scholz and Pullum (2006). (Scholz and Pullum 2006)

[4]Due to the nature of the study Simon was in, the data presented is in terms of understanding, *competence*, rather than collecting instances of utterances, *performance*.

know of that there are innate tendencies in areas of higher cognition comes from Caramazza (2010). He has shown that imagining animate and inanimate objects activate different areas in the brain (presumably because animals are more likely to be dangerous than objects) in both sighted people and people born blind. "The finding" writes Branan (2010), "challenges the long-held notion that the two separate processing areas exist solely as the result of learning to recognize the differences in the visual appearance between living and non-living things says cognitive neuroscientist Marius Peelon of Princeton University, who was not involved in the study" (Branan 2010). Note that the long-held belief is empiricist, an example of the hypothesis that the brain becomes modularized based on experience. Brannan further quotes Caramazza, "the organization of the brain has to be understood in terms of our evolutionary history" (Ibid.).

With respect to (2) it is a well known fact that the brain 'fills in the gaps' with our visual field. Everyone is born with a blind spot cased by the lack of light-detecting rods and cones where the optic nerve passes through the optic disc. In a series of clever pictures, Ramachandran (1992) shows that the brain 'fills in the gaps' where the blind spot occurs. In one instance he takes a series of intersecting lines arranged somewhat like the spokes of a bicycle wheel and blurs the center so that the picture does not show any of the lines actually intersecting. However, when this blurry center is lined up with the blind spot of one's eye, it appears as though the blurry center is no longer there and that one can see the intersections of the lines. He writes "the visual system is able to "complete" the gaps even in these relatively complex types of patterns" (Ramachandran 1992). (See his short, accessible article for the picture just described and more so you can try it yourself.)

Therefore, the idea of innate dispositions and the ability to 'fill-in the gaps' given enough data is not without precedent in the realm of both vision and, as I have attempted to demonstrate, language.

5.3 Universality of Gricean maxims?

A fact that follows from the a defense of nativism and the totality of the discussion above is that I have implied that pragmatics is universal in the same way that language is; I have assumed a nativist perspective on pragmatics.

At least as early as (Chomsky 1971) the question of whether the Maxims are universal was raised, though to the best of my knowledge the question has received no address (but see Sperber and Wilson 1986/1995 (Sperber and Wilson 1995) for an orthogonally related discussion). The discussion above suggests that they are universal because they have (presumably genetically-based) neural correlates[5] (as evidenced by the fact that they can be selectively damaged) though it is still an open question as to how ToM and the right hemisphere (both of which are shown to bear some relation to the maxims) interact to achieve pragmatic competence. While an interesting question, it is

[5]That are reliably identifiable between people and over long periods of time.

not obvious to me at this stage whether that matter has theoretical import so I will not pursue it.

It may strike some as strange to argue that the Maxims, theoretical notions, are a cognitive reality and below I will argue as to how they are (linguistic) descriptions of labor-saving strategies which have likely evolved, but first I wish to expand a little bit on exactly what I mean when I discuss the universality of the Maxims.

Anyone who has lived in a culture markedly different from their or has even traveled to a culture markedly different from theirs will have likely noticed that people use language differently than back at home. In fact, travel is not even necessary.

Griefat and Katriel (1989) identify two folk linguistic terms *musayara* and *durgi*. Musayara, practiced by Israeli Arabs, is partially characterized by conversational effusiveness which "involves a variety of interactional tactics . . . [that] include the effusive use of many "layers" of greetings, the use of multiple, accentuated deferential or affectionate forms of address, accented displays of attentiveness, and the open sharing of personal resources, in both time and effort . . . for an address form to be "heard" as involving musayara it has to be contextually interpreted as going beyond the norm"(Griefat and Katriel 1989). They give the example of a grandfather addressing his grandson as "grandfather" (to show respect) as an example of 'going beyond the norm.'

Musayara, they note, requires "an ability to use language indirectly and artfully" (Griefat and Katriel 1989) and often times requires language that is "ambiguous, debatable, and open to various interpretations" and commonly employs "traditional sayings, proverbs, stories, and passages of poetry" (Griefat and Katriel 1989).

Contrast this manner of speaking with Durgi, the conversation style of Israeli Jews. "Speaking durgi in Israeli Sarba culture involves the choice of a "direct" strategy in performing an act . . . [characterized by] explicit, forceful, and unembellished terms (rather than "softening" one's remarks through the use of some form of indirection). It is a style that is highly confrontational in tone and intent" (Griefat and Katriel 1989).

Griefat and Katriel describe the two styles as diametrically opposed and note that "Arab communication is perceived by many Jews to involve a high degree of "fabulation" . . . and to inspire little trust. . . for Arabs, the Sarba style smacks of unfathomable literal-mindedness"" (Griefat and Katriel 1989)[6].

So who is violating which of the Maxims? Your answer to that question would probably reveal something about the language community you grew up in[7]. What I wish to argue here is that nobody is violating the Maxims. In the same way that one's linguistic environment determines the language they will

[6] "Over an above conflicts of interest and general belligerency, intercultural encounters between Arabs and Jews are all too often deflected due to conflicting cultural communication styles" (Griefat and Katriel 1989).

[7] To the American ear, Musayara may be seen as violating Quality ("grandfather" for grandson), Quantity (many layers of greetings, the use of multiple forms of address), and Relevance (the use of indirect language, the implementation of stories and poetry, etc.).

speak, one's cultural; environment determines how they will use language. To borrow some vocabulary from syntactic theory, the culture sets the "parameters" of one's Maxims (that is, what counts as in keeping with or violating a Maxim). So culture shapes the form (or "determines the parameters") of the maxims and this likely begins at birth[8].

What's more this universality (with 'surface differences') is not entirely surprising. Mameli (2007) writes "Many ethnographic studies suggest that, despite being large, cultural variation is not as unconstrained as many twentieth-century researchers claimed it to be. At a deep level of analysis, there seem to be many cross-cultural constants" (Mameli 2007). The maxims, broadly conceived seem to be an example of a variable crosscultural constant.

Above I argued that the Maxims, a theoretical description of communicative rules, are innate. Well, I owe you an explanation as to how they got there, don't I? Below, I will argue that the Maxims are particular formulations of a larger, more general principle of social cooperation. Stated roughly it is: Do what is necessary to retain group membership at as little personal cost as possible.

Below I will address three of the four Maxims; I will not address Manner. The Maxim of Manner, as a reminder, states:

> Maxim of Manner: Be perspicuous. Avoid obscurity of expression Avoid ambiguity Be brief (avoid unnecessary prolixity) Be orderly. (Grice 1989)

I will not address Manner because linguists tend to ignore it (though see (Brown and Levinson 1998) for an exception). Furthermore, the notion that Manner is necessary to retain group membership status strikes me as so obvious on its face that it need not be defended. This leaves us with Quality, Quantity and Relevance which I will address in turn.

5.3.1 Quality

> Maxim of Quality: Try to make your contribution one that is true. Do not say what you believe to be false. Do not say that for which you lack adequate evidence.

Cosmides (1985) argues that humans have evolved "Darwinian algorithms" that are "keyed to focus attention on those dimensions of a situation that are evolutionarily important, and operate on them with inferential procedures that embody an appropriate evolutionary strategy[9]" (Cosmides 1985). Since humans are tremendously social animals, many of the situations in question will be

[8] A warning to the reader to not take the comparison between a linguistic parameter and the cultural parameters too seriously. I use the analogy only for didactic reasons the culturally determined parameter of a Maxim is going to be far more vague than the sharply set linguistic ones.

[9] It should be noted that using this task as an argument for cheater-detection has been called into question by Stenning and van Lambalgen (2008) (esp. see chapter 6) with respect to ASD people.

social ones. Therefore, many of the algorithms in question will be of a social nature. One in particular that Cosmides argues for (on the basis of experimental evidence) is a 'cheater-detection' algorithm.

Cosmides (1985) found "that adult subjects are very adept at detecting potential "cheaters" on a social contract, even when it is unfamiliar and culturally alien, [and this skill] stands in marked contrast to the repeated finding that they are not skilled at detecting the potential invalidity of descriptive rules, familiar or unfamiliar" (Cosmides 1985). It seems as though the ability to detect cheaters is a specialized ability.

Let's generalize a bit and restate the Maxim of Quality (without doing violence to its original spirit). Let's say its essence is captured in the following: Do not lie; do not run the risk of lying. If we can take lying as a form of cheating then it is easy to see how we can derive Quality.

Humans have evolved a cheater-detection mechanism which presumably allows us to detect lying, a form of or tactic used in, cheating during social interactions. Since, as stated in the introduction, language is a social phenomenon (emerging from the brains of its users) this mechanism would be at work in the domain of language. The employment of general cheater-detection in the realm of language is what Grice was on to when he identified Quality as a conversational Maxim.

If we spelled out the rationale in evolutionary terms we might say: don't lie because you are likely to get caught. Getting caught in a lie damages your trustworthiness which weakens your social ties to other members of your tribe. Since humans have evolved to be social creatures group membership is important, if not essential, to our survival. Actually, recasting this particular maxim in this way also sheds light on an open problem - the acceptability of bullshit.

What is bullshit? Fania Pascal, after having her tonsils removed, told her friend "I feel just like a dog that has been run over." Is that bullshit? Her friend, Ludwig Wittgenstein, seemed to think so; he replied with disgust: "You don't know what a dog that has been run over feels like." Frankfurt (2005) makes the point that Pascal's remark, while not true, is far from lying. He identifies "the essence of bullshit" as a "lack of connection to a concern with truth - [an] indifference to how things really are" (Frankfurt 2005). He identifies several features of bullshit, namely two of which are the fact that the intent of the bullshitter is not to deceive, the bullshitter "does not limit himself to inserting a certain falsehood at a specific point, and thus he is not constrained by the truths surrounding that point or intersecting it"(Frankfurt 2005), and that bullshit is phony, but not false.

He leaves an exercise to the reader:"The problem of understanding why our attitude toward bullshit is generally more benign than our attitude toward lying is an important one, which I shall leave for the reader"(Frankfurt 2005). It is particularly interesting in the light of the fact that, as Frankfurt argues, "the bullshitter is a greater enemy of the truth than liars are"(Frankfurt 2005). Under my view, the acceptability of bullshit (which is primarily a verbal phenomenon) relies, as Frankfurt points out, on our ability to determine the intent of the (potential) bullshitter (Theory of Mind) and whether they are trying to manipulate us into doing or believing something for dubious purposes (cheater-detection). Everyone is 'on the same page' so to speak. Bullshit doesn't set off our alarm bells in the same way that a lie does because, as Frankfurt notices, there is no intent to deceive.

Stone et al. (2002) identified the neural correlate of the cheater-detection mechanism, the limbic system, bilaterally. They presented two kinds of Wason-type reasoning tasks to a man who had suffered brain damage. One test involved social contracts (if you take benefit, then you must satisfy requirement) and the other did not ('precaution problems;' if you engage in hazardous activity then you must take the precaution). They found that the patient "made errors on significantly more of the social contract problems than the precaution problems" (Stone, Cosmides, Tooby, Kroll, and Knight 2002) and conclude that social reasoning, in particular the ability to detect cheaters, is an ability supported by the limbic system bilaterally.

Recall in section 3.3 where I discussed the importance of cingulated cortex to the language system but was unable to draw any conclusion due to our poor understanding of its function. Despite that poor understanding, it is a well-known fact that this area is part of the limbic system. I have argued that the cheater-detection system plays a role in the linguistic system by equivocating the Gricean notion of truth[10] with a concept of honesty. The role (though not totally well-understood) that this area plays in both the language system and the cheater detection system is suggestive evidence that these two systems are linked via the cingulated cortex.

So, there is nothing particularly special about Quality; it falls out from a general rule of thumb which calls for transparency in social interaction: don't lie.

5.3.2 Quantity and Relation

Potts (2006) writes, "[Grice's] Maxims are now widely regarded as redundant in some places and objectionably vague in others" (Potts 2006). That criticism is certainly apt for these two particular Maxims.

1. Maxim of Quantity:
   (a) Make your contribution such as informative as is required (for the current purposes of the exchange).
   (b) Do not make your contribution more informative than is required.
2. Maxim of Relation: Be relevant.

(Martinich 1980) has called relevance "superfluous" given the maxim of Quantity - his argument being that the Maxim of Relation has the same effect

[10]While Grice certainly had a different concept of truth in mind than the one I am using here, I, in principle, see no harm in bending his original intentions for this purpose by making this equivocation.

as the second submaxim of Quantity. Following this observation he proposes to enhance Grice's formulation of Relevance with the following:

1. Maxim of Relation: Be relevant.
    (a) Make your contribution one that moves the discourse towards its goal.
    (b) Express yourself in terms that will allow your hearer to tie your contribution into the conversational context.

Undoubtedly, this is not the whole story with respect to Relevance, but I do think that this more precise treatment is preferable to Grice's (even he seemed unhappy and 'exercised' by it). If we adopt this formulation, then the need for the second submaxim of Quantity becomes so minimal than we can safely drop it.

Thus we are left with the three submaxims:

1. Make your contribution such as informative as is required (for the current purposes of the exchange).
2. Make your contribution one that moves the discourse towards its goal.
3. Express yourself in terms that will allow your hearer to tie your contribution into the conversational context.

What is interesting in these maxims is the assumed role of Theory of Mind. How do we know how informative the statement must be? Obviously we must be inferring something about the mental state of the hearer in order to make this judgment. We can further collapse these three maxims into one, slightly more general, maxim: *Make your contribution as minimal as you can while satisfying what you understand the conversational needs of your partner to be.* This brings the implicit Theory of Mind component (in the form of an epistemological statement) a bit more to the foreground which I believe is appropriate given the above discussion. In the next section I will argue that the Maxims fall out from a more general energy-conservation strategy that out ancestors may have adopted.

5.3.3 The evolved maxims

I have distilled the following out of Grice's (reformulated) Maxims:

1. Do not lie; do not run the risk of lying.
2. Make your contribution as minimal as you can while satisfying what you understand the conversational needs of your partner to be.

I argued that truthfulness would have been selected for because, following Cosmides, those who could better "calculate whether the cost/benefit structure of [the desires of another person] . . . [can better] operate on the cost/benefit structure of the situation with inference procedures that define cheating and facilitate the detection of cheaters" (Cosmides 1985). Another way of stating this is that an individual was more likely to survive (and thus reproduce) if they could cooperate efficiently with other cooperators and could identify (and avoid) free-loaders.

The second rule requires a bit more discussion. It has been observed that Grice's maxims have a sort of upper and lower bound (Martuinich 1980, but esp. Horn 1996). In fact Horn 1996 reformulates Grice's Maxims (sans Quality) in the following two rules:

1. The **Q** Principle ("lower-bounding hearer-based guarantee of the sufficiency of informative content"):
   (a) "Say as much as you can, modulo **R**"
2. The **R** Principle ("an upper-bounding correlate of the Law of Least Effort dictating minimization of form"):
   (a) "Say no more than you must, modulo **Q**" (Horn 1996)

This reformulation is essentially the same as mine except mine places a bit more emphasis on Theory of Mind and is perhaps a bit more naturalistic (and perhaps a bit less precise). Either of these reformulations of the maxims makes perfect sense in evolutionary terms. Robin Dunbar describes evolutionary theory as "a balance between costs and benefits" (Dunbar 1998). Here the cost is calorie-burning through the use of energy (perhaps often not for one's own immediate gain), but the benefit is the ability to obtain calories easier, protection, and greater ease in obtaining a mate.

The cost to the individual is making the contribution at all (or the **Q** Principle in Horn's terms). While it is not entirely clear whether or how much increased cognition increases metabolic activity (see (He, Snyder, Zempel, Smyth, and Raichle 2008); (Sapolsky 2010)), it is clear that speaking consumes energy. Speaking increases respiration, metabolic activity involved with internal and external intercostal muscular contraction, gesture which is associated with speech to varying degrees in people (Merryman and Bronson 2009), as well as the laryngeal and oral muscle manipulation associated with articulation. (Even if the speech of our ancestors was a signed and not spoken one, the argument still holds, just in a different form.)

It is likely that the second part, satisfying what you understand the conversational needs of your partner to be (The **R** Principle), is a product of the central role that groups have had in human evolution. "The massive and complex information-crunching capacities of the brain were designed to help our ancestors make functional decisions in an environment that included other people as a prominent feature,"(Shaller, Park, and Kenrick 2007) write Schaller et al. (2007). That group membership is important is further evidenced by the role that touch and social contact has not only on social and emotional development (Harlow 1958) but also on overall physical development (Spitz 1945). Humans are so profoundly social that it has been hypothesized (Dunbar 1998) that the need to maintain social relationships in relatively large group sizes was the environmental condition that selected for our large brains as well as the evolution of language (Aiello and Dunbar 1993). The idea here is that by communicating, bonding, etc. you work to ensure membership in a group.

Our ancestors had a fine line to walk. On the one hand, they wanted to conserve energy which requires conserving movement (and possibly language-related cognition) which burns calories. On the other had they had social contracts that required them to perform activities that required effort. It was important for them to fulfill social contracts because it is easier and safer to live in a group than it is alone and in order to be a member of a group, one has to perform activities that ultimately benefit the group at some expense to the individual.

People who could strike this balance would have been more likely to pass on their genes than people who cooperated too much (putting themselves at greater risk for starvation, dehydration, etc.) or people who did not cooperate at all (who would not have maintained group membership for very long). Therefore natural selection chose the qualities that exhibit themselves in the verbal domain as (some version of) the Gricean Maxims.

5.4 A pragmatic module?

By attempting to single out certain brain regions that appear to be associated with pragmatic meaning, am I suggesting that these brain regions represent some kind of pragmatic module? Yes, and I am not the first to do so (though to the best of my knowledge I am the first to suggest specific brain regions).

Sperber and Wilson (2002) argue "that in hominid evolution there has been a continuous pressure towards greater cognitive efficiency" and that "[t]his pressure has affected both the general organization of the mind/brain and each of its components involved in perception, memory and inference" (Sperber and Wilson 2002).

They also argue that "[v]erbal comprehension presents special challenges, and exhibits certain regularities, not found in other domains. It therefore lends itself to the development of a dedicated comprehension module with its own particular principles and mechanisms" (Sperber and Wilson 2002). They speculate that such a "module might have evolved as a specialisation of a more general mind-reading module . . ." (Ibid).

Based on the discussion of the neurological data presented above, they may be right. There is now evidence that the pSTG should be a region of interest for future empirical work.

Actually the idea that there is a pragmatic module is not completely unproblematic, for example, it runs contrary to the Aristotelian conception of metaphor[11] (which was upheld by Grice and Searle). The Aristotelian model is a three stage model of metaphor comprehension:

1. Derive the literal meaning of the utterance.
2. Assess that meaning against the context of the utterance.
3. If the literal meaning does not make sense in context, seek an alternative meaning that does. (Glucksburg 2004)

However, Gluckburg (2003) points out that this model makes empirically testable claims. One is that literal meanings are always processed first which means that "literally intended language should be easier to understand than non-literal" and that "non-literal interpretations should be optional; they should be generated only when literal interpretations are defective in one way or another" (Glucksburg 2003).

If figurative language is processed with the same ease that literal language is then this also further suggests some kind of pragmatic module by de-prioritizing the literal.

Glucksburg (2004) showed experimental participants

> Four different kinds of sentences: literally true (e.g. *some birds are robins*); literally false (e.g. *some birds are apples*); metaphors (e.g. *some jobs are jails*, *some flutes are birds*); and scrambled metaphors (e.g. *some jobs are birds*, *some flutes are jails*). The metaphors were literally false category membership assertions, but they were readily interpretable if taken non-literally. The scrambled metaphors were also literally false, but not readily interpretable.

The Aristotelian model suggests that "metaphors should take no longer to be judged literally false than the scrambled metaphors" and that "the metaphor sentences should take longer to judge as false than their scrambled counterparts" (Glucksburg 2004).

He found, in fact, that it took longer for participants to reject the metaphor sentences (mean time 1239 msec) than either the literally false sentences (mean time 1185 msec) or the scrambled metaphors (mean time 1162 msec). This suggests that there is no primacy to literal meaning over figurative ones and implicates the presence of a "stimulus driven" (Glucksburg 2004) pragmatic module analogous to the syntactic one and semantic one. Again, the neurological data discussed above suggests that this is, in fact, the case. It is of further value because it identifies regions of interest for further goal-directed empirical work to shore up the presence (or not) of such a module.

[11] Note, that because the data at time of writing are not clear, it is still not clear that metaphor should be considered a pragmatic phenomenon (as shown in section 4.3.6) so I am relying on the historical interpretation of the phenomenon as well as Glucksburg 2003, 2004.

5.5 A formal model of pragmatics?

Another implication (though I will argue one that does not pan out) is that pragmatics, with it's reliance of the right hemisphere (the one that does not deal with structure), cannot be formally modeled. I have focused on the degree of novelty in the pragmatic use of language and on ToM (which has no formal model). The novel, or creative, use of language has been written about many times by many authors previously and it has been remarked many times that practically every sentence spoken is being spoken for the first time (all things considered). It has also been observed countless times in undergraduate classrooms everywhere that any utterance can be felicitous in the right context. Thus, the machinery needed for a formal theory of pragmatics needs to be not only very powerful but extremely flexible. However, as Pinker (1984) points out, "Mechanisms that are too powerful and flexible do not allow one to explain why the data come out the way they do, rather than in countless conceivable alternative ways" (Pinker 1984) and are therefore of little value.

So, can a formal model of pragmatics ever come to fruition? I think so. In fact, there is a nascent theory already on the books that is perfectly compatible with the discussion above, Inquisitive Pragmatics.

The feature I will focus on (but see Groenendijk and Roelofsen 2009 for the full story) is the notion of common ground. In Inquisitive Pragmatics, the common ground is not "(roughly) the set of propositions whose truth is taken for granted as part of the background of the conversation" (Kadmon 2001) but rather is conceived of as a body of knowledge that needs to be enhanced in order to achieve certain goals.

In Inquisitive Pragmatics, the common ground consists of an external common ground, that which is known to everyone and is communicated through language, an internal common ground that "contains common information of which the participants are *not aware* that they have it in common"[12] and the individual state which is knowledge that is known only to the individual and is not shared (Groenendijk and Roelofsen 2009), (emphasis original)

[12]They go on to write "Such 'unconscious' common information is rather useless in that, for instance, it cannot form a basis for coordinated action." I will argue below that this is not necessarily true.

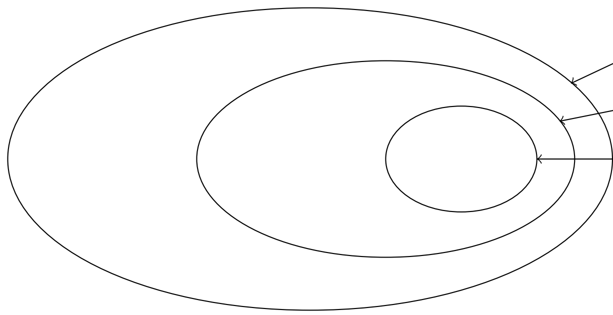

Figure 5.1: An individual information state, the internal, and the external common ground.

Pinker (2008) writes,

> Imagine Harry says, "Would you like to come up and see my etchings?" and Sally demurs. There is little or no uncertainty about Harry's intent, and none about Sally's: Sally knows that she has turned down an overture, and Harry knows that she has turned down an overture. However, Sally doesn't necessarily know that Harry knows; she might think to herself, "Maybe Harry thinks I'm naive." In addition, Harry doesn't necessarily know that Sally knows that he knows; he might think to himself, "Maybe Sally thinks I'm obtuse." Although there is individual knowledge, there is no common knowledge, and they can maintain the fiction of friendship. (Pinker, Nowak, and Lee 2009)

The external common ground is represented in the verbal information "Would you like to come up and see my etchings?," the internal common ground is the level that Pinker describes as both of them knowing that Sally has refused Harry's sexual advance. How that information is exchanged is not discussed, but here I argue that it is achieved via ToM. The individual state consists of Sally's thinking that Harry thinks she is naïve and Harry's thinking that sally is naïve. That is private information, the individual state.

The summary above is not quite in-keeping with the tenets of Inquisitive Pragmatics but I argue here that a better conception of the internal common ground may be the information that is non-verbally exchanged between communicators via ToM processes. Furthermore, that it is not "useless" but bears a relation in structuring or affecting the information that goes into the external common ground after this information has entered the internal common ground. After all, it would be a faux pas for Harry to press Sally, "Are you sure, I am quite a good etcher and I just finished one that I am particularly proud of." By examining the interplay of information exchanged at these two levels we can begin to formulate and make predictions about what will happen at the external common ground based (at least in part) on information contained in the internal common ground.

It is worth, briefly, stressing the importance of this fact. Theories of pragmatics typically focus on dimensions such as sentence meaning and speaker meaning by focusing on the interplay of language and (the usually ill-defined notion of) context. The model advocated for here changes that conception (or, if you like, partially defines the concept of 'context'). It adds a third dimension, by illuminating a second channel on which information is conveyed, information which can potentially affect the meaning of an utterance. So, whereas before there was a one channel (language) model where linguistic information was enriched by context now we have a two channel (language and ToM model) where the information in one channel (ToM) enriches or can enrich the information in the other (language)[13]. Separating these two forms of communication allows us to more closely examine what is happening on each channel, and how the channels interact to create the effect on meaning that they do. Simply put, it gives us a new lens with which to appreciate the dynamic nature of speaker meaning with respect to a given utterance-in-context.

5.6 A model of syntax? (speculation)

I would like to end this section not with a defense of another implication but rather with a bit of speculation. In the introduction I said that there is likely no theory of syntax to be found yet in this research method but there was something interesting to be said which would reopen a version of an old debate; the debate is about the context-freeness of language. Pullum and Gazdar (1982) ask, "If human languages do not have to be [Context Free Languages -*DF*], why do so many (most?) of them come so close to having the property of context-freeness" (Pullum and Gazdar 1982)? In the same article they famously disprove all prior claims of the non-context-free nature of natural languages. Since then, evidence has come to light of the non-context free nature of human languages, most famously from Swiss German (Huybregts 1976; Shieber 1985). Thus the debate is settled. Listed below are the languages for which non-context-free claims have been made:

Language	Author and Publication	Nature of Data not modelable by CFG
Swiss German	(Huybregts 1976; Shieber 1985)	cross-serial dependencies (case based)
Tagolog	(Machlachlan and Rambow 2002)	cross-serial dependencies (case based)
Persian	(Dehdari 2006)	cross-serial dependencies (case based)
Bambara	(Culy 1985)	Lexical Rule *NP-o-NP*
Dutch	(Huybregts 1976)	Semantic cross-serial dependencies

Only five[14]. Evidence from Swiss-German (cross-serial case dependencies),

[13]Or, again, if you like, you can take the information relayed through ToM processes as part of the context. I find it more useful to view it as a separate information stream and to not relegate it to an idea as diffuse and mysterious as 'context.'

[14]Note that I am omitting Higginbotham 1984 which claims that English is not context

in particular[15], led people to draw the conclusion that, not only was natural language not context-free, but natural language syntax was not context-free. It is this second assumption that I wish to pick at a little. Is it really the case that the syntactic component of Universal Grammar ("UG") is not context free, but this fact is only evidenced by a subset of grammatical features of approximately 0.0005 percent of the world's languages?

I want to suggest that there is a base component of syntax and that this base component can be represented as a context free phrase structure grammar. Of the above languages, I do not need to worry about Dutch (because it is not a syntactic issue) or Bambara ("There is evidence that the Noun-o-Noun construction belongs in the vocabulary rather than in the syntax" (Culy 1985).) which means there are only three trouble-makers remaining. However, notice that all three trouble makers are trouble makers for the same reason: case.

Sahin et al. (2009) show that there is sequential processing of inflected verbs in Broca's area. In an ICE study (recall that in ICE studies electrodes are put in direct contact with the brain and individual neurons are measure over long periods of time, days) they found that "Broca's area revealed distinct neuronal activity for lexical (200 milliseconds), grammatical (320 miliseconds), and phonological (450 milliseconds) processing identically for nouns and verbs, in a region activated in the same patients and task in functional magnetic resonance imaging" (Sahin, Pinker, Cash, Schomer, and Halgren 2009). That is the word root is processed before any inflection.

It would need to be tested experimentally on a speaker of a language with productive case (ideally Swiss-German), but if case inflection is processed in the same way as verbs and nouns, then it could be that pre-inflected syntax is context-free. Note that treating syntax as a machine with these distinct phases may also offer solutions to other inflection-related problems such as quirky case (Zaenen, Mailing and Thráinsson 1985) and may also yield insights into tense systems for both syntax and semantics. I will conclude my review of some of the implications of this work here.

5.7 Conclusion

This work supports a picture that is beginning to emerge in both the linguistics literature and the psychology literature. The days of the field of pragmatics being considered a second-class citizen or inferior or less serious, in some way, to syntax or semantics are coming to an end. Considering the fact that language-in-context - in the context of interpersonal communication - is the only known environmental stimulus needed for the acquisition of language, and pragmatics is the study of just that, it may strike some, in retrospect, as odd that so few

free because this claim is highly contested (esp. cf . (Higginbotham 1984):fn. 1). Also cf. (Pullum 1985), (Higginbotham 1985), (Pullum 1991). Cf. (Pullum 1991) for a very nice, brief, informal overview of the literature with respect to this debate.

[15] This may strike you as strange since Huybregts discovered the same fact 9 years prior. It was sort of an accident of history that this publication was overlooked at the time. See (Pullum 1991) for discussion.

linguists work in the field. This oddity may be compounded given the fact that there are very few *complete* theories ((Pinker 1984), (Tomasello 2003)) about how children learn language and that question is one of the driving forces behind much research in the field.

The picture presented here is one that treats the ability to learn language as an evolved human universal. Evidence was presented that language is learned in a way that other skills (such as reading or playing the piano) are not. Moreover, the type of evidence that supports this claim is similar to evidence that can be obtained through other cognitive domains ('filling in the gaps'). Furthermore, I continued to argue that the constraints that govern how we use language to communicate fall out from a balance which is stricken from two competing, but equally necessary, survival strategies: maintaining group membership and maximizing energy conservation.

Based on the neurological data that was examined, evidence was uncovered that identifies the pSTG as a region of interest as a potential pragmatic module, a module that has been hypothesized to exist previously in the literature. What's more, I rebuffed the implication that a Formal Pragmatic model cannot be constructed, due to the presence of RH processing. I did this by identifying a model that fits quite nicely with the data presented here.

Lastly, I speculated that the general approach proposed here has the potential to derive from the data more fine-grained theories of linguistics, such as a model of syntax. This is important because it would allow researchers to do more than just resolve interface issues which is most of what is possible given the current state of technology (though clearly, theoretically important concepts can be derived).

Chapter 6

Conclusions

“There is no reason why a psychological theory can’t be expressed in mathematical terms . . .[and] the fact that some logician is not interested in psychology does not preclude the possibility that he or she may develop a theory which can be taken as a serious candidate for a psychological theory.” – Barbara H. Partee 1978

6.1 Summing up

Here I have proposed a new investigatory approach and while not every detail is worked out, I have demonstrated that it is a viable means for collecting data with which to construct or inform a linguistic theory and that it has the potential to become an even more powerful methodology as the relevant technology becomes more sophisticated.

Specifically, I have presented a review of the literature which strongly suggests that ToM is a cognitive capacity that is intimately related to the Language Faculty and that is part of the FLB. I have attempted to identify regions of interest for future study and in so doing have added further evidence to the theory that there may be a pragmatic module analogous to the ones for syntax and semantics.

I have argued that, due to evolved hemispheric specializations, the right hemisphere should be responsible for (at least some) pragmatic uses of language. That claim was fortified with data that showed that right hemisphere damage resulted in damage to communicative competence (the Gricean Maxims), the ability to interpret non-literal speech (including specifically sarcasm, indirect speech and metaphor), as well as prosody.

I have grounded my explanations, at every stage, in evolutionary terms and used evolution to explain why some capabilities (language-learning) and practices (Gricean Maxims) are innate.

I have reflected on some of the implications of the idea put forth here and while nothing presented here is uncontroversial, there exists compelling evidence not only for what is directly argued for but also that the implications of what is presented here tie in with other aspects of the literature which complement this analysis.

6.2 Benefits of discussion

What are the immediate benefits of this discussion? There are five that are obvious to me. First, it has helped to illuminate a two-channel model of communication where linguistic and non-linguistic (ToM) communication happen concurrently. This could potentially have an effect on how (much) information needs to be accounted for in a pragmatic model. If one wanted to restrict a model of pragmatics to deal only with the exchange of linguistic data and not ToM data, then certain types of implicatures would not have to be accounted for (perhaps this is a negative aspect of what is advocated here, ToM becomes the wastebasket of linguistic-pragmatics). For example take the following:

1. Mary is running fast.
2. Oh quite well, I think; he likes his colleagues, and he hasn't been to prison yet. (Grice 1989) (said by B in response to A's inquiry as to how C likes his new job at the bank)

The implicature in (1), that Mary is running, is obviously a linguistic implicature (goes into the external common ground) because it is a literal implication; if one is running fast then they must be running. But of (2) one cannot be so sure. Grice (1989) writes, "At this point, A might well inquire what B was implying, what he was suggesting, or even what he meant by saying that C had not yet been to prison. The answer might be any one of such things as that C is the sort of person likely to yield to the temptation provided by his occupation, that C's colleagues are really very unpleasant and treacherous people, and so forth" (Grice 1989). This is a ToM implicature.

Whereas the linguistic implicatures are limited and unambiguous (they sort of follow from the facts of the sentence), ToM implicatures lend themselves to several different interpretations. Because the implicature follows less from the data and more from inferential processing in which (this is likely the case, anyway) a key word from the indirect speech act which gives rise to the implicature (in this case prison) and you compare that with the other linguistic data in the common ground as well as with your knowledge about the person to derive what you believe the person is implicating. But, crucially, the interpretation of the implicature involves an invocation of, in Groenendijk and Roelofsen's (2009) terms, the internal common ground.

Second, as just demonstrated, taking ToM into account helps us to carve up our data in new ways. As has been noted before (Potts p.c.), language is ubiquitous - therefore a model of pragmatics has to make decisions about what data it takes into account and how. Remember Pinker's observation that a model that is too powerful or flexible tells one nothing. By changing the notion of the common ground from a static one (Stalnaker 1978) to a dynamic one

(Groenendijk and Roelofsen 2009) information exchange is better represented as a complex linguistic interaction between agents. What I propose here is that common epistemic states are also shared, they are not "useless," and that there is place for them in the model of Inquisitive Pragmatics which can capture two different ways in which information behaves in a talk exchange. First, it can shed light on how the internal common ground can effect the external common ground (e.g. direct or indirect speech acts). Furthermore it can help to model how information in the internal common ground effects what information ends up in the external common ground with respect to individual agents (avoidance of faux pas, etc. become constraining factors for an individual's contribution).

Third, is that this work has offered a unique comparison between different sources of pragmatic language. Though the relationship between ToM, the RH, and The Gricean Maxims is somewhat unclear (though ToM does appear to be the stronger candidate if it is a one-or-the-other situation), I hope to have demonstrated that in-roads can be made to make distinctions between different types of pragmatic language by highlighting the advantages of analyzing how the utterance in question is novel.

Does the meaning of the utterance trade on it's novelty in the situation or does it invite the speaker to draw conclusions from information gaps that are in the digital linguistic data? The former is likely a RH process and the latter is likely a ToM process. We can then think of these different kinds of pragmatic data as existing at different levels (e.g. common grounds) in a pragmatic model. And we can do so for empirical reasons.

Fourth, this work has identified regions of interest (particularly the pSTG) for future investigation as a potential pragmatic module that has been hypothesized to exist in the literature. This implies that the field of pragmatics needs to be re-characterized from a second-class status (in some minds) to having the same import as syntax and semantics in the discipline-at-large. Furthermore, this carries with it the implication that there is no primacy of the literal; at least some non-literal utterances are not 'defective' utterances but are just as felicitous as literal statements. (Consider, without a context, the English sentence *Somebody just kicked the bucket.* Though testing would need to be done, I predict that more people would automatically process this as the equivalent of *Somebody just died* rather than providing information about the past actions of a living person or various other literally true interpretations.)

Fifth, it reformulates the notion of pragmatics from language from (1) to (2)

1. Pragmatics = language + context
2. Pragmatics= language +Theory of Mind+context

Chomsky (1980) writes, "The study of language must be concerned with the place of language in a system of structures embodying pragmatic competence, as well as structures that relate to matters of fact and belief." I have attempted to address the relationship of the language faculty to ToM in order to address the notion of pragmatic competence. Linguistic communication cannot, it seems, happen fluently when either the LF is damaged (as shown by

Broca's aphasiacs) nor with the ToM module/right hemisphere is damaged (as argued throughout). Though questions remain regarding the precise contributions of ToM and the right hemisphere (cite some) it appears as though there is some intimate connection between ToM and the Language Faculty.

6.3 Benefits of approach

In terms of the benefit of the approach advocated here, I can see at least eight advantages. The first is that it is a step towards the original conception of the cognitive revolution. Chomsky writes of the cognitive revolution, it "introduced an important change of perspective, from the study of the behavior and its products to the inner mechanisms that enter into how behavior is interpreted, understood, and produced. Behavior and its products, such as a corpus of utterances, are no longer the object of inquiry, but are just data of interest insofar as they provide evidence for what really concerns us, the inner mechanisms of mind and they ways they form and manipulate representations and use them in executing actions and interpreting experience" (Chomsky 1997).

In other words, study the behavior insofar as it sheds light on what is happening internally. One would think with the advent of brain imaging techniques that people would have begun to try to use that data to 'burn the other end of the candle' so to speak in order to begin seeing how behavior-driven theories match up with the internal workings of the brain - yet that has not happened. Most work in psycholinguistics is targeted towards specific problems and the body of data is never considered at a global level (unless it is in relation to one of those specific issues - such as a meta-study).

It should also be noted here, despite Chomsky's claim[1], that corpus studies are also a valuable methodology towards this end. Just one example of a surprising (and perhaps counter-intuitive) finding is that "about 37% of word-tokens are nouns" (Hudson 1994). Since one of the goals of linguistics is answer the question of how children acquire language knowing what data is in their environment is crucial knowledge towards that goal. Since one cannot[2] intuit what that data looks like objectively (any more than a fish can objectively describe water), corpus studies are also of vital importance.

This work is an attempt to take up the other side and see how global linguistic theory matches up with known, evolutionary-derived properties of the brain in combination with data obtained my measuring how the brain behaves when processing certain data. By comparing this to the existing work in linguistics we can begin to see how theory matches brain-function. The good news is that,

[1]He likely means the behavioristic study of corpora here rather than modern corporeal studies since he is discussing historical events.

[2]The quote from Hudson is actually the title of the paper in question. He begins it thusly, "The title of this note is a generalization that may strike the reader as a joke. . . . I am reluctant as any linguist could be to believe it; after all, the choice of word-classes in a text depends on a myriad of variable influences, from the message conveyed to the style of the author."

at least based on this, admitted preliminary, investigation, the two match up pretty well so far.

Second, this approach helps to move us away from the questionable role of intuitions (Dennett 1996; Chabris and Simons 2010). The proposal made here is to add in a new type of evidence, brain-imaging evidence and to, over the course of time, swap it out bit by bit with evidence from intuitions (a kind of weaning process). I don't suggest a weaning process out of some conciliatory nature (though that would be nice) I suggest it out of necessity; intuitions aren't going anyway for a while, but we should begin to work towards minimizing our dependence on them.

Perhaps this is a shift to a (intuition-free) science of linguistics from a (intuition-laden) philosophy of linguistics. I do not think such a breaking off would be a bad thing nor do I think that the two groups could not peacefully and productively co-exist. However, as discussed above, straws have been in the wind for some time and it is time to begin scrutinizing the presence, use, and meaning of intuitions in linguistics. Many empirical experiments (e.g. Chabris and Simons 2010 (Chabris and Simons 2010)) have shown us that our intuitions, even with respect to the visual domain, are not as reliable as once thought. Any reliance on intuitions in the linguistic domain will have to justify itself or prove beyond a doubt that it is not susceptible to the same kinds of foibles as intuitions in other domains.

Third, it shows promise for future development and expansion as the (methodology, discussion, and) speculation regarding a possible model of syntax demonstrates. In the last twenty years we have gone from no fMRI machines to advanced techniques for being able to analyze that data. Imagine where we will be in the next twenty years.

Fourth, it sets the stage for future empirical work that can be done in the present. I have suggested some specific brain regions that, based on experimental data, appear to be reasonable candidates for a pragmatic module (something that has previously been proposed in the literature). Am I correct? I don't know, but some clever experiments could help tip the scales of opinion.

Fifth, it helps to illuminate some of the implicit assumptions about inferring intentions (ToM in modern parlance) in Grice's original work, and begins to bring to bear the role that mind-reading plays in interpersonal communication. Furthermore, this interdependence between linguistic data and mind-reading data has the potential to be exploited in a number of different ways to shed light on how information transmitted on one level effects information transmitted on the other and how information on various levels interact in order to shape the direction of the talk exchange and to loosely explain and predict the contribution by each individual in said exchange.

Sixth, it recasts pragmatics (at least as conceptualized by the Maxims) into something more than just a loose set of guidelines that people follow for effective communication. Rather the Gricean Maxims are universally instantiated in people. The parameters of the maxims can be understood in terms of evolved tendencies and cultural conditions (historical factors, etc.). The universal instantiation was explained in evolutionary terms. This again recasts what pragmatics is to an extent because it gives us a different way of thinking about the conditions constraining why we communicate in the ways that we do.

As Pinker (2007), Pinker et al. (2009) show, the Gricean notion of an overarching Co-operative Principle is a bit too general. We do not always communicate in a cooperative manner. Consider the following:

> Would you like to come up and see my etchings? [a sexual come-on]
> If you could pass the guacamole, that would be awesome. [a polite request]
> Nice store you got there. Would be a real shame if something happened to it. [a threat]
> We're counting on you to show leadership in our Campaign for the Future. [a solicitation for a donation]
> Gee, officer, is there some way we could take care of the ticket here? [a bribe] (Pinker, Nowak, and Lee 2009)

They write, "a fundamental insight from evolutionary biology is that most social relationships involve combinations of cooperation and conflict" (Ibid). A notion of cooperation needs to be taken along with conflict avoidance, and "face" maintenance (Brown and Levinson 1998). Here I would add that ToM processes can be drawn on as an inference mechanism underlying the motivations of taking one approach (direct, indirect) over another.

Seventh, is that this approach explicitly integrates psychology, neurology (to a greater or lesser extent), philosophy, and linguistics into a unified framework (working towards consilience (Wilson 1998)). This approach provides a method of investigation which is complementary to different current approaches so it offers an opportunity for a synthesis of insight from other investigatory methodologies in linguistics. It also has the potential to foster interesting debates in other fields.

Synthesizing different methodologies makes the synthesis stronger than the sum of its parts. If an argument (and the evidence in support of it) in one field is found to complement or support and argument in another field (and the evidence in support of it) the new synthesis not only inherits the strength of the evidence from both fields but also, by the virtue of its being synthesized into one argument becomes an even stronger and more convincing story.

Lastly, this brings me to the final virtue of this approach which follows from the prior paragraphs. The explanation, by taking evolutionary psychology and a method of reverse-engineering and also by exploiting evolved characteristics of the brain (and those implications for the mind), I have grounded all of my argumentation in evolutionary theory which is the most accepted and attested fact in all of the natural sciences. With evolution as a bed rock for argumentation one need not worry about cracks.

Chapter 7

Appendix

It is easier and more informative to give an example than to describe this method of analysis. Take the following paragraph and analysis.

> A series of violent, bloody encounters between police and Black Panther Party members punctuated the early summer days of 1969. Soon after, a group of Black students I teach at California State College, Los Angeles, who were members of the Panther Party, began to complain of continuous harassment by law enforcement officers. Among their many grievances, they complained about receiving so many traffic citations that some were in danger of losing their driving privileges. During one lengthy discussion, we realized that all of them drove automobiles with Panther Party signs glued to their bumpers. This is a report of a study that I undertook to assess the seriousness of their charges and to determine whether we were hearing the voice of paranoia or reality. (Heussenstam, 1971,p. 32)

Proposition List for the Bumper stickers Paragraph:

1 (SERIES, ENCOUNTER)
2 (VIOLENT, ENCOUNTER)
3 (BLOODY, ENCOUNTER)
4 (BETWEEN, ENCOUNTER, POLICE, BLACK PANTHER)
5 (TIME: IN, ENCOUNTER, SUMMER)
6 (EARLY, SUMMER)
7 (TIME: IN, SUMMER, 1969)
8 (SOON, 9)
9 (AFTER, 4, 16)
10 (GROUP, STUDENT)
11 (BLACK, STUDENT)

12 (TEACH, SPEAKER, STUDENT)
13 (LOCATION: AT, 12, CAL STATE COLLEGE)
14 (LOCATION : AT, CAL STATE COLLEGE, LOS ANGELES)
15 (IS A, STUDENT, BLACK PANTHER)
16 (BEGIN, 17)
17 (COMPLAIN, STUDENT, 19)
18 (CONTINUOUS, 19)
19 (HARASS, POLICE, STUDENT)
20 (AMONG, COMPLAINT)
21 (MANY, COMPLAINT)
22 (COMPLAIN, STUDENT, 23)
23 (RECEIVE, STUDENT, TICKET)
24 (MANY, TICKET)
25 (CAUSE, 23, 27)
26 (SOME, STUDENT)
27 (IN DANGER OF, 26, 28)
28 (LOSE, 26, LICENSE)
29 (DURING, DISCUSSION, 32)
30 (LENGTHY, DISCUSSION)
31 (AND, STUDENT, SPEAKER)
32 (REALIZE, 31, 34)
33 (ALL, STUDENT)
34 (DRIVE, 33, AUTO)
35 (HAVE, AUTO, SIGN)
36 (BLACK PANTHER, SIGN)
37 (GLUED, SIGN, BUMPER)
38 (REPORT, SPEAKER, STUDY)
39 (DO, SPEAKER, STUDY)
40 (PURPOSE, STUDY, 41)
41 (ASSESS, STUDY, 42, 43)
42 (TRUE, 17)
43 (HEAR, 31, 44)
44 (OR, 45, 46)
45 (OF REALITY, VOICE)
46 (OF PARANOIA, VOICE)

Note, Lines indicate sentence boundaries. Propositions are numbered for ease of reference. Numbers as propositional arguments refer to the proposition with that number.

Bibliography

Abu-Akel, A. (2003). A neurobiologcal mapping of theory of mind. *Brain Research Reviews 43*(1), 29–40.

Adolphs, R. and M. Spezio (2009). *Handbook of Neuroscience for the Behavioral Sciences vol. 2*, Chapter Social Cognition. John Wiley and Sons, New Jersey.

Aiello, L. C. and R. Dunbar (1993). Neocortex size, group size, and the evolution of language. *Current Anthriopology 34*(2), 184–193.

Allan, K. (2004). Aristotle's footprints in the linguist's garden. *Language Sciences 26*(4), 317–342.

Anderson, J. R. (1984). *Cognitive Psychology and its Implications (6th ed.)*. New York: Worth Publishers.

Anderson, S. R. (2005). *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Andreasen, N. C. (1984). *Scale for The Positive Assessment of Symptoms (SAPS)*. Department of Psychiatry College of Medicine the University of Iowa.

Andreasen, N. C. (1985). *Psychiatry vol. 1*, Chapter Positive versus negative schizophrenia criteria and validation. New York: Plenum Publishing Corporation.

Autumn, K., Y. A. Liang, T. S. Hsieh, W. C. Zesch, P. Wai, T. W. Kenny, and R. F. ad Robert J Full (8 June 2000). Adhesive force of a single gecko foot-hair. *Nature 405*(6787), 681–685.

Autumn, K., M. Sitti, Y. A. Liang, A. M. Peattle, W. R. Hansen, S. Sponberg, T. W. Kenny, R. Fearing, J. N. Isrealachvili, and R. J. Full (September 17, 2002). Evidence for van der waals adhesion in gecko setae. *Proceedings of the National Academy of Sciences of the United States of America 99*(19), 12252–12256.

Avis, J. and P. L. Harris (1991). Belief-desire reasoning among baka children: Evidence for a universal conception of mind. *Child Development 62*(3 (Jun 1991)), 460–467.

Bach, K. (1999). *The Semantics-Pragmatics Interface from Different Points of View*, Chapter The Semantics-Pragmatics Distinction: What It Is and Why It Matters. Oxford: Elsevier.

Bar-Hillel, Y. (1971). Out of the pragmatic wastebasket. *Linguistic Inquiry 2*(3), 401–407.

Baron-Cohen, S. (1988). Without a theory of mind one cannot participate in a conversation. *Cognition 29*(1), 83–84.

Baron-Cohen, S. (1997). The study of autism reveals evolved mechanisms for mindreading. *Natural History Magazine 9*, 62–65.

Baron-Cohen, S., H. Ring, J. Moriarty, P. Shmitz, D. Costa, and P. Ell (1994). Recognition of mental state terms: a clinical study of autism, and a functional neuroimaging study of normal adults. *British Journal of Psychiatry 165*, 640–649.

Baron-Cohen, S., H. Tager-Flusberg, and D. J. Cohen (2000). *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford:Oxford University Press.

Barta, P. E., G. D. Pearlson, R. E. Powers, and L. E. . T. Stephanie Richards (November 1990). Auditory hallucinations and smaller superior temporal gyral volume in schizophrenia. *American Journal of Psychiatry*, 1457–1462.

Beeman, M. and C. Chiarello (1998). *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*. Mahwah: Lawrence Erlbaum Associates.

Beeman, M. J. (2005). Bilateral brain processes for comprehending natural language. *TRENDS in Cognitive Psychology 9*(11), 512–518.

Bennett, M. and B. H. Partee (2004). *Compositionality in Formal Semantics*, Chapter Towards the Logic of Tense and Aspect in English. Malden:Blackwell.

Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind and Language 17*(1 and 2), 37–54.

Boër, S. E. and W. G. Lycan (1980). A performadox in truth conditional semantics. *Linguistics and Philosophy 4*(1), 71–100.

Branan, N. (2010). Wired for categorization. *Scientific American Mind 20*(11).

Broca, P. (1965). Chapter ??

Brown, D. (1999). *The MIT Encyclopedia of the cognitive sciences*, Chapter Human Universals. Cambridge: MIT Press.

Brown, P. and S. Levinson (1998). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brownell, H., H. Gardner, P. Prather, and G. Martino (1995). *Handbook of neurological speech and language disorder*, Chapter Language, communication and the right hemisphere. New York: Dekker.

Brownell, H., H. Potter, A. Bihrle, and H. Gardner (1984). Sensitivity to lexical denotation and connotation in brain damaged patients: A double dissociation? *Brain and Language 22*, 253–265.

Brownell, H., H. Potter, A. Bihrle, and H. Gardner (1986). Inference deficits in right brain damaged patients. *Brain and Language 27*, 310–321.

Bruel, C. (2004). *Focus Structure in Generative Grammar: An integrated syntactic, semantic and intonational approach.* Amsterdam/Philadelphia: John Benjamins.

Brüne, M. (2005). "theory of mind"in schizophrenia: A review of the literature. *Schizophrenia Bulletin 31*(1), 21–42.

Calhoun, V., T. Adali, L. Hansen, J. Larsen, and J. Pekar (2003). *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, Chapter ICA of Functional MRI Data: An Overview. Proceedings of the ICA2003.

Carnie, A. (2002). *Syntax: A generative Introduction.* Malden:Blackwell.

Carrington, S. J. and A. J. Bailey (2009). Are there theory of mind regions in the brain?: A review of the neuroimaging literature. *Human Brain Mapping 30*(8), 2313–2335.

Chabris, C. and D. Simons (2010). *The Invisible Gorilla: And Other Ways Our Intuitions Deceive Us.* New York:Crown.

Chakabarti, B., E. Bullmore, and S. Baron-Cohen (2006). Empathizing with basic emotions: Common and discrete neural substrates. *Social Neuroscience 1*(3 and 4), 364–384.

Channon, S., A. Pellijeff, and A. Rule (2003). Social cognition after head injury: Sarcasm and theory of mind. *Brain and Language 93*(2), 123–134.

Chomsky, N. (1955/1975). *The Logical Structure of Linguistic Theory.* New York: Plenum Press.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge: MIT Press.

Chomsky, N. (1967). *Readings in the Psychology of Language*, Chapter A review of B.F. Skinner's Verbal Behavior. Englewood NJ: Prentice-Hall.

Chomsky, N. (1971). Conditions on rules of grammar. *Linguistic Analysis 2*, 303–351.

Chomsky, N. (1976). *Reflections on Language.* New York:Pantheon.

Chomsky, N. (1980/2005). *Rules and Representations.* New York:Columbia University Press.

Chomsky, N. (1997). *The Future of the Cognitive Revolution*, Chapter Language and Cognition. Oxford: Oxford University Press.

Chomsky, N. (2004a). *The Generative Enterprise Revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Noaki Fukui and Mihoko Zushi*. Berllin: Mouton de Gruyter.

Chomsky, N. (2004b). The master and his performance: An interview with noam chomsky. *Intercultural Pragmatics 1-1*, 93–111.

Clark, H. (1997). *Using Language*. Cambridge: Cambridge University Press.

Comon, P. (1994). Independent component analysis: a new concept? *Signal Processing 36*(3), 287–314.

Corballis, M. (2007). The uniqueness of human recursive thinking. *American Scientist 95*(3).

Cosmides, L. (1985). *Deduction or Darwinian Algorithms? An explanation of the "elusive" content effect on the Wason selection task*. Doctioral Dissertation, Harvard University. University Microfilms number 86-02206.

Cosmides, L. and J. Tooby (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Chapter Cognitive Adaptations for Social Exchange. New York: Oxford University Press.

Cosmides, L. and J. Tooby (1994). *Mapping the Mind: Domain specificity in cognition and culture*, Chapter Origins of domain-specificity: the evolution of functional organization. New York: Cambridge University Press.

Cosmides, L. and J. Tooby (n.d.). Evolutionary psychology: A primer. *n/a*, n/a.

Courtin, C. and A.-M. Melot (2005). Metacognitive development of deaf children: Lessons from the appearance reality and false belief tasks. *Developmental Science 8*, 16–25.

Culy, C. (1985). The complexity of the vocabulary of bambara. *Linguistics and Philosophy 8*(3), 345–351.

Dawkins, R. (1976/2006). *The Selfish Gene*. New York: Oxford University Press.

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua 117*, 1858–1878.

Dediu, D. and R. D. Ladd (2007). Linguistic tone is related to the population frequency of the adaptive halpogroups of two brain size genes, aspm and mircocephalin. *Proceedings of the National Academy of Sciences of the United States of America 104*(26), 10944–10949.

Dehaene, S., E. Spelke, P. Pinel, R. Stanescu, and S. Tsivkin (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science 284*(5416), 970–974.

Dehdari, J. (2006). *Crossing dependencies in Persian*. Ph. D. thesis, Brigham Young University.

Delis, D., W. Wapner, H. Gardner, and J. Moses (1983). The contribution of the right hemisphere to the organization of paragraphs. *Cortex 19*(1), 43–50.

Dennett, D. C. (1996). *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge: Cambridge University Press.

Doherty, M. (2008). *Theory of Mind: How Children Understand Others Thoughts and Feelings*. New York: Psychology Press.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews 6*(5), 178–190.

Eden, G. F. and T. A. Zeffiro (1998). Neural systems affected in developmental dyslexia revealed by functional neuroimaging. *Neuron 21*(2), 279–282.

Editors, T. (2006). Loud fake laugh misinterpreted as loud real laugh in critical sarcasm miscalculation. *The Onion 42-32*, August 7, 2006.

Editors, T. (2007). Is metaphor special? *Brain and Language 100*(2), 111–114.

Editors, T. (26 January 2010). Nasa concedes defeat in effort to free rover. *International Business Times*.

Ellman and Tait (2000). On the generality of thesaurally derived lexical links. *JADT 2000: 5es Journées Internationales d'Analyse Statistique des Données Textuelles*.

fen Hsu, C. and A. Karmiloff-Smith (2008). Language and williams syndrome. *Annual Review of Applied Linguistics 28*(1), 191–204.

Ferstl, E., J. Neumann, C. Bodger, and D. von Cramon (2007). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping 29*, 581–593.

Ferstl, E., M. Rinck, and D. von Cramon (2005). Emotional and temporal aspects of situation model processing during text comprehension: an event-related fmri study. *Journal of Cognitive Neuroscience 17*, 724–729.

Filler, A. G. (2009). The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: Ct, mri, and dti. *nature prepublication series http://precedings.nature.com/documents/3267/ version/5*.

Fiske, D. (2010). Watch out for semantic chimeras. *unpublished paper*.

Fitch, W. T. (2010). *The Evolution of Human Language: Biolinguistic Perspectives*, Chapter Three meanings of "recursion": key distinctions for biolinguistics. Cambridge: Cambridge University Press.

Fletcher, P., F. Happé, U. Firth, S. Baker, R. Dolan, R. Frackowiak, and C. Frith (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition 57*(2), 109–128.

Fodor, J. and J. J. Katz (1964). *The Structure of Language: Readings in the Philosophy of language*, Chapter Introduction. New Jersey: Prentice Hall.

Foldi, N. (1987). Appreciation of pragmatic interpretation of indirect commands: Comparison of right and left hemisphere brain-damaged patients. *Bran and Language 31*, 88–108.

Frankfurt, H. (2005). *On Bullshit*. Princeton: Princeton University Press.

Frith and Allen (1988). *Schizophrenia: The major issues*, Chapter Language Disorders in Schizophrenia and their implications for neuropsychlogy. Oxford: Heinemann Medical.

Frith, C. D. (1992). *The Cognitive neuropsychology of schizophrenia*. East Sussex:Lawrence Erlbaum Associates Ltd.

Gardner, H. (1976). *The Shattered Mind: the person after brain damage*. London:Routledge and Kegan Paul Ltd.

Gardner, H., H. Brownell, W. Wapner, and D. Michelow (2004). *Cognitive Processing in the right hemisphere*, Chapter Missing the point: the role of the right hemisphere in the processing of complex linguistic materials. New York : Academic Press.

Geschwind, N. (Nov. 27, 1970). The organization of language and the brain. *Science 170*(3961), 940–944.

Givón, T. (2001). *Syntax: An Introduction Volume II*. Amsterdam/Philadelphia: John Benjamins.

Glucksburg, S. (2004). *Experimental Pragmatics*, Chapter On the Automaticity of Pragmatic Processes: A Modular Proposal. Houndmills: Palgrave Macmillian.

Glucksburg, S. (February 2003). The psycholinguistics of metaphor. *TRENDS in Cognitive Science 7*(2), 92–96.

Golan, O., S. Baron-Cohen, S. Wheelwright, and J. Hill (2006). Systemizing empathy: Teaching adults with asperger syndrome and high functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology 18*(2), 589–615.

Goodglass, H. (1992). *Conduction Apasia*, Chapter Diagnosis of conduction aphasia. Hillsdale, New Jersey : Lawrence Erlbaum Associates, Inc. Publishers Hillsdale.

Goodglass, H. and M. Hunter (1970). A linguistic comparison of speech and writing in two types of aphasia. *Journal of Communicative Disorders 3*(1), 28–35.

Gray, P. (2007). *Psychology (5th ed.)*. New York: Worth Publishers.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Griefat, Y. and T. Katriel (1989). *Language, communication, and culture: Current Directions*, Chapter Life demands Musayara: Communication and Culture Among Arabs in Israel. Newbury Park, California: SAGE Publications.

Groenendijk, J. and F. Roelofsen (2009). Inquisitive semantics and pragmatics. *Presented at the Stanford Workshop on Language, Communication and Rational Agency*.

Grossman, M. A. (1980). A central processor for hierarchically-structured material: evidence from broca's aphasia. *Neuropsychologia 18*(3), 299–308.

Hajiĉová, E., B. H. Partee, and P. Sgall (1998). *Topic-focus articulation, tripartite structures, and semantic content*. Dordrecht: Kluwer.

Harris, P., M. de Rosnay, and F. Pons (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science 14*(2), 69–73.

Harris, Z. (1952). Discourse analysis. *Language 28*(1), 1–30.

Hauser, M. D., N. Chomsky, and W. T. Fitch (22 November 2002). The faculty of language: What is it, who has it, and how did it evolve? *Science 298*(5598).

He, B. J., A. Z. Snyder, J. M. Zempel, M. D. Smyth, and M. E. Raichle (2008). Electrophysiological correlates of the brains intrinsic large-scale functional architecture. *Proceedings of the National Academy of Sciences of the United States of America 104*(41), 16039–16044.

Heine, B. and T. Kuteva (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hengeveld, K. and J. L. Mackenzie (2008). *Functional Discourse Grammar*. Oxford: Oxford University Press.

Hickok, G., P. Erhard, J. Kassubek, A. K. Helms-Tillery, S. Naeve-Velguth, J. P. Strupp, P. L. Strick, and K. Ugurbil (2000). A functional magnetic resonance imaging study of the role of the left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. *Neuroscience Letters 287*(2), 156–160.

Higginbotham, J. (1984). English is not a context free language. *Linguistic Inquiry 15*(2), 225–234.

Higginbotham, J. (Spring 1985). Reply to pullum. *Linguistic Inquiry 16*(2), 298–304.

Hirst, W., J. LeDoux, and S. Stein (1984). Constraints on the processing of indirect speech acts: Evidence from apasiology. *Brain and Language 23*(1), 26–33.

Hockett, C. (1960). The origin of speech. *Scientific American 203*, 88–111.

Hof, P. R. and E. van der Grucht (2007). Structure of the cerebral cortex of the humpback whale, *Megaptera novaengliae* (cetacea, mysticeti, balaenopteridae). *The Anatomical record 290*, 1–31.

Horn, L. R. (1996). *Handbook of Contemporary Semantic theory*, Chapter Presupposition and Implicature. Malden: Blackwell.

Horwitz, B., K. Amunts, R. Bhattacharyya, D. Patkin, K. Jeffries, K. Zilles, and A. R. Braun (2003). Activation of broca's area during the production of spoken and signed language: a combined cytoarchitectonic mapping and pet analysis). *Neuropsychologia 41*(14), 1868–1876.

Hough, M. S. (2003). Narrative comprehension in adults with right and left hemisphere brain- damage: Theme organization. *Brain and Language 38*(2), 253–277.

Hubel, D. and T. Wiesel (May 19, 1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences 198*(1130), 1–59.

Hudson, R. (June 1994). About 37 % of word-tokens are nouns. *Language 70*(2), 331–339.

Huxley, J. (1942/2010). *Evolution: The Modern Synthesis*. Cambridge: MIT Press.

Huybregts, R. (1976). Overlapping dependencies in dutch. *Utrecht Working Papers in Linguistics 1.24-65.*

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jacobsen, L. K., J. N. Giedd, F. X. Castellanos, A. C. Vaituzis, S. D. Hamburger, S. Kumra, and M. L. J. L. Rapoport (May 1988). Progressive reduction of temporal lobe structures in childhood-onset schizophrenia. *Am J. Psychiatry 155*(5), 678–685.

Joanette, Y. and P. Goulet (1990). *Discourse ability and brain damage: Theoretical and empirical perspectives*, Chapter Narrative discourse in right-brain-damaged right handers. New York: Springer-Verlag.

Joanette, Y., P. Goulet, B. Ska, and J.-L. Nespoulous (1986). Informative content of narrative discourse in right-brain-damaged right-handers. *Brain and Language 29*(1), 81–105.

Judson, O. (May 4, 2010). Darwin got it going on. *The New York Times Opinionator.*

Kadmon, N. (2001). *Formal Pragmatics*. Malden:Blackwell.

Kanwisher, N., J. McDermott, and M. Chun (June 1, 1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience 17*(11), 4302–4311.

Kaplan, J. A., H. H. Brownell, J. R. Jacobs, and H. Gardner (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain and Language 38*(2), 315–333.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge: MIT Press.

Kintsch and vanDijk (1978). Toward a model of text comprehension and production. *Psychological Review 85*, 363–394.

Klein, S. B. (2000). *Biological Psychology.* New Jersey: Prentice Hall.

Krifka, M. (1989). *Semantics and Contextual Expression*, Chapter Nominal Reference, Temporal Constitution and Quantification in Event Semantics. Dordrecht: Foris Publication.

Kumaran, D., J. J. Summerfield, D. Hassabis, and E. A. Maguire (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron 63*(6), 889–901.

Labov, W. (1975). *The Scope of American Linguistics: Papers of the First Golden Anniversary Symposium of the Linguistic Society of America*, Chapter Empirical foundations of linguistic theory. Lisse: Peter de Ridder.

Labov, W. (1996). *Papers from the Parasession on Theory and Data in Linguistics*, Chapter When Intuitions Fail. Chicago: Chicago Linguistic Society.

Lancker, D. V. and J. Cummings (1999). Expletives: neurolinguistic and nuerobehavioral perspectives on swearing. *Brain Research Reviews 31*(1), 83–104.

Landman, F. (2008). *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, Chapter 1066. On the differences between the tense-perspective-aspect systems of English and Dutch. Amsterdam: John Benjamins.

Landman, F. and S. Rothstein (2009). *Syntax, Lexical Semantics and Event Structure*, Chapter Incremental homogeneity in the semantics of aspectual for-phrases. Amsterdam: John Benjamins.

Lawson, J., S. Baron-Cohen, and S. Wheelwright (2004). Empathising and systemising in adults with and without asperger syndrome. *Journal of Autism and Developmental Disorders 34*, 301–310.

Lebedev, M. A. and M. A. Nicolelis (2006). Brain-machine interfaces: past, present and future. *TRENDS in Neuroscience 29*, 536–546.

Leslie, A. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review* (94), 412–426.

Leslie, A. (1994). *Mapping the Mind: Domain specificity in cognition and culture*, Chapter ToMM, ToBY, and Agency: Core architecture and domain specificity. Cambridge: Cambridge University Press.

Leslie, A. and D. Roth (1993). *Understanding other minds: Perspectives from autism*, Chapter What autism teaches us about metarepresentation. Oxford: Oxford University Press.

Levinson, S. (2000). *Pragmatics*. Cambridge: Cambridge University Press.

Ley, R. G. and M. Brydean (1982). A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain and Cognition 1*(1), 3–9.

Li, C., P. B. Umbanhowar, H. Komsuoglu, D. E. Kodltschek, and D. I. Goldman (2009). Sensitive dependence of the motion of a legged robot on granular media. *Proceedings of the National Academy of Sciences of the United States 106*(24), 3029–3034.

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology Review 16*(3), 131–148.

Lombardo, M., J. Barnes, S. Wheelwright, and S. Baron-Cohen (2007). Self-referential cognition and empathy in autism. *PLoS ONE 2*(9).

Ma, D., R. Rabionet, I. Konidari, J. C. Jawoski, H. Wright, R. Abrason, J. Gilbert, M. Cuccaro, M. Pericak-Vance, and E. Martin (2010). Association and gene-gene interaction of slc6a4 and itgb3 in autism. *Am J Med. Genet Part B 153B*, 477–483.

Machlachlan, A. and O. Rambow (2002). Cross-serial dependencies in tagalog. *Proceedings of the Sixth International Workshop on tree Adjoining Grammar and Related Frameworks (TAG+6)*, 100–104.

MacNeilage, P. (2009). The origin of the left and right brain. *Scientific American 301*(7), 48–55.

Mameli, M. (2007). *Oxford Handbook of Evolutionary Psychology*, Chapter Evolution and Psychology in philosophical perspective. Oxford: Oxford University Press.

Mar, R. A. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologia 42*(10), 1414–1434.

Martin, I. and S. McDonald (2003). Weak coherence, no theory of mind, or executive dysfunction? solving the puzzle of pragmatic language disorders. *Brain and Language 85*(3), 451–466.

Martinich, A. (1980). Conversational maxims and some philosophical problems. *The Philosophical Quarterly 30*(120), 215–228.

Mashal, N. and M. Faust (2008). The role of the theory-of-mind cortical network in the comprehension of narratives. *Brain and Language 104*(2), 103–112.

Mason, R. A. and M. A. Just (2009). Right hemisphere sensitivity to novel metaphoric relations: Application of the signal detection theory. *Language and Linguistics Compass 3*(1), 157–174.

Matthews, P. (1993). *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

McCarley, R. W., D. F. Saltsbury, Y. Hirayasu, D. A. Yurgelun-Todd, M. Tohen, C. Zarate, R. Kikinis, F. A. Jolesz, and M. E. Shenton (2002). Association between smaller left temporal posterior superior temporal gyrus volume on magnetic resonance imaging and smaller left temporal p300 amplitude in first-episode schizophrenia. *Arch. Gen. Psychiatry 59*(4), 321–331.

McGrath, J. (1991). Ordering thoughts on thought disorder. *British Journal of psychiatry* (158), 307–316.

McGurk, H. and J. MacDonald (1976). Hearing lips and seeing voice. *Nature 264*, 746 – 748.

Meaburn, E., P. Dale, I. Craig, and R. Plomin (2002). Language-impaired children: No sign of the foxp2 mutation. *Neuro Report (Cognitive Neuroscience and neuropsychology) 13*(8), 1075–1077.

Menon, R. R., P. E. Barta, E. H. Aylward, S. S. Richards, D. D. Vaugn, A. Y. Tien, G. J. Harris, and G. D. Pearlson (1995). Posterior superior temporal gyrus in schizophrenia: grey matter changes and clinical correlates. *Schizophrenia research 16*(2), 127–135.

Merryman, G. D. and P. A. Bronson (Dec. 232009). Smart people gesture more when they talk-so will kids be smarter if they gesture? *Newsweek (Nurture Shock)*.

Morgan, G. and J. Kegl (2006). Nicaraguan sign language and theory of mind: the issue of critical periods an abilities. *Journal of Child Psychology and Psychiatry 47*(8), 811–819.

Morris, C. W. (1938). *Foundations of the Theory of Signs*. Chicago: University of Chicago.

Myers, P. and R. Bookshire (1996). Effect of visual and inferential variables on scene descriptions by right-hemisphere-damaged and non-brain-damaged adults. *Journal of Speech and Hearing Research* (39), 870–880.

Myers, P. S. (1978). *Clinical Aphasiology: Proceedings of the Conference*, Chapter Analysis of Right Hemisphere Communication Deficits: Implications for Speech Pathology. Minneapolis, Minnesota:BRH Publishers.

Nestor, P. G., M. E. Shenton, R. W. McCarley, J. Haimson, R. S. Smith, B. O'Donnell, M. Kimble, R. Kikinis, and F. A. Jolesz (1993). Neuropsychological correlates of mri temporal lobe abnormalities in schizophrenia. *Am. J. Psychiatry 150*(12), 1849–1855.

Newbury, D., E. Bonora, J. Lamb, S. Fischer, C. Lai, G. Baird, L. Jannoun, V. Slonims, C. Stott, M. Merricks, P. Bolton, A. Bailey, A. Monaco, and the International Molecular Genertic Consortium (2002). Foxp2 is not a major susceptibility gene for autism or specific language impairment. *American Journal of Human genetics 70*(5), 1318–1327.

Newport, E. L. (1999). *Language Creation and Language Change: Creolization, Diachrony, and Development*, Chapter Reduced Input in the Acquisition of Signed Languages: Contributions to the Study of Creolization. Cambridge: MIT Press.

Ogawa, S., T. Lee, A. Nayak, and P. Glynn (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine 14*, 68–78.

Peattie, A. and R. Full (2007). Phylogenetic analysis of the scaling of wet and dry biological fibrillar adhesives. *Proceedings of the National academy of Sciences of the United States of America 104*(47), 18595–18600.

Peterson, C. (2004). Theory-of-mind development in oral deaf children with cochlear implants or conventional hearing aids. *Journal of Child Psychology and Psychiatry 45*(6), 1096–1106.

Philippe, A., M. Martinez, M. Guilloud-Bataille, C. Gilberg, M. Røastam, E. Sponheim, M. Coleman, M. Zapella, H. Aschauer, L. van Maldergem, C. Penet, J. Feingold, A. Brice, M. Leboyer, and the Paris Autism Research International Sibpair Study (1999). Genome-wide scan for autism susceptibility genes. *Human Molecular Genetics 8*(5), 805–812.

Pinker, S. (1984). *Language Learnability and Language development*. Cambridge: Harvard University Press.

Pinker, S. (1994/2007a). *The Language Instinct (P.S. edition)*. New York: Harper Collins.

Pinker, S. (1997). *How the Mind Works*. New York: W.W. Norton and Co.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. London: Weidenfeld and Nicolson.

Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. New York: Penguin.

Pinker, S. (2007b). *The Stuff of Thought:Language as a Window int Human Thought*. New York: Penguin.

Pinker, S. and P. Bloom (1990). Natural language and natural selection. *Behavioral and Bran Sciences 13*(4), 707–784.

Pinker, S., M. A. Nowak, and J. J. Lee (2009). The logic of indirect speech. *Proceedings of the National academy of Sciences of the United States of America 105*(3), 833–838.

Potts, C. (2006). Review of siobhan chapman, paul grice: Philosopher and linguist. *Mind 115*(459), 743–747.

Potts, C. (2010). *Routeledge Encyclopedia of Pragmatics*, Chapter Formal Pragmatics. London: Routledge.

Premack, D. and G. Woodruff (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci. 1*(4), 515–526.

Pullum, G. K. (1985). Remarks and replies: Such that clauses and the context-freeness of english. *Linguistic Inquiry 16*(2), 291–298.

Pullum, G. K. (1991). *The Great Eskimo Vocabulary Hoax: and other irreverent Essays on the Study of Language(chapters 16 p. 131-138 and 17 p.139-146)*. Chicago: University of Chicago Press.

Pullum, G. K. and G. Gazdar (1982). Natural languages and context-free languages. *Linguistics and Philosophy 4*(4), 471–504.

Pullum, G. K. and B. C. Scholz (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review 18*(1-2), 9–50.

Pyers, J. E. and A. Senghas (2009). Language promotes false-belief understanding: Evidence from learners of a new sign language. *Psychological Science 20*, 805.

Quigg, M. and N. B. Fountain (1999). Conduction aphasia elicited by stimulation of the left posterior superior temporal gyrus. *J. Neurol Neurosurg Psychiatry 66*(3), 393–396.

Ramachandran, V. (1992). Filling in gaps in perception: Part i. current directions in psychological science. *Current Directions in Psychological Science 1*(6), 199–205.

Ramachandran, V. (2004). *A brief tour of Human consciousness: from Impostor poodles to Purple numbers*. New York: Pi Press.

Ramchand, G. and C. Reiss (2007). *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press.

Renzi, E. D., D. Perani, G. Carlesimo, M. Silvers, and F. Fazio (1994). Prosopagnosia can be associated with damage confined to the right hemisphere - an mri and pet study and a review of the literature. *Neuropsychologia 32*(8), 893–902.

Rorden, C. and H. Karnath (2004). Using human brain lesions to infer function - a relic from a past age in the fmri age? *Nature Reviews Neuroscience 5*, 813–819.

Ross, E. D., J. A. Edmondson, G. B. Seibert, and R. W. Homan (1988). Acoustic analysis of affective prosody during right-sided wada test: A within-subjects verification of the right hemisphere's role in language. *Brain and Language 33*(1), 128–145.

Ross, J. R. (1967). *Constraints on variables in syntax*. Ph. D. thesis, M.I.T.

Sabbagh, M. A. (1999). Communicative intentions and language: Evidence from right hemisphere damage and autism. *Brain and Language 70*(1), 29–69.

Sahin, N. T., S. Pinker, S. S. Cash, D. Schomer, and E. Halgren (2009). Sequential processing of lexical, grammatical, and phonological information within broca's area. *Science 326*(5951), 445–449.

Sampson, G. (1987). Untitled review of chomsky's system of ideas by fred d'agostino. *The Philosophical Quarterly 37*(149), 477–481.

Sampson, G. (2001). *Empirical Linguistics*. London: Continuum.

Sandler, W. (i.p.). *The Oxford International Encyclopedia of Linguistics*, Chapter Sign Language Phonology. Oxford: Oxford University Press.

Sapolsky, R. (2010). Personal communication april 18, 2010.

Saxe, R., S. Carey, and N. Kanwisher (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Ann Rev. Psychol. 55*, 87–124.

Saxe, R. and L. J. Powell (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science 17*(8), 692–699.

Schmithorst, V. J., S. K. Holland, and E. Plante (2006). Cognitive modules utilized for narrative comprehension in children: A functional magnetic resonance imaging study. *NeuroImage 29*(1), 254–266.

Schmitt, D. P. and J. J. Pilcher (2004). Evaluating evidence of psychological adaptation: How do we know one when we see one? *Psychological Science 15*(10), 643–649.

Scholz, B. C. and G. K. Pullum (2006). *Contemporary Debates in Cognitive Science*, Chapter Irrtational nativist exuberance. Oxford: Blackwell.

Schultz, J., H. Imamizu, M. Kawato, and C. D. Frith (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of Cognitive Neuroscience 16*(10), 1695–1705.

Selkirk, E. (2002). *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody*, Chapter Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English.

Senghas, A., S. Kita, and A. Özyürek (2004). Children creating core properties of language: evidence from an emerging sign language in nicaragua. *Science 305*(5702), 1779–1782.

Shaller, M., J. H. Park, and D. T. Kenrick (2007). *Oxford Handbook of Evolutionary Psychology*, Chapter Human evolution and social cognition. Oxford: Oxford University Press.

Shapin, S. (2010). The darwin show. *London review of books 32*(17), 1695–1705.

Shapiro and Danly (1985). The role of the right hemisphere in the hemisphere in the control of speech prosody in prepositional and affective contexts. *Brain and Language 25*(1), 19–36.

Shaywitz, S. E., B. A. Shaywitz, K. R. Pugh, R. K. Fulbright, R. T. Constable, W. E. Mencl, D. P. Shankweiler, A. M. Liberman, P. Skudlarski, J. M. Fletcher, L. Katz, K. E. Marchione, C. Lacadie, C. Gatenby, and J. C. Gore (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National academy of Sciences of the United States of America 95*(5), 2636–2641.

Shieber, S. M. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy 8*, 333–343.

Silverman, I. (2008). *Narrative comprehension in adults with right hemisphere brain damage: the role of coherence and theme organization.* Ph. D. thesis, Brown University.

Spencer, N. (1972). Differences between linguists and nonlinguists in intuitions of grammaticality-acceptability. *Journal of psycholinguistic research 2*(2), 83–98.

Sperber, D. and D. Wilson (1986/1995). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell Ltd.

Sperber, D. and D. Wilson (2002). Pragmatics, modularity and mind-reading. *Mind and Language 17*(1 and 2), 3–23.

Stone, J. (2005). *Encyclopedia of Statistics in Behavioral Science Volume 2*, Chapter A Brief Introduction to Independent Component Analysis. John Wiley and Sons.

Stone, V., S. Baron-Cohen, and R. T. Knight (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience 10*(5), 640–656.

Stone, V. E., L. Cosmides, J. Tooby, N. Kroll, and R. T. Knight (2002). Selective impairment of reasoning about social exchange in patient with bilateral limbic system damage. *Proceedings of the National academy of Sciences of the United States of America 99*(7), 11531–11536.

Stringaris, A. K., N. C. Medford, V. Giampietro, M. J. Brammer, and A. S. David (2007). Deriving meaning: Distinct neural mechanisms for metaphoric, literal, and non-meaningful sentences. *Brain and Language 100*(2), 150–162.

Surian, L., S. Baron-Cohen, and H. V. der Lily (1996). Are children with autism deaf to gricean maxims? *Cognitive Neuropsychiatry 1*(1), 55–71.

Surian, L. and M. Siegal (2001). Sources of performance on theory of mind tasks in right hemisphere-damaged patients. *Brain and Language 78*(2), 224–232.

Szabó, Z. G. (2005). *Semantics versus Pragmatics*. Oxford: Oxford University Press.

Szabolci, A. (2001). *The Handbook of Contemporary syntactic theory*, Chapter The Syntax of Scope. Oxford: Blackwell.

Tager-Flusberg, H. and R. M. Joseph (2003). *Autism: mind and brain*, Chapter Identifying neurocognitive phenotypes in autism. Oxford: Oxford University Press.

Taylor, I. and M. M. Taylor (1990). *Psycholinguistics: Learning and using Language*. Englewood N.J.: Prentice-Hall International.

Tomasello, M. (2003). *Constructing a Language: A Usage based theory of Language acquisition*. Cambridge: Harvard University Press.

Turner, K. (1999). *The Semantics-Pragmatics Interface from Different Points of View*. Oxford: Elsevier.

Twist, J. (2005). Universe 'too queer' to grasp. *BBC News 12 July 2005.*

von Heusinger, K. and K. Turner (2005). *Where Semantics meets pragmatics.* The Michigan State University Papers.

Wasow, T. and J. Arnold (2005). Intuitions in linguistic argumentation. *Lingua 115*, 1481–1496.

Wassink, T. H., J. Piven, V. J. Vieland, J. Pietila, R. J. Goedken, S. E. Folstein, and V. C. Sheffield (2002). Evaluation of foxp2 as an autism susceptibility gene. *American Journal of Medical Genetics (Neuropsychiatric Genetics) 114*, 556–569.

Wernicke, C. (1874). *Classics in Psycholinguitics*, Chapter Der aphasische Symptomencomplex. Eine psychologishe Studie auf anatomischer Basis. Amsterdam:John Benjamins.

Wernicke, C. (1885). *Das Urwindungssytem des menslichen Gehirns.* ??

Wernicke, C. (1906). *Der Aphasie Symptomenkomplex.* ??

Weylman, S., H. Brownell, M. Roman, and H. Gardner (1989). Appreciation of indirect requests by left- and right-brain-damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language 36*(4), 580–591.

Wilmer, J. B., L. Germine, C. F. Chabris, K. Williams, E. Loken, K. Nakayama, and B. Duchaine (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences of the United States of America 107*(11), 5238–5241.

Wilson, E. (1998). *Consilience: The unity of knowledge.* New York: Vintage Books.

Wilson, S. J., K. Parsons, and D. C. Reutens (2006). Preserved singing in aphasia: A case study of the efficacy of melodic intonation therapy. *Music Perception 24*(1), 23–36.

Winkler, S. (1997). *Focus and secondary Predication.* Berlin: Mouton de Gruyter.

Xu, J., S. Kemeny, G. Park, C. Frattali, and A. Braun (2005). Language in context:emergent features of word, sentence, and narrative comprehension. *NeuroImage* (25), 1002–1015.

Yamada, J. E. (1990). *Laura: A case study for the Modularity of Language.* Cambridge: The MIT Press.

Yngve, V. H. (1996). *From grammar to science: new foundations for general linguistics.* Amsterdam:Benjamins.

Zaenen, A., J. Mailing, and H. Thráinsson (1985). Case and grammatical functions: The icelandic passive. *Natural Language and Linguistic Theory 3*, 441–483.

Zubizarreta, M. J. (1998). *Prosody, Focus, and Word Order.* Cambridge: MIT Press.

Start by marking â€œToward a Theory of Cultural Linguisticsâ€ as Want to Read: Want to Read savingâ€¦ Want to Read.Â Palmer begins by showing how cognitive grammar complements the traditional anthropological approaches of Boasian linguistics, ethnosemantics, and the ethnography of speaking. He then applies his cultural theory to a wealth of case studies, including Bedouin lamentations, spatial organization in Coeur d'Alene place names and anatomical terms, Kuna narrative sequence, honorifics in Japanese sales language, the domain of ancestral spirits in Proto-Bantu noun-classifiers, Chinese counterfactuals, the non-arbitrariness of Spanish verb forms, and perspective schemas in English discourse. Evolutionary linguistics is a subfield of psycholinguistics that studies the psychosocial and cultural factors involved in the origin of language and the development of linguistic universals. The main challenge in this research is the lack of empirical data: spoken language leaves practically no traces. This led to the abandonment of the field for more than a century, despite the common origins of language hinted at by the relationships among individual languages established by the field of historical Evolutionary Linguistic Theory (ELT) is an international peer-reviewed journal intended as a platform for discussing the question of the origin and development of the language faculty understood as a specifically dedicated part of the human mind/brain and its connection with the human cognition. The specificity of the journal is to contribute to the ongoing debate on language origin from an explicitly linguistic viewpoint which examines its complex subject from a well-grounded knowledge in theoretical linguistics (with its subsystems, psycholinguistics, neurolinguistics, language acquisition a