THE BURDEN OF SYNTAX

by

Joseph James Charles Thompson

B.A. (Hons.), Simon Fraser University, 2007

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
IN THE
DEPARTMENT OF PHILOSOPHY
FACULTY OF ARTS AND SOCIAL SCIENCES

© Joseph James Charles Thompson 2011

SIMON FRASER UNIVERSITY

Summer 2011

APPROVAL

Name: Joseph James Charles Thompson

Degree: Master of Arts

Title of Thesis: The Burden of Syntax

Examining Committee: Dr. L. Shapiro (Chair)

_____

Dr. R.E. Jennings
Professor of Philosophy
Senior Supervisor

_____

Dr. M. Hahn
Associate Professor of Philosophy
Supervisor

_____

Dr. C. Eliasmith
Associate Professor of Philosophy
and Systems Design Engineering
University of Waterloo
External Examiner

_____

Dr. T.P. Racine
Assistant Professor of Psychology
Internal External Examiner

Date Approved: April 27th, 2011

Declaration of
Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Abstract

A theorist's responsibility for the theory he proposes is inseparable from his responsibility for the language in which the theory is expressed. Language science is no different in this regard. Theorists agree that language is the product of a population, and individual humans participate in linguistic social activity. But disagreements about typology cannot be mediated solely by aprioristic considerations. It has become a norm of scientific discourse that theoretical idiom is to be vindicated empirically. Evidence is needed both to serve as common ground and to make disciplined inferences about the individual sequencing demands of colloquy. I take up Philip Lieberman's proposal that circuits projecting to and from the basal ganglia could underlie aspects of motor, linguistic, and perhaps even cognitive sequencing. The evidence illustrates how the study of individual sequencing requirements might proceed without aprioristic typology.

To Michelle

Acknowledgments

Contents

List of Figures

List of Abbreviations

CFG	Context-free Grammar
CSC	Cortical-Striatal-Cortical Circuit
FLN	Narrow faculty of Language
GPe	Globus Pallidus (exterior)
GPi	Globus Pallidus (interior)
LOT	Language of thought
LOT-RCE	A language of thought having sentences which contain centre-embedded recursive elements
OMO	Odd-Man-Out task
PD	Parkinson's Disease
SNr	Substantia Nigra pars Reticulata
SNc	Substantia Nigra pars Compacta
STN	Subthalamic nucleus
TMS	Test of meaning from syntax
VOT	Voice onset time
WSCT	Wisconsin card sort

Chapter 1

Introduction

A theorist's responsibility[1] for the theory he proposes is inseparable from his responsibility for the language in which the theory is expressed. Thus virtually any language that finds its way into empirical discourse is the responsibility of those who participate in the discourse. The rejection of the language of *phlogiston* and that of *caloric fluid* went beyond recording the falsity of sentences expressed in those languages. The choice of theoretical language need not be grounded *merely* in convention or in current fashion. Evidence and theory more generally make 'electron' a respectable topic for empirical discussion, and establish mathematically definable elements of electrocardiographic records as suitable proposals for new diagnostic tools. This is why theorists who would insist upon a theoretical language that mentions angels, demons, and Ouija boards, face an uphill battle for acceptance.

Our interests here lies in identifying a suitable idiom for studying the evolution of syntactically complex sentences, and where possible vindicating this idiom empirically. Much of this discussion will consider what demands syntactically complex sentences place on the individual, as traditional approaches to language evolution have argued that participating in more syntactically complex languages requires fancier neural machinery. This inference is not obvious, and can only be made against background theoretical assumptions. What is uncontentious is that language is the product of a population, and individual humans

---

[1] By *responsibility* I just mean something like *accountability*.

simply *participate* in a complex social activity, which we conversationally refer to as *colloquy.* In what follows, I make every effort to be theory neutral with respect to what, exactly, constitutes a *language.* But it will be useful, at least for pedagogical purposes, to reveal this author's biases. It will also be useful in getting some sense of who will benefit most from this work.

Some readers will wonder why we do not take up the familiar Chomskian and Saussurean starting point. This would begin with an ominous reminder of the 'heterogeneity,' or extensive variation, of colloquy (Saussure, 1959). The standard method for making a respectable science out of this unwieldy construct is to distinguish an abstract set of sentences from the messy performative attempts at producing sentences. It is then added that linguistic competence lies with mechanisms allowing the capacity to produce and comprehend those sentences. On such a view, participation in a syntactically complex language requires special neural machinery to support the comprehension and production of sentences of the right degree of complexity. These assumptions are easy to find in the study of language origins. Old representations of natural language as 'hierarchically' complex (Chomsky, 1957; Chomsky and Miller, 1963) have occasioned comparative psychological research into what animals posses the capacity to 'learn' formal grammars of 'hierarchical' complexity (Fitch and Hauser, 2004; Gentner et al., 2006). We will return to Chomsky's typology below.

This work will be of more immediate interest to those who think that the study of language origins is best left as a study of a peculiar *activity* of human populations. This is not to deny, of course, that human beings can learn hierarchical formal languages. But the capacity to converse about formal grammars is not the capacity to participate in syntactically complex language. For some theorists, such as Bickerton (2009), syntactic complexity is, for the most part, an artifact of analysis. Some rudimentary sequencing abilities are certainly required to participate in syntactically complex colloquy, and theorists are happy to admit as much (Bickerton, 2009; Tomasello, 1998, 2005, 2008), but these sequencing requirements need not correspond to our abstract theories of sentence complexity. If *colloquy* is the fundamental explanandum of language science, then any ensuing typology of colloquy into sentences, while perhaps useful for some purposes, need not reveal much about colloquy's demands on her participants.

We begin, then, with some choices in theoretical language that, as a practical matter, are unexceptionable. There is little doubt that the idiom of evolutionary biology is well-suited to the study of humans and human origins. Some theorists might insist that psychology requires its own *independent* idiom, perhaps because of the special complexity of the subject matter. There might also be disagreement about the extent to which theorists abuse the idiom of *evolutionary biology*. But both biologists and psychologists, at the very least, are obliged to co-opt the idiom of *populations*. As other animals vary, so do humans, and it is inescapable that that diversity is as crucial to the human domain as it is elsewhere in biology.

There is no reason to suppose that even theories of short-term language change can or must be different from other theories in this regard. As far as this author knows, every theorist referred to in this work would grant (at any rate once it is pointed out) that incidental patterns of effects of earlier linguistic constructions can be exploited and can become the main patterns of effect pattern of later productions. Though one might characterize the effect of *going* (as in *I am going to the well*) as suggesting motion towards, this, far from preventing, must also create the suggestion of *duration*. Thus duration is residual, and perhaps even primary, in such uses as *I am going to stay where I am* (Jennings and Thompson, submitted for publication). Similarly, *have* can be relieved of its capacity to suggest possession and instead used to suggest attainment of a state (*I have done nothing)* (Jennings and Thompson, submitted for publication). Such observations about language are, on their face, consistent with any of the positions ranging from those of Skinner (1981), which are founded on (non-mentalistic) general learning abilities, to those of Steven Pinker (1994), which are founded on an innate language of thought (also see chapter 6).

The language of *populations*, with its tolerance for diversity, provides an attractive starting point for an inquiry into linguistic behaviour, even when our goal is to describe the *structure* of this behaviour. Colloquy, that is the general phenomenon of linguistic exchange, is typically a temporally extended interaction involving speech, gesture, external cues, and more than one human. The features of *colloquy*, like the features of organic life generally, are features of a population. Variation within the population is to be expected, and the precise extension of non-arbitrary features, such as *dog*, are typically unclear (if a precise extension exists at all). While the structural features of formal grammars have precise

definitions, it does not seem impossible that some sorts of *grammatical species* might exhibit the kind of vagueness typical of properties in biology. One such example might be the notion of well-formed or a *fluent* speech. These would not be grammatical features as is usually understood, as they belong to the realm of colloquy rather than to the syntax of a language. But aside from this taxonomic point, the features of colloquy could be as unamenable to definition as any biological feature. It is only when we impose further typology upon colloquy that genuine disagreement arises.

It should be noted that some have suggested that colloquy, even prior to a distinction between language and pragmatics (Saussure, 1959), exhibits structural features which are available for exploitation by its participants (Turnbull, 2003).

> The structural nature of the orderly methods for doing talk derives from the adjacent positioning of utterances and of turns. A turn may consist of one utterance or several utterances, each of which is adjacent to a specific other utterance. Further, each turn is adjacent to a specific other turn. Thus, the major method for doing talk is the sequential location of an utterance in a turn and of a turn in a sequence of turns (Turnbull, 2003, p. 141).

Such accounts need not, such as Grice's did, emphasize communicative goals of colloquy (Turnbull, 2003), and they certainly do not assume that the *purpose* of colloquy is to communicate information. They also do not impose a strict understanding of the 'sentence,' as colloquy, on such a view, is driven primarily by *turns* at talk. Turns at talk need not be made up of sentences. They could be composed as much by a suitably placed pause as by speech:

1. Student: do you respond to email?

   (1.5 second pause)

2. Instructor: sometimes (adapted from, Turnbull, 2003, p. 94).

While such structural descriptions have been defended, the more familiar Chomskian start-

ing point remains dominant. The variation in social interaction, according to this approach, is either too prominent or too misleading for fruitful study. Saussure, in particular, argued that we should introduce an abstract notion of *language* in order to circumvent the 'heterogeneity of colloquy.'

> Taken as a whole, speech is many-sided and heterogeneous; straddling several areas simultaneously — physical, physiological and psychological — it belongs both to the individual and to society; we cannot put it into any category of human facts, and we cannot discover its unity (Saussure, 1959, 9).

This touches on a deep controversy surrounding the choice of idiom for studying colloquy. Many invoke a notion of *language* in order to abstract away from the performative facts of colloquy more generally. The majority of variation in colloquy probably falls under parole for Saussure, and classified as misleading variation that potentially confounds any study of language.[2]

Saussure's typology was a great boon for generative linguistics, and Chomsky recognized as much (Chomsky, 1964, p. 52). It sets aside the majority of variation in colloquy as *non-linguistic*, and this is crucial for most formal accounts of sentence composition. The notion of *syntactic composition*, when applied to naturally occurring linguistic sequences, is problematic. The syntactic composition of a formal language is guaranteed by recursively defined conditions for well-formed sentences. The semantics is defined along this recursion to guarantee that the semantic value of whole sentences is determined by the meaning of their parts. Formal languages also enjoy, unlike natural languages, a well-defined class of atoms. This allows one to identify the form of a sentence with its set of substitution instances. New atoms are introduced into natural language all the time in the form of idioms. Consider, for example,

1. He's a diamond in the rough.

---

[2] It is important to remember that how we choose to type colloquy is independent of the question as to whether there are any linguistic universals. Theorists can argue about the pervasiveness of a linguistic feature while agreeing that many utterances may be of the same linguistic type, despite non-linguistic differences which are due to social, psychological or other pragmatic factors.

2. He's a rhinestone in the rough (Dorothy Parker).

Insofar as we represent such items as semantically atomic, we are compelled to view natural language as lacking a stable class of atoms (Thompson, 2010b). The way to circumvent such problems, if they can be avoided at all, is to provide an understanding of *language* such that sentences are composed from a fixed stock of atoms, regardless of performative variation.

This discussion boils down to what is, in effect, a tautology. Providing structural descriptions of colloquy using virtually any formal grammar requires, at the very least, typing of colloquy into atomic parts, and if these atomic parts are not readily found within colloquy, then a distinction between *language* and *parole* is useful indeed.

The moral is that while theorists freely adopt the language of *population* and *variation*, they disagree about what additional structural vocabulary is needed for language science. Theorists need structural types to recognize structural similarities. If the variation exhibited by colloquy is problematic, then it is problematic because it leaves us without the precise categories required for disciplined inquiry.

No one seems to be under the illusion that a suitable, and yet well-understood, idiom is available. Those who have settled on some theoretical vocabulary will be quick to acknowledge that this vocabulary is poorly understood, or at least awaiting a satisfactory definition or explication.

Theorists cannot, for example, claim to know what syntactic form *is*, in the case of natural language. We are no where near a complete theory of grammar, and prominent linguists admit as much (Jackendoff, 1994, p. 26). Yet a near-complete grammatical theory is necessary for an understanding of *form* with respect to a natural language.

Difficulties with the notion of *uniform substitution* nicely illustrate why *form* requires a developed theory of grammar. The form of an expression, in the case of formal languages, is typically defined by reference to the expression's substitution instances. But uniform substitution is problematic in the case of natural language. For example, it is fine to claim that

'you have coffee or you have tea'

is a disjunction. However,

'You *may* have coffee or you *may* have tea'

is not a disjunction at all (Jennings, 2005). From the latter, one is permitted to infer that 'you may have coffee,' which, on a disjunctive reading, is to infer a disjunct from a disjunction. All paradoxes of this sort must be avoided if one is to have a complete understanding of uniform substitution with respect to natural language. But it seems that such paradoxes can be avoided only with a complete, or near-complete, theory of grammar. So while formal languages can identify the form of a sentence with the corresponding set of substitution instances, *an incomplete grammatical theory precludes an analogous identification for natural language.* This is, at least partially, due to the fact that, when representing a natural language in a formal one, form is not cleanly distinguished from semantics. One needs to consider the various semantic uses of *or*, before they can comfortably represent an expression with the $\vee$.

Some might argue that the idiom of *sequencing* is somehow better off than the idiom of *linguistic form* insofar as only the latter is restricted to the domain of language. Is is true that 'sequencing' applies perfectly well outside the domain of language, such as in motor control. Some kind of *sequencing ability* is certainly required for an organism to participate in complex behavioural sequences. 'Motor acts, phonemes, words, phrases, sentences, turns at talk, subvocal phonological loops, and even strategic dispositions to behave could all be the subject of sequencing' (Thompson, 2010b).

However, the more general idiom of *sequencing* does not avoid the aforementioned difficulties. While we may have a rough idea of how to characterize the parts of a motor sequence (a pirouette, a somersault, and so on), one still needs to specify the atomic components if we are to identify sequences of similar structure. Motor sequences cannot be idiomatic if we want to distinguish their forms.

Confusion surrounding the language of *sequencing* is brought to the fore particularly in discussions of *hierarchical complexity*. The complexity of the sequences in an organism's

behavioural repertory is supposed to reveal something about the organism's neural re-
sources. Joseph Devlin (2006), in his review of Lieberman (2006), takes up this principle in
objecting to the view that a single neural system could provide both motor and linguistic
sequencing.

> Motor actions such as walking and dancing tend to have linear structure, whereas
> even simple sentences are based on a hierarchical structure. Consequently, se-
> quencing in dance and language is likely to require fundamentally different mech-
> anisms, and equating the two will be misleading. In other words, it is not se-
> quencing, per se, that is important for language, but the ability to produce and
> comprehend *hierarchically structured* sequences (emphasis added, Devlin, 2006,
> p. 314).

This argument relies on the claim that motor control lacks hierarchical complexity. It is
ironic that Lieberman has also suffered through an objection presupposing the *opposite*.

> [. . . If] aspects of language such as grammatical categories, constraints on co-
> reference, and inflectional morphology had parallels in the motor system, this
> would support the view that language evolved out of the speech motor control.
> [Lieberman] makes no argument for such parallels; indeed, once one put aside
> gross commonalities such as *hierarchical structure and sequencing* (properties
> that are shared by virtually all complex systems), these domains have nothing
> in common (emphasis added, Bloom, 1992, p. 383).

That two evidently inconsistent taxonomies could be assumed as obvious suggests a much
deeper confusion with the notion of *hierarchical complexity* and *sequencing*, one that perhaps
rivals confusion about *form*.

We unabashedly pursue inquiry into the structure of colloquy, even when we are either
(a) unsure what idiom is most suitable for our study or (b) fail to understand the idiom
they have settled upon. Fortunately, a theorist is not *expected* to provide a definition for
every word used outside the 'methods' section of a technical paper. It is often assumed,

especially in psychology, that scientific inquiry can be conducted even when we have only a rudimentary understanding of the properties we postulate.

On the dominant view, there is no pressing burden to provide a complete semantic account for every property of theoretical interest (Cronbach and Meehl, 1955). Mentalistic vocabulary constitutes perhaps the most infamous example of vocabulary which plays a theoretical role without being given a complete semantic account. What *is* necessary is to fix on a method for *measuring*, or at least tracking, properties of interest. The especially troubling properties, and this presumably includes 'sequencing,' are often called *hypothetical constructs*, and much of the literature presupposes two methodological claims.

   i Whether a test X is a valid measure of theoretical construct C is an empirical question which can be evaluated by further research.

   ii Empirical research will allow theorists to formulate more precise reconstruals of theoretical constructs which already have representatives in common discourse (this may or may not include folk-psychological constructs such as *intelligence, belief, desire, and so on).*

Cronbach and Meehl (1955) defend an extremely popular account which adopts i and ii. Determining whether a test is a good measure of some theoretical construct, which is to say the test has 'construct validity,' is an enterprise extending beyond the scope of a single study. A theorist may study whether a test measures intelligence by checking how test-performance relates to academic success (Cronbach and Meehl, 1955). This may further construct validation, because one would expect a relationship between intelligence and academic success. But observed relations between test performance and academic success would not *demonstrate* validity because no one believes that academic success is a perfect measure of *intelligence*. There is typically no single variable (or 'criterion') which could be used to establish construct validity (Cronbach and Meehl, 1955). Demonstrating construct validity is, therefore, a theory-wide enterprise. Theorists will often need, for instance, to appeal to a number of criterion variables which, according to our best theory, are expected to relate to intelligence. At first, this may include a teacher's perception of intelligence. Theorists will probably need to see how their test relates to a variety of measures before

they can have much confidence that the test tracks intelligence more accurately than, say, the measures of academic success.

The theorist who argues that temperature is reflected in mercury-expansion may be in a similar predicament (Cronbach and Meehl, 1955). Demonstrating the that mercury-thermometers measure temperature, at least in early theorizing, must make reference to imperfect measures, such as the scientist's relative perception of *hot* and *cold*. Once a large enough body of evidence is acquired, theorists may determine that mercury-thermometers are more reliable measures of temperature than their perceptions of warmth.

Cronbach and Meehl (1955) account for ii by positing a 'nomological network' of scientific laws³ which, taken as a whole, constrain our capacity to explicate theoretical terms. '[...Even a] vague, avowedly incomplete network still gives the constructs whatever meaning [*sic.*] they do have' (Cronbach and Meehl, 1955, p. 12). One develops these empirical laws as science proceeds, and these laws jointly specify the observable consequences for the satisfaction of a construct-property. Modern physics has a nomological network which entails that temperature is more accurately tracked by mercury-thermometer than by common sense intuitions of *warm* and *cool*. Of course, if a theorist is to view nomological networks as actually constraining the *meaning* of terms such as 'temperature,' then construct validity theory obviously presupposes a particular philosophy of language which need not be adopted here.

1.1 A Proposal

Perhaps what is most attractive about construct validity theory is the fact that it relieves theorists from the immediate burden of giving semantic accounts for theoretical-vocabulary. It allows theorists to *study* intelligence while putting off answering the question as to what intelligence *is*. It should be noted, however, that Cronbach and Meehl (1955) insisted that *some* common ground was necessary for construct validation.

Unless substantially the same nomological net is accepted by the several users

---

³ These laws, according to Cronbach and Meehl (1955) could be statistical or deterministic.

> of the construct, public validation is impossible. If A uses aggressiveness to
> mean overt assault on others, and B's usage includes repressed hostile reactions,
> evidence which convinces B that a test measures aggressiveness convinces A that
> the test does not. Hence, the investigator who proposes to establish a test as
> a measure of a construct must specify his network or theory sufficiently clearly
> that others can accept or reject it [. . .] (Cronbach and Meehl, 1955, p. 291).

The disagreements among, for instance, Hauser, Chomsky, and Fitch (2002), Pinker and
Jackendoff (2005), Tomasello (1998, 2005, 2008), Cowley (2002, 2004), Lieberman (2006),
Turnbull (2003), and Jennings (1994, 2005, 2007) are presumably too foundational for passive investigation alone to reveal the appropriate idiom for describing the structure of colloquy. Construct validation requires more common ground between theorists. This is further
justification for taking colloquy, at least tentatively, as the primary explanandum of language science.

We need not adopt the particular picture of Cronbach and Meehl (1955), nor need we
adopt the philosophy of language implicit in it.[4] We should be interested, however, in using
empirical research to evaluate any structural idiom used for typing colloquy. As long as we
have a philosophy of science that allows us to tell when our measures are measuring what
we want them to measure, and allows us to reconstrue our idiom[5] to fit our needs, then
neuroscientific research could inform our views on the structure of colloquy. Those who
rely on these assumptions in the study of mental and biological properties should not be
squeamish about applying them to the study of sequencing.

My specific proposal is to focus on the notion of *hierarchical complexity*, to be discussed in
chapters 2 and 3, as it constitutes the idiom of choice for many inquiries into the evolution of
language. It is perhaps most directly appealed to in exchanges among Hauser, Chomsky, and
Fitch (2002) and Pinker and Jackendoff (2005).[6] Of especial interest are the basal ganglia, a

---

[4] Some philosophers might, for instance, object to the idea that empirical observation could assist in uncovering the meaning of a term (also see 5.3).

[5] Whether *reconstrual* constitutes explication of the meaning of a term or the introduction of entirely new vocabulary is irrelevant for our purposes.

[6] Also see Fitch et al. (2005)

collection of sub-cortical structures, and their apparent involvement in hierarchical linguistic sequencing. In conjunction with the neural circuits they form with the cortex, the basal ganglia allow for sequencing in a variety of domains. The project, which seems implicit in Lieberman (2006), and which I will be explore in this work, is to reconstrue and evaluate our chosen structural idiom in light of evidence from disorders affecting the basal ganglia.

1.2   Chapter Outline

The discussion has the following structure. Chapter 2 will discuss, in more detail, what is meant by *hierarchy* and *recursion*. Both are frequently discussed in the study of language origins (Hauser, Chomsky, and Fitch, 2002). The argument for Hierarchical complexity in English often appeals to relative clauses. Consider, for example,

> The following conversation, which took place between the two friends in the pump-room one morning, after an acquaintance of eight or nine days, is given as a specimen of their very warm attachment [...] (Austen, 1870).

The capacity to engage with such linguistic items is supposed by many to reveal something about the sequencing powers of individual organisms and, ultimately, their brains. While a detailed discussion of hierarchical structure is left for chapter 2, it is not hard to see more immediately why relative clauses might entail special sequencing demands for an organism. Subjects may bear long-distance relationships with main verbs in centre-embedded relative clauses, and understanding these sentences seems first to require sensitivity to this relationship, even where the subject and main verb are far apart.

Hierarchical structure is especially important to Hauser, Chomsky, and Fitch (2002), who have proposed that the capacity for recursion was a milestone in language evolution. They proposed research into the formal grammar 'learning' capacities available to animals. The goal of this research is to judge whether non-humans are capable of learning formal grammars of hierarchical complexity (Hauser et al., 2002; Fitch and Hauser, 2004; Fitch et al., 2005; Gentner et al., 2006).

One prominent notion of hierarchical structure is rooted in the class of context-free languages, which occupy a particular position in the *Chomsky hierarchy*. This class of languages get their hierarchical status from meta-theoretic proofs about a particular relationship between classes of automata and classes of formal grammar. The class of languages representable in context-free grammars, for example, are co-extensive with the class of language which can be 'understood,' or accepted, by a pushdown automata (Hopcroft and Ullman, 1979, p. 219-253). Chomsky (1957) argued that English cannot suitably be represented in a finite state language, a class of grammars occupying a proper subset of the context-free languages. We will also discuss the notion of a *recursive embedding* on which Chomsky's argument relies.

The rest of chapter 2 considers other notions of hierarchial structure. Some theorists do not seem to take recursive embeddings to be crucial to hierarchical structure. Instead, they take the hierarchical structure of natural language to be guaranteed by its convenient representation in the appealing phrase-structure-trees (see figure 2.3) that are often used to *describe* the sentences of a context-free grammar. Chapter 2 concludes with a discussion of Lieberman's (2006) argument that we should represent motor sequences as hierarchical, and consider what notion of hierarchical complexity he has in mind.

In chapter 3, it will be argued that many, perhaps most, structural theories of language cannot be insulated from theories of neural architecture. For some, this will simply amount to the claim that our structural descriptions of colloquy are, themselves, theoretical constructs whose 'meaning' is to be malleated by empirical inquiry. For others, this will simply be an example of *modus tollendo tollens*. The literature of interest already assumes that our structural representations of colloquy have consequences for neural architecture. It is therefore plausible that our structural representations may come into question inasmuch as they generate insuperable anomalies.

Chapter 3 also makes the crucial distinction between structural descriptions of natural language and structural descriptions of mental representations. It is one thing, for instance, to claim that *English* contains grammatical categories (*nouns, verbs, and so forth*). It is another thing entirely to claim that some class of compositional mental representations contains nouns and verbs. Those theorists who posit compositional mental representations

(for our purposes this is all that is meant by a language of thought) might provide entirely different descriptions for English and 'Mentalese,' as long as this did not hinder their theories of sentence comprehension.

A detailed discussion of Lieberman's work is postponed until chapter 4, where we consider basal ganglia architecture in more detail and consider the claim that the basal ganglia, or circuits projecting through the basal ganglia, are involved in the understanding of hierarchically complex sentences. Of especial importance here is the hypothesis that similar neural processes might play a 'syntactic role' in a variety of motor, linguistic, and perhaps even cognitive sequences.

Chapter 4 also presents a portion of the evidence that Lieberman uses to evaluate the idiom of *recursion*. This evidence stems from inquiry into Parkinson's disease and hypoxia, disorders which affect basal ganglia functioning. When taken together with a well developed understanding of basal ganglia architecture, this evidence suggests that similar neural sequencing processes might underlie the apprehension[7] of a special class of relative clauses, the sequencing of motor acts such as dancing, and the modification of plans or strategies given negative feedback.

Lieberman uses this evidence to evaluate the structural idiom of *recursion*. Lieberman, like Chomsky, advances a theory about the minimal individual requirements for participating in centre embedded relative clauses. Lieberman (2006) calls this minimal capacity 'reiteration,' a form of sequencing supported by cortical-subcortical-cortical circuits. However, 'reiteration' is a form of sequencing which is not sensitive to *recursive elements*. In Lieberman's opinion, the representation of sentences as containing recursive elements does a poor job of accounting for how the brain supports production and understanding of relative clauses.

> It is an open question as to whether Chomsky's theories have any biologic validity. Can we really be sure that the relative, prepositional, and adverbial clauses

---

[7] I resist the vocabulary of *understanding* here in order to avoid the presumption that 'understanding a sentence' requires the tokening of a mental representation or the parsing of that sentence into a formal language. Unless otherwise specified, I do not use to *understanding* in this more technical, mentalistic, sense.

> that occur in complex sentences are derived from embedded sentences? [...]
> In short, recursion as defined by Hauser, Chomsky, and Fitch (2002) follows
> from hypothetical syntactic structures, hypothetical insertions of sentences and
> phrase nodes, and hypothetical theory-specific rewriting operations on these
> insertions to form the sentences we actually hear or read (Lieberman, 2006,
> p.359-360).

The phrase *Biologic validity* seems to be a species of *hypothetical construct validity*, and 'hypothetical syntactic structures' seem to constitute 'hypothetical constructs.' The constructs serve as a theoretical idiom that is to be evaluated, or perhaps refined, empirically. Reiteration, according to Lieberman, is simply the hypothetical structural construct with better evidential support. It might even be called a 'biological construct,' because it has been evaluated primarily with regard to neuroscience and pathology research.

More generally, the claim that similar processes underwrite such a wide array of sequencing behaviours may turn out to be problematic for some theories of language structure. It is especially relevant to those who, like Devlin, represent linguistic sequences as exhibiting a level of complexity over and above that of motor sequences. Such claims must be reconciled with any findings that suggest similar neural sequencing powers in motor and linguistic systems. One cannot maintain, without argument, that evolution produced a human motor sequencing mechanism that was substantially more powerful than was necessary to perform its task.

In chapter 5 we will consider how theorists might respond to this evidence. They may (sections 5-5.2) attack Lieberman's empirical argument or try to reconcile their structural theories with the data (5.4). A few responses to the argument are especially important.

The most obvious response to the arguments presented in chapter 4 will be to deny the validity of Lieberman's measures. Given the presupposition of a philosophy of science where we can evaluate the quality of our measures by means of empirical observation, demonstrating test validity is a task best left to psychologists. We will, therefore, be able to defend Lieberman's measures only in outline. The force of Lieberman's empirical argument rests on a number of empirical assumptions regarding what various tests actually measure.

Chapter 5 also considers some more specific worries. The argument presented in chapter 4 presupposes that a *single* neural system underlies both motor and linguistic sequencing. Lieberman (2006), himself, seems to advocate this when he describes the basal ganglia as a 'sequencing engine,' which suggests a unitary entity.[8] However, Lieberman recognizes a longstanding hypothesis that the basal ganglia may be organized into a number of anatomically and functionally segregated cortical-basal-ganglionic-cortical *circuits*. Even if disorders of the basal ganglia engender motor, linguistic, and cognitive sequencing difficulties, it would not follow that a single system is responsible for all three tasks. Damage to the basal ganglia may affect a multitude of neural systems simultaneously.

It will also be argued that basal ganglia research could still be relevant to our structural theories of language, even if linguistic sequencing was supported by a dedicated and discrete cortical-subcortical-cortical circuit. It remains plausible that segregated subcortical-cortical circuits have fundamentally similar architectures. The functional individuation of circuits, if this hypothesis is true, lies in their respective *cortical pairings*, rather than their basal ganglionic components. The segregation of linguistic and motor systems, then, does not preclude the possibility that similar neural processes, realized in separate systems, support a wide range of sequencing behaviours. This conjecture is not insulated from empirical revision, and it will become clear what kinds of studies could call it into question.

Chapter 6 briefly considers the implications of basal ganglia research for theories of language evolution. Lieberman (2006), of course, considers this question in far more detail. I am interested in whether this evidence should be seen as in tension with the more general view that multi-clause sentences reflect something deeper about the complexity of colloquy (Chomsky and Miller, 1963). One might question, as Lieberman seems to, whether relative clauses should be represented as containing recursively embedded sentences, a matter crucial

---

[8] One can take neural systems to be functional descriptions of some physical substrate. That is, one can individuate neural systems with respect to their functions (and theorists typically individuate *function* by either causal or evolutionary role). There is no requirement that neural systems be innate, modular, or restricted to a particular brain region. It is therefore possible that systems overlap with one another. On this rather weak understanding of *system*, the basal ganglia might be a sequencing engine which is comprised of many segregated subsystems, each supporting entirely different abilities. On such an understanding, Lieberman's claim cannot be considered contentious.

to the question of whether a language is hierarchical. But even if a different structural idiom (such as *reiteration*) is chosen to represent the individual sequencing demands of multi-clause constructions, relative clauses may yet reflect something interesting about the structure of colloquy.

The structural idiom of reiteration is well fitted to Lieberman's evolutionary hypothesis. Lieberman believes that the neural systems allowing for the execution of syntactically complex *motor sequences*, and the systems allowing for the apprehension of syntactically complex language, are cognate (Lieberman, 2006). The suggestion is that ancestral motor sequencing abilities engendered modern human syntactic abilities. This evolutionary story is also well fitted to one conjecture regarding the evolution of relative clauses: that while spoken conversation and relative clauses are taxonomically remote, relative clauses are cognate with a more general structural feature of colloquy, requiring some sequencing capacity such as *reiteration*.

Chapter 2

Recursion, Reiteration, and Hierarchy

2.1 *Hierarchical Structure*

The research proposed by Hauser, Chomsky, and Fitch (2002) is intended to uncover the origins of individual sequencing powers for participation in colloquy. They begin with some preconceptions about what is required of participants, and then attempt to track the evolution of participatory capacities. Hauser, Chomsky, and Fitch (2002) take the hierarchical status of linguistic sequences as an uncontentious starting point for this investigation.

> Natural languages go beyond purely local structure by including a capacity for recursive centre embedding of phrases within phrases, which can lead to statistical regularities that are separated by an arbitrary number of words or phrases. Such long-distance, hierarchical relationships are found in all natural languages for which, at a minimum, a 'phrase-structure grammar'[1] is necessary (Hauser, Chomsky, and Fitch, 2002, p. 1577).

---

[1] For our purposes, we can identify phrase-structure grammars with context-free grammars.

If this is right, then it is natural to ask where the capacity to partake in complex colloquy came from. This opens up comparative questions about whether non-humans can 'learn' formal grammar (Fitch and Hauser, 2004; Gentner et al., 2006; Hochmann et al., 2008; Herbranson and Shimp, 2008). Fitch and Hauser (2004), for example, have argued that cotton-top tamarins cannot 'learn' simple hierarchical grammars. In the idiom of Hauser, Chomsky, and Fitch (2002), the syntactic requirements for participation in a language is satisfied by a capacity for *recursion*. They advance two claims.

1. *Recursion* satisfies crucial syntactic requirements necessary for participating in hierarchical conversation.

2. *Recursion* is both unique to language and to human beings.

Claim 2 has been criticized more extensively elsewhere (Pinker and Jackendoff, 2005), and we will not discuss it in detail. Our interest is primarily with claim 1, which Lieberman (2006) seems to call into question on empirical grounds.

Controversy also surrounds the notion of *hierarchical structure*. Even though this structural feature is associated with a well-defined class of formal grammars (which we will discuss shortly), theorists disagree on basic questions regarding what *sequences in nature* are to be represented as hierarchical.

Joseph Devlin, for instance, has attempted to get substantial *kilomerage* out of the hierarchical structure of linguistic sequences (Devlin, 2006). He suggests that this typing of the linguistic data introduces serious problems for so called 'motor control' theories of syntax, which accept that neural systems supporting motor sequencing could be co-opted for the purposes of linguistic sequencing.[2] Devlin argues that since linguistic sequences have hierarchical structure, and motor sequences do not, we should expect fundamentally different neural systems to underwrite the two abilities.

> Motor actions such as walking and dancing tend to have linear structure, whereas even simple sentences are based on a hierarchical structure. Consequently, se-

---

[2] This view roots in Lashley (1951).

> quencing in dance and language is likely to require fundamentally different mech-
> anisms, and equating the two will be misleading. In other words, it is not se-
> quencing, per se, that is important for language, but the ability to produce and
> comprehend *hierarchically structured* sequences (emphasis added, Devlin, 2006,
> p. 314).

We will have the theoretical means of evaluating Devlin's argument shortly. For now, I merely point out that the hierarchical structure of motor sequences has also been conceded as *obvious*, even by Lieberman's detractors. Consider Bloom's response to the idea that a single general mechanism might perform motor sequencing (for acts such as dancing) and linguistic sequencing.

> [... If] aspects of language such as grammatical categories, constraints on co-
> reference, and inflectional morphology had parallels in the motor system, this
> would support the view that language evolved out of the speech motor control.
> [Lieberman] makes no argument for such parallels; indeed, once one puts aside
> gross commonalities such as *hierarchical structure and sequencing* (properties
> that are shared by virtually all complex systems), these domains have nothing
> in common (emphasis added, Bloom, 1992, p. 383).

Such conceptual embarrassments ought to prompt a fuller discussion of *hierarchical structure.* We will then have the means to gain some understanding what is meant by *recursion.*

2.2   A Formal Account of Hierarchy

One might see languages as hierarchical in a variety of ways. All hierarchies must place their elements in some kind of ordering where lower elements are seen as somehow depen-dent on higher elements. Perhaps the place of individual sentences is determined by the minimum number of calls to a function required to generate the sentence. Consider the simple grammar G, which takes p as its only atom. The set of well-formed formulae are closed under the following conditions.

[G_1] p is a well-formed formula

[G_2] For any wff α, *pα is a wff

[G_3] For any wff α, *αp is a wff

[G_4] Nothing else is a wff

Even this grammar is representable in a hierarchical tree structure. This would generate tiers of well-formed sentences, differing in structural complexity. For example,
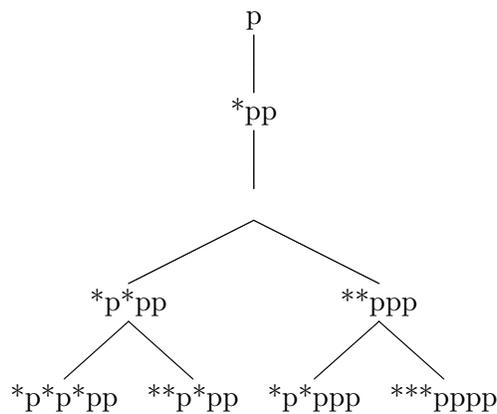
Figure 2.1: *A simple notion of hierarchical structure*

The notion of hierarchy illustrated by figure 2 is uninteresting at best. It reflects nothing more than the fact that the set of well-formed sentences are picked out by a recursive definition of the language. A more interesting notion of hierarchical structure would pick out only a subset of the grammars which satisfy this weak notion of hierarchy.

Hierarchical structure is typically represented in context-free grammars (or CFGs), which satisfy the following properties (adapted from Hopcroft and Ullman, 1979, p. 171):

1. CFGs include a set of sentences arranged by concatenated atomic parts that belong to *the terminal vocabulary*. CFGs also include a set of *non-terminal* vocabulary which, in the representation of natural languages, stand for structural descriptions of sentences in the terminal vocabulary.

2. The non-terminal vocabulary is non-empty, and contains, at a minimum, a starting symbol, S.

3. CFGs include a set of generative rules that include the metalinguistic symbol '⇒,' where A⇒B is to be read as 'if A is an acceptable string (of either terminal or non-terminal symbols) in the language, then B is an acceptable string.'

4. CFGs have only rules of the form A⇒ ϕ (where ϕ is an arbitrary string of symbols) and never the form χAψ⇒ $\chi\phi\psi$ (where χ and ψ are non-empty).[3]

5. A is a non-terminal symbol iff there exists a rule such that A⇒ ϕ[4]

One might think that phrase-tree structures have something to do with the hierarchical structure of *natural* language. This is because phrase-tree structures assign sentences grammatical descriptions which are conveniently depicted in a hierarchical tree structure. Consider the following grammar, G2.

S ⇒ NP + VP
NP ⇒ D + N
VP ⇒ V + NP
D ⇒ the,
N ⇒ elephant, duck, cow,
V ⇒ ate,

Each of the sentences in G2 can be given a structural description such as that of figure 2.2.

CFGs are often represented as providing structural descriptions to their sentences, including noun phrases, verb phrases, and so on.[5] An adequate representation (or complete theory)

---

[3] Relaxing this constraint and allowing context-sensitive rules of the form χAψ⇒ $\chi\phi\psi$ (where χ and ψ are non-empty) captures a larger class of grammars (Chomsky and Miller, 1963).

[4] For a full definition, see Hopcroft and Ullman (1979)

[5] To avoid equivocation we use the abbreviations of NP, VP, and so on as non-terminal symbols in a formal language. Theorists may (or may not) take these to represent noun phrases, verb phrases, and so on.
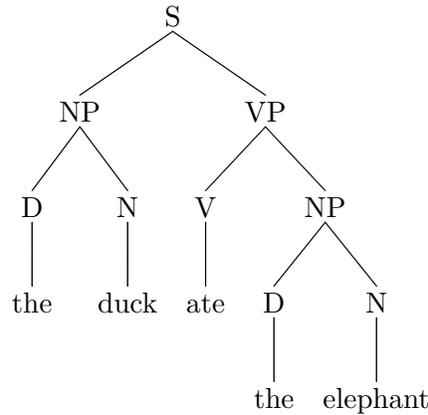
Figure 2.2: *A sentence of G2*

of language is supposed to provide such structural descriptions (Chomsky and Miller, 1963, p. 285).

Context-free grammars naturally are the focus of much discussion. This is largely because the class of context-free grammars hold a particular place in the Chomsky hierarchy of formal grammars (Chomsky and Miller, 1963). The complexity of a formal grammar is judged by looking to the relationship between classes of grammars and classes of automata. The class of languages defined by context-free grammars, for instance, *is identical to* the class of languages that can be 'understood' or accepted by some pushdown automaton (Hopcroft and Ullman, 1979, p. 219-253). We need not concern ourselves with the definition of finite state machines or pushdown automata here. It is important only to recognize that the class of finite state languages is a subset of the context-free languages.[6]

One famous context-free grammar is the $A^n B^n$ grammar, whose well-formed sentences contain all and only sentences with n items from some A class followed by n items from some B class (eg. ab, aabb, aaabbb . . .). Fitch and Hauser (2004) hoped to ascertain whether cotton-top tamarin monkeys possessed hierarchical sequencing capacity by asking whether they could 'learn'[7] this grammar. Their choice was not a coincidence. Chomsky (1957)

---

[6] We are only interested in the context-free languages characterized by strings of terminal symbols.

[7] We should be careful about the idiom of 'learning' in such studies. Fitch and Hauser (2004) used a popular habituation paradigm where tamarins were familiarized to one formal grammar, and then presented with

used the $A^n B^n$ grammar in his argument that English was not a finite state language, but one with *hierarchical complexity*, better represented (though perhaps not perfectly represented) in a context-free grammar. The $A^n B^n$ grammar is easily captured in a context-free grammar, that allows for centre embedding (eg. S⇒ASB and S⇒AB). It is tempting to represent English sentences as containing similar embeddings (for example, *the man who thought that S_1 was mistaken*).

> It is clear, then, that in English we can find a sequence $a + S_1 + b$, where there is a dependency between a and b, and we an select as S_1 another sequence containing $c + S_1 + d$, where there is a dependency between c and d, then select as S_2 another sequence of this form, etc. (Chomsky, 1957, p. 22)

This alone, of course, is no argument that English is not a finite-state language. The complete argument relies on metatheoretic proofs about the extent to which finite state grammars can tolerate embeddings (see, for example, Hopcroft and Ullman, 1979; Chomsky, 1957).

When theorists talk about hierarchical sequences they often mean sequences exhibiting *long distance dependencies.* These are sequences whose syntactic categories may bear long distance relationships with one another. The paradigm is a centre-embedded relative clause where the subject is separated from the main verb by a long clause (or a series of clauses).

> The following conversation, which took place between the two friends in the pump-room one morning, after an acquaintance of eight or nine days, is given as a specimen of their very warm attachment [. . .](Austen, 1870).

Understanding this sentence is supposed to require the recognition that *is given as a specimen of their very warm attachment* takes the *conversation* as its subject, despite the embedded clauses. Chomsky draws upon the intuition that there is no non-arbitrary upper bound on the number of clauses which can be embedded into a grammatical sentence. (eg.

---

sequences violating that grammar. Animals 'learned' the grammar if they were more likely to shift overt attention to the speakers when presented with a violation.

'the boy who believes that he believes that he believes ... in fairy tales can fly'). Note that these arbitrarily long distance dependencies are supposed to be an important source of linguistic *productivity* (Fodor, 2008, p. 103). A finite number of elements can, by centre embedding, generate an infinite number of sentences. Chapter 3 considers the central role of linguistic productivity to theories of neural architecture.

Only a sub-set of the CFGs are expected to provide decent representations of hierarchical sentences.[8] Consider figure 2.3 as an example.[9]
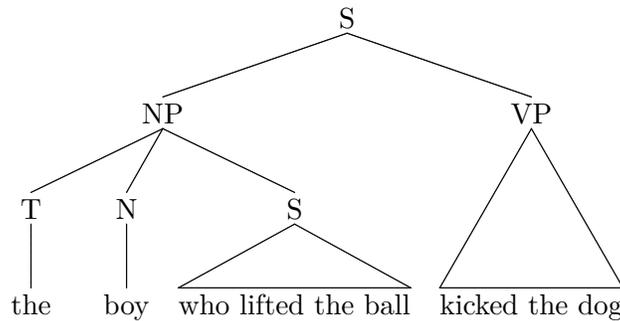
Figure 2.3: *A recursively centre-embedded sentence*

It is worthwhile, therefore, to make a distinction between context-free grammars that permit *self-embedding* (or *recursive centre embedding*) and those that do not. This notion depends on the notion of a *recursive element.*[10]

---

[8] It is important to realize that some context-free grammars do not permit hierarchical sentences. Trivial counter-examples are grammars that permit only strings of two atoms, such as

S $\Rightarrow$ AB
A $\Rightarrow$ the
B $\Rightarrow$ dog

Of course, such languages are suitably represented in simpler grammars.

[9] A triangle branching from a non-terminal symbol stands for collection of lower branches.

[10] It does not matter in what follows whether we actually represent embedded clauses as containing *sentences*. All that matters is that they are represented with *recursive elements*. A theorist could propose a separate grammatical category, *embedded-clause*, to account for such items. An embedded-clause would still be a

> A sentence in a context-free grammar has a recursive element iff the sentence is assigned a phrase-structure-tree such that some node on the tree corresponds to a grammatical category (or non-terminal symbol) which dominates a branch where the same non-terminal symbol appears (Chomsky and Miller, 1963, p. 289-290)

Once again, relative clauses provide the most plausible example, assuming we are willing to actually represent clauses as *embedded sentences*.

We can now define centre-embedding in terms of recursive elements. We simply need to distinguish trees with recursive elements appearing only on left-most-branches and right-most-branches from trees (Chomsky and Miller, 1963, p. 290).
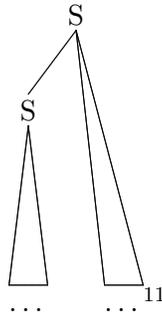

Figure 2.4: *A left-branching recursive element*
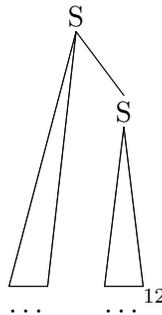
Figure 2.5: *A right-branching recursive element*

---

recursive element if it was partially composed of further clauses.

Yet even with a definition of self-embedding, there is still a sense in which *individual sequences* are not hierarchical. Any language containing a single sentence with a finite number of elements can be represented in a finite state language. The same is apparently true of languages in which the number of nested embeddings is limited to some integer n (Chomsky, 1957, p. 23). But even if English could be represented as a finite state language, Chomsky argues that it would be an inferior representation.

> We might arbitrarily decree that such processes of sentence formation in English as those we are discussing cannot be carried out more than n times, for some fixed n. This would of course make English a finite state language, as for example, would a limitation of English sentences to length of less than a million words. Such arbitrary limitations serve no useful purposes, however. The point is that there are processes of sentence formation that finite state grammars are intrinsically not equipped to handle (Chomsky, 1957, p. 23).

So even if English, or perhaps the portion of English containing the sentences that an individual could possibly utter, inscribe, or think in a human lifetime, were somehow to be represented in a finite state grammar, one could still foresee an argument to the effect that participation in English requires special sequencing powers. Fodor (1987) presented such an argument when he urged that even if language were not *productive*, or did not contain sentences of all finite lengths, then it would still be *systematic*. To say that a language is *systematic* is to say that '[the] ability to produce/understand some of the sentences is intrinsically connected to the ability to produce/understand many of the others' (Fodor, 1987, p.149). The requirements for understanding 'the dog that was fat bit the cat' are related to the requirements for understanding 'the cat that was fat bit the dog' (for a more systematic understanding of systematicity, see Hadley, 1994). Linguistic systematicity is just another point of departure for the argument that a sentence production and comprehension mechanism, if it is to be computationally[13] tractable, must somehow exploit the *composition* of the language, and by *composition* theorists have in mind *processes of sentence formation* at least as complex as those in a context-free grammar.

---

[13]We return to the notion of *computation* in chapter 3.

Unless otherwise specified, I will use the term *hierarchy* to refer to a language which requires, at a minimum, representation in a context-free grammar. I will not bother distinguishing the various ways in which the term *recursion* is used. It will do to take the *capacity for recursion* to require that the organism possesses certain computational mechanisms that are required for understanding hierarchical languages.

2.3 Hierarchical Sequences in Nature

We now have the resources to revisit the question as to whether naturally occurring sequences, outside the domain of language, should be considered as hierarchical. The answers to such questions are important, as they ground hypotheses about neural architecture. Recall Devlin's inference to the 'best' neural explanation. This inference involves drawing conclusions about neural processing by appealing to the supposed structural distinction between motor and linguistic sequencing. Of course, it does not *directly* follow that distinct neural systems are involved in these different forms of sequencing. It is logically possible, for instance, that motor sequencing mechanisms are more powerful than they need to be. At best, the structural differences make the existence of segregated sequencing systems *plausible*.

Philip Lieberman, contrary to Devlin, is happy to accept that motor sequences and linguistic sequences are hierarchically structured (Lieberman, 2006, 2007). His favoured illustration is the heel strike.[14] The strike exhibits selectional constraints, as it must occur at the proper time relative to the rest of the gait sequence. Furthermore, the precise moment of heel-strike varies across environmental conditions.

Motor control is also subject to temporal constraints. The articulation of a sentence requires careful management of abdominal muscles and the intercostals during expiration. Otherwise, the natural elastic collapse of the lungs would force air out of the airway without giving the speaker an opportunity to articulate speech sounds (Lieberman, 2006). The system needs to have access to the length of the sentence prior to the onset of phonation,

---

[14]The point in gait when the heel first touches ground (Lieberman, 2006)

if an individual is to produce a sentence normally (Lieberman, 2006). This implies that phonation also exhibits selectional constraints, as what comes next during speech must be constrained by the current state of expiration.

One might wonder whether these observations are sufficient for us to represent motor control as hierarchical. Such constraints are similar to the selectional restrictions on what kind of subject can given for the predicate 'green' (you can have a green cow but not a green idea etc.)[15]. One might be able to see how selectional constraints require grammatical categories (since a system needs to recognize subjects and predicates before it can place special restrictions on which nouns go with which predicates), and *perhaps* context-free grammars, with their non-terminal symbols, can be used to capture these grammatical categories in a theory of motor sequence composition. It remains unclear, however, whether such constraints alone imply long-distance dependency relations or recursive elements.

If *hierarchy* means *convenient representability* in a CFG, then the question comes down to whether we should posit structural categories corresponding to nonterminal symbols. If *hierarchy* means representation in a CFG with recursive centre-embedded elements, then the question comes down to whether Lieberman would represent motor sequences as possessing recursive elements. It seems Lieberman would resist both as he seems skeptical about the value to neuroscience of the '⇒' (Lieberman, 2006).

> It is an open question as to whether Chomsky's theories have any biologic valid-
> ity. Can we really be sure that the relative, prepositional, and adverbial clauses
> that occur in complex sentences are derived from embedded sentences? [...]
> In short, recursion as defined by Hauser, Chomsky, and Fitch (2002) follows
> from hypothetical syntactic structures, hypothetical insertions of sentences and
> phrase nodes, and hypothetical theory-specific rewriting operations on these
> insertions to form the sentences we actually here or read (Lieberman, 2006,
> p.359-360).[16]

---

[15]Resnik (1996)

[16]One might wonder here whether Lieberman is actually interested in representations of motor sequences or in structural representations of *mental representations that undergo computations to allow for movement.* For my purposes it does not really matter. We will see in sections 4 and 5 that Lieberman's evidence has

Lieberman seems to be suggesting that the '⇒' symbol is taken to reflect a computational re-writing operation that is of little or no relevance to actual neural processing. But recursive elements and the '⇒' aside, Lieberman seems unconvinced that we should provide linguistic sequencings with structural descriptions along the lines of noun phrases, verb phrases, and so on. He seems in no hurry to assign grammatical categories to motor sequences. This suggests that he is also wary of representing linguistic sequences with a context-free grammar.

Consequently, it is unclear whether Lieberman wants motor sequences to exhibit long distance dependencies. While Lieberman does want to claim that motor sequences are hier-archical (Lieberman, 2007), the existence of a long distance dependency seems to imply a typing of the sequence into grammatical categories. This is because long distance depen-dencies are long distance relationships between grammatical categories (Hochmann et al., 2008).[17]

There seems to be, then, two ways of reading Lieberman's views. First, it may be that Lieberman has a different notion of *hierarchical structure*. He has claimed that "[. . .] any set of ordered 'rules' inherently yields a hierarchical structure" (personal communication). If what Lieberman means by an 'ordered rules that yeild hierarchical structure' is that a well-formed sentence's composition can be described by the sequence of calls to well-formedness conditions, then it seems he is after the minimal notion of hierarchical structure already discussed, which applies even to grammars which do not include non-terminal symbols to serve as structural descriptions.

Perhaps Lieberman believes we *should* assign structural categories to motor and linguis-tic sequences, but rejects the structural descriptions of *noun phrase, verb phrase, and so on*. When Lieberman questions the *biologic validity* of these descriptions, he seems to be suggesting that our structural descriptions of natural sequences be *evaluated empirically*. Lieberman is comfortable with the grammatical category of *sentence*, for instance, because

---

consequences even for those who do not believe in compositional mental representations.

[17]One might worry about the potential for equivocation around the notion of *long-distanced dependency*. We have described long distance dependencies as long distance relationships between grammatical categories. This assumes, of course, a typing of sequences to categories. It does not assume, however, that the relationships have to hold between noun phrases and verb phrases.

the anatomical details of expiration demand that a system have access to the rough length of an utterance prior to phonation (Lieberman, 2006). Regardless of where Lieberman departs from his competitors, some common ground is available to all parties. The theorists we are concerned with agree that structural theories make at least some predictions in the domains of neuroscience and psychology. This common ground is the subject of chapter 3. My general concern in chapters 4 and 5 is whether the study of the brain poses problems for certain structural descriptions of natural language or thought.

Chapter 3

Structural Theory and Neural Mechanism

3.1 Introduction

In chapter 2, we outlined an inferential strategy that begins with a typing of colloquy in a formal language and ends with the postulation of mechanisms to account for an individual's participation in a language of the corresponding complexity. One might think this suspect. Colloquy and language (whatever that is) are products of a *population*, not an individual, and it is therefore unclear what the structure of natural language (if there is such a thing) entails about the neural requirements for an *individual's* participation in colloquy.

However, all of the theorists we are concerned with here do believe that theories of linguistic structure should make *some* predictions about the nature of individual brains, and the participatory requirements for colloquy. Their theories, therefore, are not insulated from empirical evidence, and may be confirmed or disconfirmed by scientific experiment.

One crucial difference between theorists, however, is precisely *what* languages are suitable prospects for revision. One must keep *English* apart from the mental languages which occupy the ontology of some theorists. Sections 3.2 and 3.3 will spell out why neither

language is insulated from theoretical reconstrual.

The claim that structural theories can be revised in light of empirical evidence is easy to defend from the purview of Cronbach and Meehl (1955). One needs only to assume that the structural features of a natural language, or of mental representations, are *hypothetical constructs*, whose validity must be made apparent by empirical inquiry. This view already plays a foundational role in much, if not all, of psycholinguistics. Consider, for instance, the claim that dysfluencies in speech sometimes reflect cognitive processes. Goldman-Eisler (1958) did not treat this claim as a mere methodological *assumption*. Henderson, Goldman-Eisler, and Skarbek (1966, p. 207) viewed the earlier work of Goldman-Eisler (1958, 1961) as a *defence* of the claim that one could use 'hesitation in speech' as 'an index of cognitive activity.' Later, Goldman-Eisler (1972) argued that we should use construct validity theory to evaluate the distinction between subordinate and co-ordinate clauses. In responding to a defence of the distinction, she writes

> The tool of pause measurement in spontaneous speech enables us to examine the validity of statements such as the above in terms of objective behaviour and psychological reality. If in the flow of spontaneous speech the transitions between its various constituent structures, between words, clauses — co-ordinate and subordinate, and of the latter, relative and others (adverbial) — as well as between sentences, if the transitions between these structures are of characteristic duration, then we can speak of their differential psychological reality and draw conclusions as to the degree of integration and independence of any of these units (Goldman-Eisler, 1972, p. 103).

Note that one can adopt this strategy whether they are interested in the structural features of *English* or *mental representations*. Pauses in speech, according to Goldman-Eisler (1972), track a structural distinction in language. That they do so is an empirical question. One can also see how, on a strong version of construct-validity theory, validating measures that track structural features actually improves our understanding of what syntax *is*.

Lieberman's argument for the reality of the *sentence* provides another example of how an hypothesis about neural architecture might inform structural theory. He argues that some

notion of *sentence* is required to understand neural processing in as much as systems need to 'know' something about the length of the utterance prior to phonation in order to produce a sentence without dysfluencies.[1]

We begin by teasing out some of the implications of Chomsky's modern usage of *recursion* in his account of the capacity for recursion. Afterwards we will consider the language of thought hypothesis, a theory of neural architecture which owes much to generative linguistics.

3.2 Recursion and Mechanism

When theories of linguistic structure inform theories of an individual's *ability* to participate in colloquy, they encroach into the domain of psychology, especially if they are intended to explain an individual's participatory *capacity* by making reference to a functional system. Consider their description of the narrow faculty of language (FLN), a (uniquely human) device supporting linguistic ability (and only linguistic ability).

> We assume, putting aside the precise mechanisms, that a key component of FLN is a computational system (narrow syntax) that generates internal representations and maps them into the sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the (formal) semantic system [. . .] (Hauser et al., 2002, p.3).

Whatever is meant by a 'system,' it is clear that Hauser, Chomsky, and Fitch wish to make a claim about the sequencing powers a brain would need to satisfy the demands of participation in colloquy. They manage to 'put aside' questions of *precise* mechanisms because they are operating at a higher level of abstraction. But their claims are not *divorced* from concerns regarding mechanism.

---

[1] Respiratory constraints, of course, do not licence an understanding of *sentences* as well-defined structural entities. A *sentence* for Lieberman might just be an utterance, or a turn at talk (Turnbull, 2003).

Hauser, Chomsky, and Fitch do more than merely ascribe an entire organism with sensory-motor and 'conceptual' abilities. They wish to say something about how theorists should, in functional terms, view the architecture of nervous systems. Indeed, the value in suggesting that FLN is a system performing computations on 'mental representations' which are then taken as input by phonological or '(formal) semantic systems' is precisely that it makes predictions about neural architecture. It entails, for example, that the behaviour of whatever realizes the '(formal) semantic system' must be sensitive to the behaviour of whatever realizes the FLN.[2] These claims must eventually be supported by theories regarding the neural supports of these various systems, and by evidence that these neural supports influence one another in the way that the functional account of Chomsky, Hauser, and Fitch predicts.

Perhaps the best illustration of structural representations generating empirical research is the comparative work inspired by Hauser, Chomsky, and Fitch (2002). The studies asked whether cotton-top tamarin monkeys (Fitch and Hauser, 2004) or songbirds (Gentner et al., 2006) possess the capacity to 'learn' hierarchical formal grammars. When studying whether animals might possess the recursive capacities in question, Fitch and Hauser (2004) chose the $A^n B^n$ grammar (which contains strings of the form $ab, aabb, aaabbb$, and so on) discussed in chapter 2.

> The $A^n B^n$ grammar requires additional computational machinery beyond a finite-state automaton. In psychological terms, it requires some way to recognize a correspondence between either the groups formed by the As and Bs (e.g., counting) or between specific As and corresponding Bs (e.g., long-distance dependencies). This [context-free grammar] thus provides the ideal grammar for the empirical issue addressed by this study by allowing us to focus on the generative power of the system without introducing extraneous performance variables

---

[2] One can speak about systems without suggesting they are modules of the Fodor (1983) variety. The systems may be unencapsulated or inaccessible (the system may draw upon information in addition to its input and information from within the system may be available for exploitation by other systems). They might not be neurally specific (they might be instantiated very differently across healthy individuals). They need not be genetically grounded. But Hauser, Chomsky, and Fitch are still saying something about the (functional) architecture of brains, and the mechanisms allowing for participation in colloquy.

(e.g., memory capacity or referentiality)(Fitch and Hauser, 2004, p.3).[3]

Chomsky's (1957) argument that English is not a finite-state language, taken alone, implies little about individual mechanism. The argument directly addresses only the complexity of *English*, not the demands on individual speakers. But research projects such as that of Fitch and Hauser (2004) are explicit about the implications of such arguments for individual brains. Humans have the neural machinery to learn hierarchical languages, and a goal of the Hauser et al. (2002) research programme is to judge whether the same computational powers were available to non-human ancestors.

Sponsors of structural theories share responsibility for the research their theories inspire. In chapters 4 and 5 it will be argued that some structural theories may run up against negative evidence. This class of theories will certainly include the likes of Devlin's (2006). Devlin believes that language is hierarchical, while motor sequences are not. Structural accounts will have to be reconciled with theories about how motor, linguistic, and cognitive sequencing powers are actually supported by the brain.

3.3 Hierarchy and the Language of Thought

One especially prominent class of theories that should be brought into the purview of basal ganglia research are those which posit compositional mental representations. The language of thought (or LOT) hypothesis is a theory of mental architecture which explains behaviour with reference to *structured mental representations* and *computational systems* that operate over those mental representations (Fodor, 1975, 2008).

The sentences of the language of thought (or sentences of Mentalese) must be both semantically and syntactically compositional. The expressive powers of a Mentalese sentence is determined by the semantic values of its atomic parts and its composition. A prevailing intuition, however, is that computational systems are directly sensitive to syntactic, but

---

[3] It is unimportant for our current purposes to evaluate the validity of the Fitch and Hauser (2004), study. This question is more carefully considered in Gentner et al. (2006), Hochmann et al. (2008), and Herbranson and Shimp (2008).

not semantic, features. This means that mental representations owe their causal powers to purely physical, syntactic features. For many theorists, compositional mental representations serve as causally efficacious contents for propositional attitudes (*beliefs that P, desires that P, and so on...*) and, more importantly, serve as the objects of mental processes. Consider the belief that 'granny left and auntie stayed' (Fodor, 2008).

> That the logical syntax of the thought is conjunctive (partially) determines, on the one hand, its truth-conditions and its behavior in inference and, on the other hand, its causal/computational role in mental processes. I think that this bringing of logic and logical syntax together with a theory of mental processes is the foundation of our cognitive science [...] (Fodor, 2008, p. 21).

It is popular to view the LOT hypothesis as an ally of belief-desire psychology. It provides causally efficacious propositional attitude *contents* for what seems to be a modest ontological price. But one does not *need* to advance a philosophical theory about propositional attitudes in order to posit a LOT. A vision scientist may draw upon the LOT hypothesis in advancing a theory of how the visual system performs its duties by performing computations on (structured) mental representations of environmental features. So while Fodor (1975, 2008) uses mental representations to account for propositional attitudes, we will assume no such account in what follows.[4]

However, if understanding an English sentence requires the tokening of a mental representation, then a LOT is convenient in its accounting for linguistic behaviour.[5] Our best representations of colloquy can be appropriated in descriptions of mental representation composition. Misunderstandings due to syntactic ambiguity in the English sentence *no trees have fallen over here* can be accounted for with reference to unique Mentalese sentences.

---

[4] See, for example, Carruthers (2006) for a very different, and in most respects non-fodorean, mental architecture which is still friendly to LOT.

[5] We should distinguish *accounting* from *explanation*. Note that *ease of accounting*, in the sense that I am using it, is not always a virtue for a theory. Religion may easily account for the existence of dinosaur bones without necessarily providing *explanation*.

(no trees have fallen) (over here).

(No trees have fallen over) (here).[6]

A LOT theorist could even go as far as to provide every grammatical category in English with a corresponding grammatical category in the LOT. The composition of the mental representation of 'the duck ate the elephant' might look a lot like figure 2.2 with different labels on the non-terminal symbols. Of course, the LOT hypothesis has to do more than just account for syntactic ambiguities. Many of the jobs previously associated with semantics (for example, Fodor, 1987) must also be accommodated by the LOT theorist. The syntax of Mentalese must be sufficient to distinguish the various readings of semantically ambiguous sentences (*I walked to the bank*) if computational systems are to have a way of distinguishing such readings (see, for example, Fodor, 2008, p. 58).

It is not obvious that Lieberman is at odds with *all* LOT theorists. Recall that LOT theorists will explain the comprehension of sentences by positing computational operations that are defined over mental representations. It seems that Lieberman will only disagree with LOT theorists who think these computational operations reflect the traditional theories in generative linguistics.[7]

> There is no reason to assume that neural physiology shifts to the serial algorithmic processes that, with little success, have marked linguistic studies for the past 50 years. Indeed, theoretical linguists know that they have failed at even being able to describe the grammar of any present human language using these algorithmic processes (Lieberman, 2006, p. 362).

Lieberman's argument (which we will discuss in chapter 4) will apply most directly to the LOT hypothesis, we will argue, if the sentences of Mentalese contain *recursive centre-embedded elements*. Mentalese, being just another language, is susceptible to formal representation. Some of the phrase trees of this language might contain recursive-elements and

---

[6] This example is due to Mary Shaw

[7] Thanks goes to Trey Boone, however, for the suggestion that Lieberman's argument might have implications for the composition of mental representations.

look like figure 2.3. Let us call the supposition of a language of thought with recursive centre embedding the LOT-RCE hypothesis.

A caveat is in order regarding those LOT-RCE sentences which are to represent propositions — and in so doing account for how we can understand an English sentence. Even if the representations of propositions are hierarchically complex, it does not follow that the mental representations that underlie motor sequences (assuming there are any) must be hierarchically complex. Those who believe in compositional mental representations need not believe in *the* language of thought, as if every computational system were sensitive to the same language. There is no requirement that every computational subsystem must be able to take as input any sentence of the LOT. My printer needs to accept instructions it receives from my motherboard, but information sent to the printer and my hard drive need not be in the same format. Even LOT theorists, therefore, can retain a structural demarcation in the LOT sentences representing movements and the sentences representing propositions.

If LOT theorists accept that natural languages contain grammatical sentences with an arbitrary number of embedded clauses or elements, then it should not be hard to see why LOT theorists will be tempted to adopt the LOT-RCE. The LOT theorist will be reluctant to allow natural language to exhibit a structural complexity that is not present in Mentalese, as every distinct reading of a sentence should correspond to a unique Mentalese expression. Otherwise *computational systems* could not distinguish various readings and, consequently, the behaviour of individuals would not be sensitive to such distinctions.

Furthermore, recursive centre-embedded elements are supposed to provide the LOT theorist with a chief advantage over one of its oldest competitors, radical associationism — the view that linguistic behaviour can be explained by reference to associative pairings between stimuli, responses, or ideas. Fodor (2008) still believes that the temporal-sensitivity of associative pairings constitutes a bane of associationist views.

> [. . .] Hume held that ideas became associated as a function of the temporal contiguity of their tokenings. [. . .] Likewise, according to Skinnerian theory, responses become conditioned to stimuli as a function of their temporal conti-

> guity to reinforcers. By contrast, Chomsky argued that the mind is sensitive to relations among interdependent elements of mental or linguistic representations that may be arbitrarily far apart (Fodor, 2008, p. 103).

Fodor, here, suggests positing arbitrarily long distanced relationships between *mental representations.* Such arbitrarily long distance relationships would seem to require, at a minimum, representation in a context-free grammar with recursive elements. Positing a LOT-RCE, then, is supposed to allow cognitive scientists to account for hierarchical productivity. Ascribing *composition* to mental representations allows the LOT theorist to account for the generative-novelty (or productivity) of linguistic behaviour while ascribing centre-embedded recursive elements allows the LOT theorist to account for productivity within embedded clauses.

> On the one hand, there is no reason why computational relations need to be contiguity-sensitive, so if mental processes are computational they can be sensitive to 'long distance' dependencies between the mental (/linguistic) objects over which they are defined. On the other hand, we know that computations can be implemented by causal mechanisms; that's what computers do for a living. [. . .] In short, a serious cognitive science requires a theory of mental processes that is compatible with the productivity of mental representations, and [the computational theory of mind] obliges by providing such a theory (Fodor, 2008, p. 105).

The question as to whether linguistic and motor sequencing are implemented in similar computational systems is especially pressing to the LOT-RCE theorist. This is because the LOT-RCE theorist *must* propose a theory where recursive elements directly fit into the causal explanation of linguistic behaviour. The whole point of positing composition in Mentalese sentences is that computations can be sensitive to syntactic composition. Positing a LOT-RCE, therefore, has consequences for one's computational theory of *sentence processing.* Failing to posit recursive elements in motor representations is just as significant. Such predictions are ripe for empirical evaluation.

It would be a great boon for LOT theorists if research into the basal ganglia could reveal something about the composition of mental representations. Even those who already adopt the LOT hypothesis have difficulty drawing disciplined inferences about the composition of Mentalese.

There are, of course, plenty of arguments of the following form.

> This organism, system, or neural population must be able to represent X in order to do Y.

This is precisely the kind of argument Fodor makes for the LOT hypothesis. The 'mind' needs to be "sensitive to 'long distance' dependencies between [. . .] mental (/linguistic) objects," (Fodor, 2008, p.105) in order for humans to participate in colloquy. But these arguments always rely on assumptions about what is possible with and without mental representations of a particular structure. Fodor and Pylyshyn (1988)[8] can be read as an argument that connectionist systems without *compositional* mental representations cannot account for systematicity (see 2.2) and thus cannot account for linguistic behaviour. Such assumptions, themselves, tend to pose difficult questions, and they are almost always difficult to evaluate empirically.[9] LOT theorists committed to a particular structural theory of LOT often find themselves implicitly relying on assumptions about the limits of a computational system without mental representations of a given composition. Perhaps this is why some computationally minded theorists see the LOT hypothesis as empirically vindicated, while others view it as lacking any kind of empirical support.

> This heavy commitment to a syntactical picture is often made willingly enough, but if it is made, it is solely on the basis of aprioristic reasoning, for so far as I can see, while there has been plenty of interesting speculation, there is almost no empirical evidence yet that tends to confirm any substantive hypothesis about

---

[8] Also see Fodor and McLaughlin (1990)

[9] But see Hadley (1994, 2004) for a more useful way of framing questions about the capacity of connectionist systems to account for *systematically*.

the nature of this supposed syntax of mental representation (Dennett, 1983, p.
213).

A related, and perhaps deeper, problem is that theorists avoid commitments about how
computational systems are implemented in the brain. When confronted with the question
as to how much brain is needed to realize a 'mental file' written in Mentalese, Fodor (2008)
is forthright. 'How big they have to be is anybody's guess; maybe as large as neural nets,
maybe as small as sub-neuronal molecules. Nobody knows, though many pretend to' (p.
94). Theorists interested in the details of how computational systems are implemented
are certainly refreshing. Eliasmith and colleagues, for instance, try to keep the number of
neurons required to realize their computational models in check (Stewart and Eliasmith,
forthcoming), and press for the creation of neurologically plausible computational models
(Eliasmith, 2007).

What is neat about Lieberman's argument is that the work seems relevant both to LOT and
to non-LOT theorists, and its importance does not depend upon contentious assumptions
about the limits of non-representational and non-LOT-representational systems. Instead it
takes the form of a kind of paradox. It places longstanding intuitions about the complex-
ity of motor sequencing against an account of the sequencing powers offered by cortical-
subcortical-cortical circuits. As evidence for the latter hypothesis mounts, theorists will be
obliged to accommodate it, and in doing so they are likely to make interesting concessions.

It must be made transparent that theorists who, like Fodor, operate at a 'high level of
abstraction' must nevertheless make *some* predictions if they claim to advance an empirical
hypothesis. One cannot, for example, adopt some eccentric vocabulary into one's theoretical
language, note that the vocabulary is not definable in the language of, say, biochemistry,
and claim to be operating at a higher level of abstraction. *Abstractions* are *abstractions from*
something, and abstractions must be integrated with lower levels of description, *at least in
some fashion.* If a theorist is providing an empirical theory, then the integration between
computational systems and their neural implementation forces certain kinds of predictions.
With respect to the structural typology of LOT-RCE theorists, this forces predictions about
the relationship between linguistic and motor sequences, on the one hand, and the mental
representations underlying linguistic and motor sequencing (assuming there are any) on the

other. Evidence from neuroscience *becomes* relevant to such theorists as soon as the right sort of (computational) description of cortical-subcortical-cortical circuits is in place (see 5.4.2).

Chapter 4

The Basal Ganglia and Sequencing

Central to Philip Lieberman's thesis is an hypothesis about the functional architecture of complex neural circuits involving the basal ganglia. These cortical-striatal-cortical circuits (or CSCs) project from the cortex, through a number of structures in the basal ganglia (including the striatum), to the thalamus, and back to the cortex. He argues that these circuits underwrite important sequencing abilities in the domains of cognition, motor control, and language. However, even a grossly simplified account of CSC functional architecture demands a cursory neuroanatomical discussion. I will consider the evidence suggesting cortical-subcortical-cortical involvement in motor, cognitive, and linguistic sequencing in 4.2 and 4.3.

4.1 Cortical-Striatal-Cortical Circuit Architecture

CSCs contain two neural pathways (see figure 4.2), which depart at the striatum. Excitation of neural populations in the direct pathway serve to relieve the thalamus from the inhibition provided by the interior globus pallidus (GPi) and substantia nigra reticulata (SNr) (Alexander and Crutcher, 1990). Since the thalamus has excitatory projections to the cortex, relieving it from inhibition can serve to promote cortical activity. Activation of the direct pathway is thought, ceteris paribus, to facilitate cortical activity (Alexander and Crutcher, 1990). The indirect pathway projects through a longer route from the striatum through the exterior globus pallidus (GPe), to the subthalamic nucleus (STN), and thence to the interior globus pallidus (GPi) and substantia nigra reticulata (SNr) (Alexander and Crutcher, 1990). This pathway seems to have the capacity to inhibit activity of the cortex (Alexander and Crutcher, 1990). The two pathways are often represented respectively as having the capacity to facilitate or inhibit 'movements initiated by the cortex' (Marsden and Obeso, 1994, p. 889).[1]

Dopamine plays a key role in this simplified model of CSC architecture. Delong (1990), for instance, used the influence of dopamine to explain a number of features of Parkinson's disease. Because the neural populations representing the direct and indirect pathways bear different *dopamine receptors*, dopamine from the SNc (substantia nigra pars compacta) can simultaneously promote the activity of the direct pathway and inhibit the activity of the indirect pathway. It follows that the Substantia nigra Pars Compacta, which has dopaminergic projections to the striatum, constrains and modulates the activity of Cortical-Striatal-Cortical circuits.

A classic feature of Parkinson's disease (or PD) is bradykinesia, a general slowness of voluntary movement.[2] This sometimes results in, for instance, loss of facial gestures (Chou and

---

[1] Basal ganglia architecture is, of course, more complicated than this suggests. I am ignoring, for instance, research which suggests an important pathway projecting from the motor cortex directly to the globus pallidus (Nambu et al., 2002).

[2] Understanding the motor difficulties involved in PD, and other motor disorders, requires employing much finer distinctions than are made here. Theorists need to distinguish, for example, the initiation of movement from the amplitude of movement, and *rigidity* (which implies muscle resistance to even passive
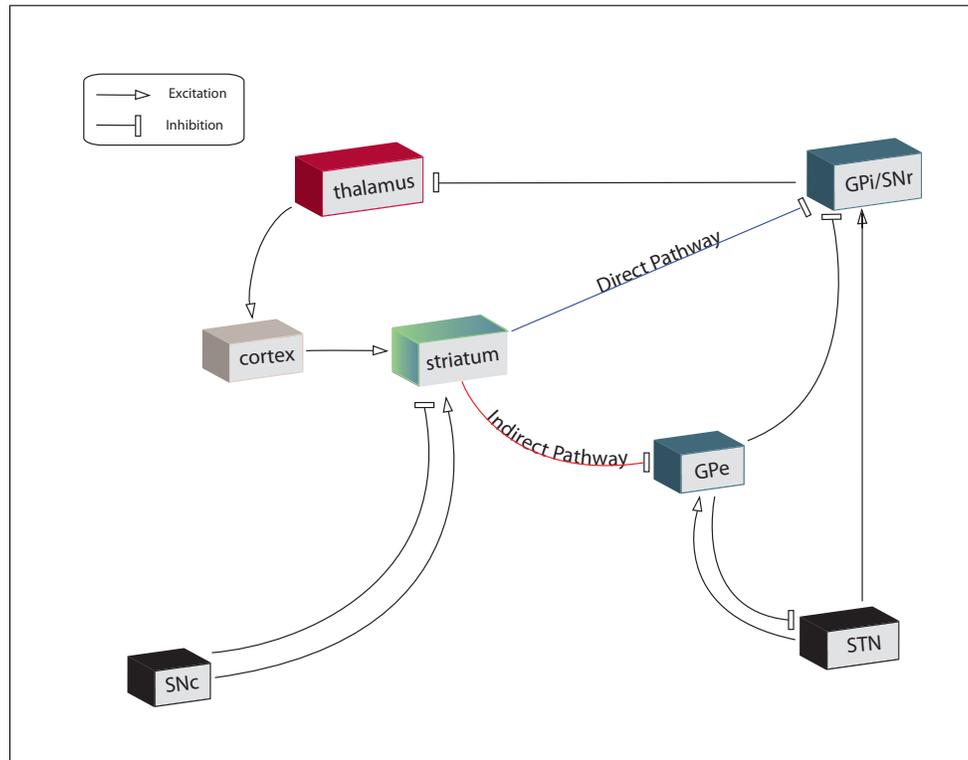
Figure 4.2: A simplified account of cortical-striatal-cortical circuit architecture (adapted from, Yin and Knowlton, 2006, p. 465)

Hurtig, 2005). Delong (1990) explained the lack of voluntary movement found in untreated Parkinson's disease by turning to this simplified model of cortical-subcortical-cortical pathways.

One of the prominent features of Parkinson's disease is massive cell death in the Substantia Nigra pars Compacta, and untreated Parkinson's patients have reduced dopamine levels in the striatum (Ebadi and Pfeiffer, 2005). The lack of dopamine in the striatum, according to this simplified model, would lead to an over-activation of the indirect pathway. The consequence would be a general inhibition of 'cortically initiated movements'. The model can also used to explain why *excess* dopamine levels in the striatum, brought on by L-dopa

---

movement) from *spasticity* (which implies more muscle resistance to high velocity movements than to low velocity movements) (Chou and Hurtig, 2005).

treatments for PD, could occasion *excess voluntary movement.*[3]

The past twenty years of re-
search have confirmed the im-
portance of the basal ganglia
to motor sequencing (for ex-
ample, Doyon et al., 2009).
They seem important for
motor execution, learning,
and adaptive motor control
(Doyon et al., 2009). Rats, for
example, perform a number
of grooming movements that,
when executed individually,
do not seem to significantly
involve the basal ganglia mo-
tor circuits.   However, rats
sometimes produce a series of
temporally ordered grooming
movements, or *chains.* Lesion
studies and single cell record-
ings have suggested that the
basal ganglia are crucial for

Figure 4.1: The basal ganglia (image from Siegel and Sapru, 2006, p. 214)

the production of movement chains, but not the individual movements (Aldridge et al.,
1993). I cannot here do justice to the colossal literature on the basal ganglia's role in motor
control.[4] It is worth, however, taking a moment to discuss voice-onset time (or VOT),
which Lieberman has used to measure motor-sequencing deficits in Parkinson's disease and

---

[3] I do not claim here that this simplified model is anywhere near complete. It might explain features of
bradykinesia, but it is less helpful, at least given what has been said here, in explaining why Parkinson's
disease is often accompanied by *tremor* (this point is thanks to Chou and Hurtig, 2005) which is, loosely
speaking, an uncontrollable shaking, especially in limbs and extremities (Chou and Hurtig, 2005).

[4] See Doyon et al. (2009) for a fuller discussion.

hypoxia.

When speakers produce stop consonants, they must close their lips obstructing the passing of air, and then open their mouths, allowing a burst of air to escape the mouth (Lieberman, 2006). The beginning of phonation requires manipulation of the laryngeal muscles after the opening of the mouth. However, distinguishing stop consonants such as [p] and [b] requires that the onset of phonation occur at a particular time (Lieberman, 2006). Producing the sound [b] requires phonation within 20 milliseconds from the burst of air, and [p] requires a longer delay (Lieberman, 2006). Theorists can measure overlap between the voice-onset-times of an individual's attempts at [p] and [b] sounds. The wrong sort of overlap[5] indicates a motor-sequencing deficit.

> One of the primary speech deficits of Broca's syndrome and compromised basal ganglia function is a breakdown in the sequencing of the motor commands necessary to produce stop consonants (Lieberman, 2006, p. 177).

People with Parkinson's disease or hypoxia (both of which influence the basal ganglia) often have unusual VOTs, suggesting similar motor sequencing difficulties in both disorders (Lieberman et al., 1992; Hochstadt et al., 2006; Lieberman et al., 2005, 1995).

4.2 Cognitive Sequencing

The notion of *executive function*, not surprisingly, writhes in confusion. As with most hypothetical constructs, psychologists have hoped to explicate this notion in the course of developing a valid measure of executive function. While no validated measure of executive function exists, theorists tend to view *the Wisconsin Card Sort* (or WCST) as relatively good at capturing a few important components of executive function (Ebadi and Pfeiffer, 2005).

---

[5] VOT overlap can only be taken to indicate a motor-sequencing deficit where the right confounds are controlled for. An inappropriate VOT of a [b] sound, if it is to indicate a motor-sequencing deficit, must not be due to, for instance, generally slower speech (Lieberman, 2006, p. 177-179).

Often discussion of executive function dissolves into discussion of *cognitive sets*. Possession of *cognitive set*, even on the most deflationary readings, implies a certain kind of strategic disposition to behave (see Flowers and Robertson, 1985, p. 517; Taylor and Saint-Cyr, 1995, p. 283; Ebadi and Pfeiffer, 2005, p. 350). These dispositions need to be shifted when changing environmental demands render a strategic disposition to behave inappropriate or unhelpful. Some theorists go further and identify set with a brain state.

> One kind of behavioural disturbance affecting cognitive function in Parkinson's disease is commonly described as a failure of mental 'set'. Set may be defined as a state of brain activity which predisposes a subject to respond in one way when several alternatives are available. Control of set involves both maintaining one predisposition or strategy against other competing possibilities (that is, the motor equivalent of attention), and also changing the strategy when circumstances change (Flowers and Robertson, 1985, p.517).

This says little about the constraints on what kinds of sets can be formulated, and does relatively little to clear up what a cognitive set *is*. It remains unclear whether cognitive sets are literally *sets of propositions* or, more accurately, *sets of mental representations* about the immediate environment. Those who are serious about describing 'sets' as '*cognitive* sets' may argue that sets can contain representations of anything representable by cognition. Theorists hope, perhaps fondly, that further testing will reveal answers to these questions. For now, theorists are left to cling to their best putative measures of cognitive set switching.

Consider, for example, the odd-man-out task (or OMO), a WSCT-like task that has been introduced for the purposes of studying Parkinson's disease (Flowers and Robertson, 1985). On the first trial, individuals are given a card with three pictures. The three items can be classified in two ways, and each classification leaves one item out as distinct. For example, there may be three letters on the card, with two items sharing the same shape/letter and a distinct pair sharing the same size. When asked which item is the odd one out, participants are forced to 'choose their own rule' for determining which item is the odd one out. They must determine the odd one out in virtue of, for example, letter type or size. The participants are then asked to use the same rule they had chosen for the next ten stimuli. Ten

trials later, they are asked to change the rule by which they chose the odd one out. After another ten cards, they are asked to go back to the original rule. OMO performance is impaired impaired in PD, and a similar set shifting difficulty may exist in hypoxia (Flowers and Robertson, 1985; Lieberman et al., 2005).

The WSCT differs from the OMO in the number of items to be matched, and the number of classifications that can be used to sort cards. It seems likely that both tests involve the following capacities, though perhaps to greater and less degrees.

(A) Formulate an appropriate set or strategic disposition to respond[6]

(B) Maintain an appropriate set or strategic disposition to respond

(C) Recognize negative feedback necessitating a change in set or strategic disposition

(D) Shift away from an inappropriate set or change strategic disposition in response to negative feedback[7]

Theorists have appealed to the influence of PD on cognition in their theorizing about the role of CSCs in cognition. Marsden and Obeso (1994) suggested that the basal ganglia might confer a sequencing power that allows for flexibility in both motor control and cognition.

> Most of the time [the basal ganglia] allow and help cortically determined movements to run smoothly. But on occasions, in special contexts, they respond to unusual circumstances to reorder the cortical control of movement (Marsden and Obeso, 1994, p. 889).
>
> [...] Perhaps the basal ganglia are an elaborate machine, within the overall frontal lobe distributed system, that allows routine thought and action, but

---

[6] It is still unclear regarding how many of these components of WSCT performance actually involve the basal ganglia (Monchi and Petrides, 2001; Monchi et al., 2004, 2007). CSC dysfunction, as seen in PD, has implications for much of the cortex. It is therefore possible that many aspects of executive function are due exclusively to the cortex — the prefrontal cortex in particular.

[7] This particular functional division of WSCT task is thanks to Monchi and Petrides (2001).

which responds to new circumstances to allow a change in direction of ideas and movement (Marsden and Obeso, 1994, p. 893).

This view entails the following predictions.

1. Those with Parkinson's should exhibit a 'cognitive-sequencing' deficit.

2. Individuals with damage to the basal ganglia, especially damage to motor-thalamic or pallidal lesions, should exhibit a 'cognitive-sequencing' deficit.

3. Healthy individuals should recruit the basal ganglia for the purposes of cognitive sequencing

As Lieberman (2006) points out, confirmation of prediction (1) is seen in the performance of Parkinson's patients on tests that measure the ability to shift cognitive sets (such as the Wisconsin card sorting task and Odd-Man-Out task)(Ebadi and Pfeiffer, 2005; Lezak et al., 2004). Confirmation of (2) might be seen in climbers on Mount Everest (Lieberman et al., 2005), where hypoxia targets the basal ganglia (Jeong et al., 2002; Lieberman, 2006). Evidence for (3) comes from imaging studies finding activation in the basal ganglia during a cognitive sequencing task (Monchi and Petrides, 2001; Monchi et al., 2004).[8]

4.3 Linguistic Sequencing

A number of studies have reported sentence comprehension difficulties in Parkinson's disease (Lieberman, 2006; Lieberman et al., 1990, 1992; Pickett, 1998; Grossman et al., 2000; Grossman, Lee, Morris, Stern, and Hurtig, 2002; Grossman, Zurif, Lee, Prather, Kalmanson, Stern, and Hurtig, 2002; Hochstadt et al., 2006). However, it is contentious whether

---

[8] While it remains plausible that cortical-striatal-cortical circuit dysfunction could be responsible for some of the cognitive deficits in Parkinson's disease, it must be noted that other explanations have been proposed. The cognitive deficits may be due to a general reduction of neurotransmitter levels in the prefrontal cortex, perhaps due to a loss of dopaminergic innervation from the mesolimbic-cortical pathway (Mattay et al., 2002).

these difficulties actually reflect a difficulty with *linguistic sequencing*. That is, it is unclear whether Parkinson's patients have difficulty exploiting *structural information* to apprehend a sentence. Grossman et al. (2002), for instance, proposed that the linguistic difficulties of PD are due to a problem with executive resources rather than a genuine sequencing deficit. Others have made use of the test of meaning from syntax (TMS) in their attempt to track a genuine linguistic-sequencing deficit in Parkinson's disease (Hochstadt et al., 2006).

One might have thought that if there are comprehension deficits in PD, then these deficits could simply be due to working memory limitations, as card sort tasks are thought to require working memory capacities to some degree (see, for example, Flowers and Robertson, 1985). Furthermore, working memory deficits, or at least some portion of them, are arguably not limitations in *sequencing*.

Suppose one is working with a computational model where working memory, or at least a portion of it, is a form of *memory storage* that, in principle, could be expanded without changing the computations performed by systems (Fodor and Pylyshyn, 1988). On such a view the sequencing powers of a system could be constrained by the complexity of the functions called. Some memory deficits, on such a view, need not be considered sequencing deficits at all. Consider a program which produces sentences of the $A^n B^n$ grammar in accordance with the rule S$\Rightarrow$ASB. The length of sentences which the system can produce will be limited by memory constraints. But it is not obvious that adding additional memory *increases* its sequencing powers (Fodor and Pylyshyn, 1988).

Whether a memory deficit qualifies as a *sequencing* deficit will come down to the competence/performance distinction imposed by the theorist. An intuitive approach is to argue that, where expandable memory systems can be distinguished from computation, the system's sequencing powers are determined by the latter (Fodor and Pylyshyn, 1988). A similar competence/performance distinction can be held in the case of motor control.

Of course, it is not obvious that *working memory* is partially constituted by an expandable memory system. Memory is not cleanly distinguished from computation in some architectures (Fodor and Pylyshyn, 1988). Furthermore, verbal working memory capacity is sometimes thought to contain an executive component (Baddeley and Hitch, 1974) which

can manipulate stored information. A sentence comprehension deficit due to an *executive* working memory deficit might still be considered a sequencing deficit.

Fortunately, the question as to whether the comprehension deficits in Parkinson's disease are genuine sequencing deficits is not only a matter of aprioristic typology. Theorists can introduce measures to clear up the nature of the PD comprehension deficits. The test of meaning from syntax, for example, is supposed to measure an individual's ability to derive the significance of a sentence from its scope and other grammatical properties alone (Pickett, 1998; Hochstadt et al., 2006). Participants are given a sentence and a set of images. Their task is to match up the sentence with the picture that best captures the sentences significance or meaning. Sentences varied across three crucial dimensions.

1. Sentences were either the active or passive voice ('the beagle kicked the dog', 'the beagle was kicked by the dog').

2. The sentence could be devoid of relative clauses, include a final relative clause (such as 'the dog is a beagle that stinks'), or a centre embedded clause ('the dog that stinks is a beagle').

3. The significance of sentences could be semantically constrained or unconstrained ('the dog bit the pear'/'the dog bit the lion').

It was important to vary sentences across active and passive voice, as many of the sentence comprehension difficulties could have been due solely to working-memory limitations, or perhaps some other cognitive but non-linguistic deficit. Hochstadt et al. (2006) did report that Parkinson's patients had much more difficulty understanding sentences of the passive voice, and this difficulty was well accounted for by alleged working-memory measures. This, of course, is consistent with the results of Grossman et al. (2002).

It was also important to vary items across the dimension of *semantic constraint* in order to tease out whether the difficulty with sentence comprehension actually reflected a difficulty with linguistic sequencing. Semantically constrained sentences, such as 'the cook who was fat pulled the box,' can be matched with the correct picture by appealing to background knowledge regarding cooks and boxes. One might match the sentence to a picture correctly

simply because boxes don't typically pull cooks. Matching unconstrained sentences, such as 'the cook who was fat pulls the king', arguably requires a degree of linguistic sequencing ability, as one needs to determine, with no semantic cueing, who was fat and who pulled whome. There is some evidence that this condition is important, as participants tended to be better at apprehending the complex semantically constrained sentences than apprehending unconstrained ones (Hochstadt et al., 2006).

The reason for varying items across clausal structure should be obvious, as theorists were interested in whether the participants had a specific difficulty with centre-embedded sentences. Participants did have more difficulty with the structurally complex sentences but, as Grossman et al. (2002) would presumably have predicted, many of these difficulties were well accounted for by working memory measures.

Of especial interest, however, were semantically unconstrained, centre-embedded, sentences of the active voice. Participants had significantly more difficulty with this class of sentences. Furthermore, one's score on working memory measures was not particularly good at predicting difficulty with centre-embedded sentences of this sort. In fact, the only significant predictor of difficulty with centre-embedded, active-voiced, semantically unconstrained sentences was difficulty with cognitive-set shifting, measured by the Odd man out task (Hochstadt et al., 2006).

The Hochstadt et al. (2006) results suggest that there may be more than memory limitations involved in the PD sentence comprehension deficits. Otherwise, we would expect working memory measures to be strongly related to each of the PD sentence comprehension deficits. It seems likely, given the OMO's emphasis on task switching, that some kind of set sequencing deficit underlies the PD comprehension deficits.

Note that even if this comprehension deficit can be considered a *sequencing* deficit, more evidence is required to consider it as a *linguistic*-sequencing deficit. It should be obvious, at this point, that I am not going to give a semantic account for *linguistic sequencing.* For now, we are drawing only on the limited evidence (and the intuition) that tests such as the TMS have something to do with linguistic sequencing. Nevertheless, there are some things that linguistic sequencing *cannot* be, at least insofar as we are avoiding contentious typological

commitments. Grossman and colleagues, presumably because of an implicit distinction between language and cognition, are free to distinguish *genuine* linguistic-sequencing deficits from cognitive impairments. LOT theorists, or anyone adopting mental representations for that matter, are free to distinguish *language systems* from *motor systems* on the basis of mental representations. The domain of a system, for instance, might be individuated by the representational content of the system's input (Fodor, 1975, 2008). Where these assumptions are unavailable, a *linguistic sequencing deficit* is simply an incapacity to meet colloquy's structural demands.

4.4 The Argument: A Provisional Formulation

Recall that the goal of this thesis is to explore how research into other sequencing domains could influence structural theories of language and motor control. We cannot provide a complete review of the empirical literature (See Lieberman (2006) for a fuller account of CSC architecture), and there can be no doubt that further research is needed.[9]. But the current evidence does invite speculation about the processes underwriting sequencing in these different domains. If the same processes underlie different domains of sequencing, then this fact must be rectified with an account of the sequencing requirements imposed by these domains. CSCs, according to one plausible view, are sufficiently powerful to satisfy at least some of the sequencing requirements of language, cognition, and motor control.

These kinds of theories could change our view on the sequencing requirements of motor control and language as whole. Theorists would need to consider whether or not CSCs could satisfy colloquy's more strenuous sequencing demands. If they can, then theorists should either grant that motor sequences and linguistic sequences impose similar individual requirements, or provide an evolutionary account according to which some motor control requirements came to be satisfied by more powerful sequencing machines than were necessary for the task.

Finally, theories about individual participatory requirements can influence structural theo-

---

[9] However, recent computational models of basal ganglia architecture are especially promising (**?**)

ries as a whole (*English*, *Dance*, *Mentalese*, and so on), especially if those structural theories predict a disparity in the mechanisms supporting motor and linguistic sequencing. We will return to this kind of inference in 4.4.1 and 4.4.2.

Readers may recognize that the argument of the last few paragraphs can also be run with 'cognitive sequencing' replacing 'motor sequencing.' But aside from the card sort tasks described in the previous section, it seems unclear what 'cognitive sequencing' *is*, how it could be related to 'linguistic sequencing,' and whether the two can even be distinguished. While the answers to these questions are beyond our current scope, they must be given only in the light of evidence.

The familiar response to this provisional formulation of the argument is to claim that participation in colloquy requires more robust sequencing powers than motor control, and that special linguistic mechanisms fulfill the sequencing requirements that CSCs cannot. One obvious possibility is that Broca's area supports such a system. Lieberman responds to this line of thought by questioning the presupposition that the linguistic deficits of Broca's aphasia are always produced by purely cortical damage. Lieberman rejects this premise (Lieberman, 2006, p. 162). But regardless of the strength of Lieberman's response, the structure of the dialectic is clear. The claim that Broca's area offers superior sequencing powers to the domain of language needs to be supported by the further hypothesis that its sequencing powers are *not* due, in part, to subcortical circuits. It may be, for instance, that Ullman (2006) is right in claiming that Broca's area possesses its own cortical-subcortical-cortical circuit.

It must be acknowledged that this gloss of the problem equivocates between two kinds of structural theories. The first are theories of *natural language* structure. We may, for example, wish to bring our structural representations of English and our structural representations of cognitive sequencing into line with one another. A very different kind of structural theory would be those of *mental representational* structure. The language of thought, as discussed in 3.3, presupposes a theory of how mental representations are compositionally structured. The two types of structural theories will have to confront this evidence independently.

4.4.1   Problems for the LOT hypothesis

From the purview of the LOT hypothesis CSCs will presumably be viewed as *computational systems*. Recall that compositional structure in mental representations allows neural systems to be sensitive to purely physical, syntactic, features of mental representations. This would allow them to perform computations over mental representations in an ontologically respectable manner.

A LOT theorist, then, might take the basal ganglia research to imply the following.

> Similar computations, perhaps the same computations, underlie cognitive and linguistic sequencing.

If true, this claim would provide a serious challenge to those positing recursively centre-embedded elements in mental representations. If CSCs support the syntactic processing of centre-embedded relative clauses, then the centre-embedded recursive elements of mental representations had better play a significant role in the computations performed by CSCs. Recall from 3.3 that the whole point of positing recursive elements is to help account for *productivity in relative clauses.* Sensitivity to recursive elements is supposed to allow computational systems to parse sequences with arbitrarily long embedded clauses (see section 3.3).[10] But if linguistic and cognitive sequencing are accomplished through similar computations, then LOT theorists must reconcile this fact with their structural representation of Mentalese.

There are ways for the LOT theorist to accommodate the apparent relationship between linguistic and cognitive sequencing, though these responses will not be insulated from revision. For example, one might propose a theory of set-switching where *cognitive sets* are compositional mental representations of propositions. On such a view, it would be no sur-

---

[10] Of course, memory constraints would prevent an individual from actually understanding a sentence that, written in 10 point font, would span five kilometers. If, however, one has a theory of neural architecture in which computations are clearly distinguished from a system's memory, then theorists can claim, without obvious incoherence, that such sentences *could* be understood by a system performing the same computations with better memory capacities (Fodor and Pylyshyn, 1988).

prise if representations within cognitive sets contained recursive elements. The propositions composing a set could bear this structure.

The problem with such a response is that it is viciously ad hoc. It is one thing to argue that representations need a certain kind of composition in order for behaviour to be explained, but this is legitimate only where compositional structure has been shown to aid in explanation. The problem with reflexively conceding hierarchicality to the representations underwriting the OMO task is that set sequencing implies a kind of behavioural flexibility of the sort that recursive elements cannot purchase. Furthermore, it is unclear how recursive elements make accounting for this phenomenon any easier. I will return to this argument in chapter 5.4.

Much the same can be said of motor control. If the evidence really suggests similar computations underlying motor and linguistic sequencing, then many LOT theorists will want to represent motor and cognitive sequences as highly complex, else they will have to give up on the LOT-RCE theorist's alleged advantage in explaining how humans deal with hierarchically structured sentences.

4.4.2   Problems for other structural theories

The basal ganglia research may press us to revise our structural descriptions of linguistic, motor, and cognitive sequences. I have already suggested that, if one posits a LOT, then theorists can try to accommodate productivity within relative clauses by revising their structural descriptions of Mentalese. This hypothesis seems to encounter anomalies (see chapter 5). However, even those who are not LOT theorists want their structural theories of colloquy to make predictions about neural architecture. A purported difference in structural complexity between motor and linguistic sequences has prompted Devlin (2006) to predict a similar disparity in motor sequencing systems and linguistic sequencing systems. This is in tension with the empirical claim that similar neural processes underwrite linguistic and motor sequencing.

Even if we avoid compositional mental representations, then, the neuroscientific research may have implications for our theories of natural language grammar. This might amount

to treating syntactic features as hypothetical constructs, to be revised in light of evidence, or it may simply be a case of *modus tollens.* If our predictions, which are largely informed by our structural theories of language structure, turn out to be false, then we must consider revising our structural theories. Consequently, it may turn out that we have to revise our views on English, or Mentalese, grammar in order to accommodate an apparent relationship between cognitive, linguistic, and motor sequencing.

Of course, Devlin (2006) could simply have been wrong to suggest that motor sequences lack hierarchical structure. It seems clear that Bloom (1992) thinks that Devlin was. But note that this concession *must* reflect badly on theorists who believe that intuitions about language structure bear remotely on individual participatory requirements. Devlin's (2006) appraisal *was* intuitive. *My* intuitions testify to the presence of long-distance dependencies in relative clauses, and they make no such testimony in the case of motor sequences. If a closer appreciation of the evidence suggests that motor sequences *are* hierarchical, then then we should be skeptical of uneducated intuitions about sequence structure.

Of course, revising one's structural representations of linguistic, motor, and cognitive sequencing constitutes only one kind of solution to this simple formulation of the problem. In chapter 5, I will outline some ways in which a theorist can respond to this apparently negative evidence. Some of these responses will attack Lieberman's empirical argument, pointing to methodological difficulties stemming from study design. In chapter 6 we will consider more radical revisions to our structural representations of language.

Chapter 5

Responses to the Provisional Formulation

In chapter 4 I provided a simple formulation of how one might treat the individual sequencing demands of colloquy as an hypothetical construct to be investigated. Representations of natural language which are not insulated from empirical evidence will have to be reconciled with what is known about the relationship between linguistic, cognitive, and motor sequencing.

For example, the following claims are in tension.

1. Sentences, but not motor sequences, are hierarchically complex (chapter 2).

2. Our representation of natural language structure bears directly on individual sequencing demands (chapter 3).

3. The same neural, and for some theorists computational, process underlies linguistic and motor sequencing (chapter 4).

The tension is this. *If similar processes underlie both motor and linguistic sequencing, then we must reconcile such facts with both our structural representations of these sequences and*

our theory of the demands such sequences impose. In the event that this tension becomes insupportable, some assumptions will need to be relaxed. Reconciliation may, for example, involve the concession that motor sequencing is more difficult than previously imagined, or it may involve the concession that motor sequencing systems are more powerful than they need to be. Relationships between linguistic and non-linguistic, 'cognitive,' forms of sequencing could force the same sorts of theoretical revisions. In this chapter, we consider some problems involved in using data from Parkinson's disease to study individual sequencing demands.

5.1 Difficulties with the Empirical Argument

5.1.1 Methodological concerns

The easiest way to respond to the argument presented in chapter 4 is to question its empirical premises. There are a number of places where the empirical argument is simply too speculative to force major theoretical revisions.

In chapter 1 I emphasized that we would not defend the construct validity of any tests referred to in this work. It is nevertheless implicit in my argument, and Lieberman's, that the Odd man out Task, Voice Onset Time, and the Test-of-Meaning-from-Syntax are good measures of cognitive, motor, and linguistic sequencing respectively. I take it that the TMS is in the most need of construct validation, and so this is an obvious place for dissenters to look for flaws in the empirical argument. Here I can only proceed as on the assumption that these measures tap into the properties they seem to, and that further research will confirm their validity.

It is worth mentioning the methodological difficulties that come with theorizing about healthy brain architecture by looking to individuals with Parkinson's disease, though these are precisely the kinds of confundities that construct validity theory supposes we can eventually overcome. Inferences from pathology are always perilous, partly because it is difficult to isolate precise locations of brain damage and partly because it is difficult to disambiguate effects directly caused by, for instance, cell death in the substantia nigra from effects due

to longer term *consequences* of such lesions.

Parkinson's disease is especially problematic because of the influence of CSC circuits on the cortex. The CSCs of individuals with Parkinson's disease are not obliterated. They are better characterized as *dysfunctional*. Marsden and Obeso (1994) noticed that performing targeted lesions which do functionally obliterate CSC pathways can *improve* the motor symptoms of Parkinson's disease. This suggests that the PD cortex is better off without the "'noisy' disturbed input of the sort likely to be present in Parkinsonism" (Marsden and Obeso, 1994, p. 886). Many of the disturbances evident in PD may be due to CSC dysfunction, rather than obliteration. PD patients should not be viewed as examples of otherwise healthy individuals who lack a basal ganglionic sequencing system. Basal ganglia dysfunction has consequences for much of the cortex. Dopamine loss in the substantia nigra pars compacta, alone, could potentially effect the entire brain, as this structure seems to have a modulatory influence over CSCs. More problematic still, many symptoms of PD may not be due to dysfunction in these circuits. Patients with PD may suffer from a general reduction of neurotransmitter levels in the prefrontal cortex. This may be due, in some part, to loss of dopaminergic innervation from the *mesolimbic-cortical pathway* (Mattay et al., 2002).

The difference between short term and long term consequences in PD, and the difference between damage and dysfunction in the basal ganglia, makes it important to distinguish between various *stages* of PD. A complete understanding of the cognitive and linguistic difficulties of Parkinson's disease, and their neural underpinnings, probably requires understanding PD as a developmental *process*.

Comorbidity provides another challenge for studies into the linguistic disturbances of PD. People with Parkinson's disease, for instance, are more likely to develop of a form of dementia (Ebadi and Pfeiffer, 2005). Studies on cognitive and linguistic deficits in Parkinson's disease attempt to screen for dementia in attempts to distinguish the cognitive and linguistic deficits associated with dementia from those associated with mild-moderate PD (see, for example Grossman et al., 2002).

Putative relationships between performance on cognitive, motor, and linguistic sequencing

must therefore be considered with especial caution, until such relationships are confirmed in healthy individuals. This suggests another needed line of empirical work. While the basal ganglia seem important to cognitive and motor sequencing in healthy individuals (Doyon et al., 2009; Monchi and Petrides, 2001; Monchi et al., 2004), more research is needed on the importance of the basal ganglia to the linguistic sequencing *of healthy brains.* Some evidence comes from PET studies implicating the basal ganglia in the comprehension of ambiguous sentences (Stowe et al., 2004). Also relevant is the work of (Grossman et al., 2003), who argued, using fMRI evidence, that their healthy participants, to a greater degree than their PD participants, recruited the striatum in the comprehension of a class of centre-embedded clauses.

5.1.2   The existence of a single sequencing engine

These general methodological concerns aside, one might worry that empirical evidence from basal ganglia research will influence our structural theories only if one very contentious hypothesis turns out to be true, namely the hypothesis that a single sequencing system underwrites different domains of sequencing. Of particular relevance here is a longstanding hypothesis regarding *segregation* between different CSCs (Alexander et al., 1986; Alexander and Crutcher, 1990; Middleton and Strick, 2000a).[1] At least five segregated CSCs have been proposed[2](see figure 5.1).

This might help explain the following observation. Although individuals with Parkinson's disease tend to have difficulties with motor (Ebadi and Pfeiffer, 2005), linguistic (Hochstadt et al., 2006), and cognitive sequencing (for example, Flowers and Robertson, 1985), measures of motor sequencing seem to be poor at predicting difficulty with particular kinds of sentences[3] (Hochstadt et al., 2006). Part of the answer may be that the basal ganglia dysfunction of Parkinson's disease affects a number of CSCs, including both motor and cog-

---

[1] I do not take functional and anatomical segregation of circuits to imply that segregated circuits constitute *modules* in any interesting sense. However, it will not effect my argument if we take modularity in the basal ganglia to be a realistic possibility.

[2] Alexander et al. (1986)

[3] I am referring to centre-embedded, semantically constrained, relative clauses of the active voice (see 4.3).

nitive sequencers. However, the inability to predict linguistic sequencing difficulties from motor sequencing measures suggests that these forms of sequencing may be supported by segregated CSCs.

> The differential roles of impaired set-shifting and articulatory rehearsal in PD sentence comprehension deficits are supported by the absence of significant correlation between OMO performance and mean VOT separation. This is consistent with their being supported by segregated cortico-striato-cortical circuits that will not necessarily be affected together in a given PD patient (Hochstadt et al., 2006, p. 253).
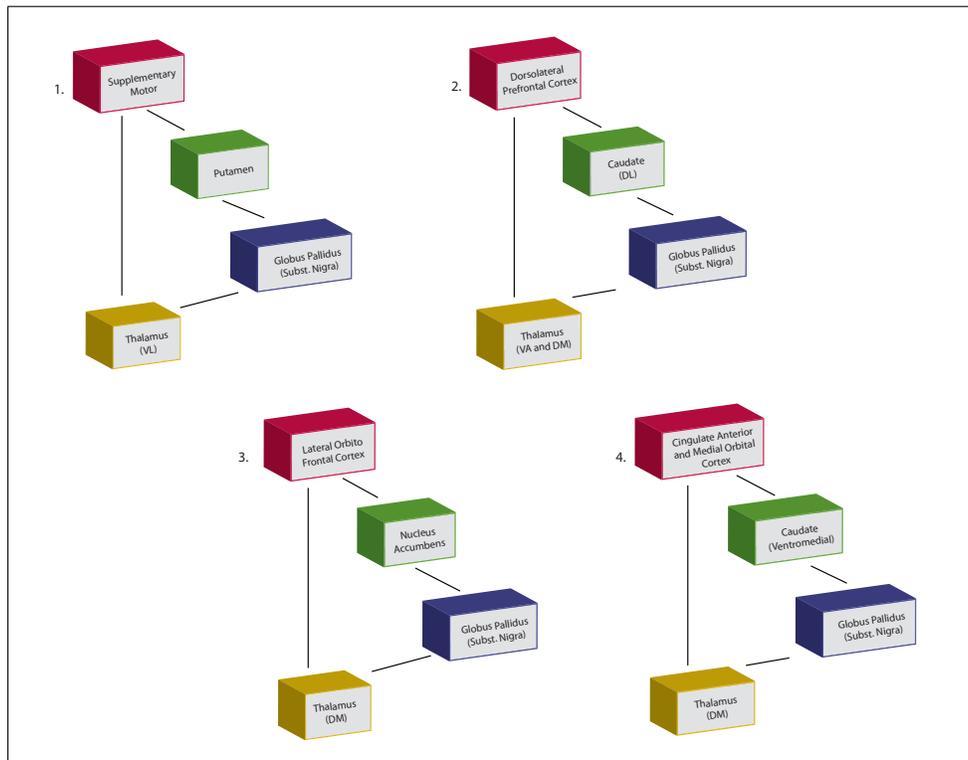

Figure 5.1: Segregated cortical-striatal-cortical circuits (adapted from, Maia et al., 1999)

It will require much more research to demonstrate that a single system underlies linguistic and cognitive sequencing. One needs to consider, for instance, the arguments of Ullman

(2006) for the existence of a segregated CSC involving Broca's area.[4]

However, the argument in chapter 4 can even be run *without* the assumption that a *single* sequencing system performs cognitive and linguistic sequencing. If CSCs are so segregated from one another functionally, then one must ask whether various CSCs owe their functional differences to their respective *cortical projections*, or to actual differences in CSC architecture or operations. If one adopts the former view, then it will still be possible that segregated CSCs perform the same operations on *different cortical input*.[5] This idea seems to have struck Alexander et al. (1986) as plausible.

> Because of the parallel nature of the basal ganglia-thalamocortical circuits and the apparent uniformity of synaptic organization at corresponding levels of these functionally segregated pathways [...] it would seem likely that similar neuronal operations are performed at comparable stages of each of the five proposed circuits (Alexander et al., 1986, p.361).[6]

The idea that the functioning of particular CSCs should reflect the functions associated with their particular cortical targets survives some scrutiny. Where some cortical area and some neural population within the GPi (or substantia nigra pars reticulate) belong to the same CSC, the functional significance of the neural population seems to bear some relationship with the functional significance of the cortical area. Middleton and Strick (2000b) argued for this claim by appealing to tracer studies and cell recordings. They also predicted that lesions to basal ganglia neural populations belonging to a CSC involved in X should have similar consequences to lesions of cortical areas belonging to the same circuit. 'One prediction from our results is that a lesion or disturbance of a particular subcortical loop should produce a behavioural disturbance that resembles the disturbance seen after damage to the

---

[4] For an objection to this argument see Lieberman (2006, p. 174)

[5] 'input' is a metaphor here. I do not wish to presuppose a LOT in this discussion. I leave it for non-LOT theorists to translate my claim into their idiom of choice, and for LOT theorists to argue that attempting such a translation is misguided.

[6] Note that the existence of some anatomical overlap between CSCs would not necessarily make this 'segregation hypothesis' false. While *extensive* overlap would presumably be a problem, the circuits might exhibit some interesting kind of functional segregation despite anatomical overlap.

cortical area subserved by that loop' (Middleton and Strick, 2000b, p.243-244). It may be a consequence of basal ganglia organization, then, that lesions to very specific neural populations in the interior globus pallidus may produce disturbances that have *more* in common with a cortical lesion than with lesions to contiguous GPi neural populations. This hypothesis is particularly plausible if Lieberman is right that the evolution of human CSCs was largely a story of co-opting the sequencing powers of motor circuits for other forms of sequencing (Lieberman, 2006).

It follows that claims about about the recursive elements involved in sentences (or their mental representations) may still be problematic, *even if distinct systems underlie cognitive, motor, and linguistic sequencing.* Various forms of sequencing would still be accomplished by the employment of similar, or even the same, neural processes. The arguments in sections 4.4 and 4.5 could, therefore, be run with the claim that similar neural processes underlie motor, linguistic, and cognitive sequencing.

This would be an interesting outcome, as some have taken Lieberman's argument to require a single general sequencing system. This interpretation is, no doubt, encouraged by Lieberman's depiction of the basal ganglia as a 'sequencing engine' (Lieberman, 2006), which suggests a unitary entity. His appreciation of the complexity of motor control provides further encouragement. 'Brain mechanisms adapted to handle the complex sequential operations necessary for speech production would have no difficulty in handling the comparatively simple problems of syntax' (Lieberman, 2000, p. 107-108). Furthermore, Lieberman markets himself as an opponent of modularity theories generally (Lieberman et al., 1992; Lieberman, 2006). It is therefore important to distinguish this unitary sequencing system hypothesis, whether Lieberman believes it or not, from his attack on recursive centre-embedding.

5.2 Competing Theories

Without a doubt, more evidence will be needed to put some of Lieberman's opponents to rest. Unfortunately, some *a priori* work may be needed to sort out who, exactly, Lieberman's competitors are. Grossman et al. (2002) is one such example, as he seems to argue that the 'linguistic' difficulties in Parkinson's disease are actually non-linguistic, cognitive,

difficulties in sentence comprehension. It is unclear, however, that Grossman et al. (2002) and Lieberman (2006) are working with the same understanding of *syntax*. Grossman et al. (2002) invoke the following as a foundational in their introduction.

> Sentence comprehension is a complex process that involves at least knowledge of the syntactic rules that govern the long-distance syntactic relationships between words and phrases in a sentence, as well as the processes needed to implement this knowledge in a fashion that efficiently supports rapid day-to-day communication (Grossman et al., 2002, p. 604).

It is not obvious what Grossman et al. (2002) have in mind when they talk about 'rules that govern the long-distance syntactic relationships between words and phrases in a sentence.' If by 'long distance relationships' they mean long-distance relationships between grammatical categories that are generated by recursively embedded elements, then they may be assuming a structural representation that Lieberman denies (see chapter 2). More importantly, deficits which are clearly non-syntactic for Grossman et al. (2002), may qualify as 'syntactic' for Lieberman. Recall that Lieberman (2000, 2006) is comfortable with modeling a portion of human linguistic sequencing after *cognitive set-switching*, a putative component of executive function. There is a sense in which (Lieberman, 2006) would be happy with the view that the linguistic deficits of Parkinson's disease are not *syntactic*, on a certain understanding of the word. But he might well argue that humans don't need syntax of this sort: cognitive sequencing will do. Many of the disagreements between Grossman and Lieberman, then, may be due to different background assumptions.

Grossman and Lieberman even seem to hypothesize similar neural mechanisms in accounting for the sentence comprehension errors in Parkinson's disease. Grossman et al. (2002), for example, suggests that difficulty in comprehending the sentence 'the boy that the girl chased is friendly' in a Parkinson's disease patient might be due, in part, to a problem in *inhibiting* a possible (but ultimately incorrect) reading of the sentence.

> Inhibition is necessary to block canonical grammatical role assignment in sentences with atypical word orders (in this example, blocking interpretation of

> the head noun phrase 'The boy' as the subject of the verb 'chased') (Grossman
> et al., 2002, p. 604)

Depending on what is being inhibited, such a process could amount to cognitive set-switching. This bears striking similarities to Lieberman's account. Recall the suggestion from Hochstadt et al. (2006) that a centre-embedded sentence "may require the listener to switch away from the 'canonical' subject-verb-object order of English sentences at the boundary between the main and the embedded clause" (Hochstadt et al., 2006, p. 245). It may be that Grossman and Lieberman differ primarily in their fundamental assumptions about language, and not in the general sort of explanations they would like to offer.[7]

5.3 Objections to Construct Validity Theory

In chapter 1 I distinguished two claims that are entailed by construct validity theory (Cronbach and Meehl, 1955).

(a) Whether a test X is a valid measure of theoretical construct C is an empirical question which can be evaluated by further research.

(b) Empirical research will allow theorists to formulate more precise explications for theoretical constructs which already have representatives in common discourse (this includes folk-psychological constructs such as *mind, belief, desire, and so on).*

(a) is crucial for evaluating the empirical argument presented here. If we are to be moved by Lieberman's arguments, then we must eventually become confident with our relevant measures of cognitive and linguistic sequencing. It is not obvious that the odd-man-out task measures cognitive sequencing. It is also not obvious that the TMS really measures a capacity to understand centre-embedded relative clauses from structural information alone. The hope is that such claims can be supported by further empirical research.

---

[7] Lieberman's and Grossman's accounts do differ in some details. For Lieberman, the rules of syntax are acquired by general cognitive processes, and it is unclear that Grossman shares this view.

The importance of (b) turns on how serious we are about viewing cognitive sequencing as genuinely *cognitive.* If theorists are comfortable viewing sets in non-mentalistic terms, then (b) may not be important to the project.

One might wish to object to construct validity theory, perhaps by advocating a theory of language which is inconsistent with (b). It is easy to see, for instance, that claim (b) is in tension with the (wittgensteinian) intuition that regular speakers of the language already understand the vocabulary of *cognition* and *mental* as well as anyone possibly can. Understanding the term may be guaranteed by the capacity to participate in the right sort of practices.

A theorist might also complain that construct validity is circular, as it requires that the notions of *meaning* and *construct* are also constructs. Since the notion of a *construct* is still poorly understood, one might come to question its theorists' record theories record for actually clarifying its cental concepts.

I do not wish to formulate such objections in any detail here. I point out only that, while some might attack construct validity theory, my competitors probably do not have this option. The LOT theorist will rely on construct validity theory to provide measures which reflect the computations relevant to mental processes (see section 3.3 and Thompson, 2010a). Those LOT theorists, such as Fodor (2008), who think that having a belief that P[8] amounts to some computational relationship between an organism and its mental contents, need some way of empirically tracking the computations that correspond to beliefs. They cannot simply stare at the brain until the dawn of an appropriate computational theory. Consider, more generally, theorists who would be happy to claim that a survey designed to study beliefs about abortion *actually measure causally efficacious entities.* Such claims obviously need to be made even where theorists are unable to spell out how *beliefs* and *desires* differ in computational terms. Construct validity theory, at least of the Cronbach and Meehl (1955) variety, arguably excuses theorists from such a burden, at least tentatively.

---

[8] It is a curious thing that the LOT hypothesis demands that mental contents be compositional, and yet theorists are still tempted to represent them by P, which conventionally refers to an atomic element. It would be more appropriate to describe propositional attitude contents with a metalogical variable, such as 'α.'

5.4 Accommodating the Evidence

5.4.1 Could RCE help account for set-shifting?

Not likely. Rather than denying the hierarchical structure of linguistic sequences, LOT theorists might prefer to posit centre-embedded recursive elements in the representations of motor and cognitive sequences. This avenue of attack owes some of its plausibility to the confusion surrounding the notion of a *mental set*. I do not know, for instance, if we are supposed to view mental sets as mental representations, or sets of mental representations composed of recursive elements.

The specific proposal I am concerned with here holds that CSCs realize computational systems that owe their sequencing powers primarily to the fact that their computations are sensitive to (centre-embedded) recursive elements.

While I have nothing to say against such proposals in principle, they do not *resolve* the anomalies presented in chapter 4, but merely speak to them. Suppose that very similar neural processes underlie sequencing in a variety of domains. An admixture of recursive elements in cognitive and motor representations might seem to reconcile our structural representations with our architectural hypotheses. But this ad hoc move fails to afford the theorist any additional explanatory power. Cognitive sequencing systems account for a kind of behavioural flexibility that recursion[9] alone simply cannot purchase.

Recursive centre-embedding purchases linguistic productivity, specifically productivity within multi-clause sentences, as in 'John believes that (she believes that)^n the sky is falling'. Productivity, of course, is a very particular aspect of linguistic novelty. Competent speakers seem to have a capacity to produce and comprehend, with relative ease, sentences which have never been uttered. Even Chomsky (1964, p. 59) recognized that this form of novelty is

---

[9] Recall that we distinguish between recursively centre-embedded elements and *recursive functions*. The former are artifacts of a relatively narrow class of formal languages, while the latter reflect a class of functions. Languages can be defined with recursive rules that do not contain any recursive elements. Even if we forget about context-free grammars all together, some could still represent basal ganglia functioning in terms of recursive functions.

distinct from the 'rule-changing creativity',[10] which pervades common discourse. Consider the following example of wit from John Austin.

> 'There are burgundy wines and burgundy-type wines. This, I should say, is a burgundy-type-type wine'(John Austin)

The capacity to continue on with '... or a burgundy-type-type-type wine' guarantees neither wit nor novelty (Jennings and Thompson, submitted for publication). One is more inclined to say that the wit of such utterances decreases monotonically with further iterations. No one ever claimed that productivity — or, for that matter, systematicity — could substitute for genuine novelty. In fact, there is reason to think that no recursive definition of well-formedness can ever encompass the class of genuinely novel productions, particularly if the definition appeals to a well-defined stock of atoms (Jennings et al., 2010). If we represent idioms as semantically atomic, then we introduce new atoms all the time.[11] Consider,

1. He's a diamond in the rough.

2. He's a rhinestone in the rough (Dorothy Parker).

There is no direct evidence that CSCs are involved in the introduction of semantic atoms or even wit generally. However, individuals must modify failing strategies in response to changing environmental demands, and CSCs seem to be involved at least upon receiving negative environmental feedback (Monchi and Petrides, 2001; Monchi et al., 2004, 2007). I take it that this is just the kind of ability that makes LOT theorists worry that cognitive processes are *computationally untractable* (Fodor, 2001; Carruthers, 2006). Freely representing motor and cognitive sequences in terms of recursive elements is, in principle,

---

[10]He might better have said 'constraint-challenging creativity.' Rules belong to formal languages - constraints to natural biological systems.

[11]The LOT theorist can make this line of argument more difficult by conceding an open-class of atoms for English, but denying humans the power to introduce new atoms in Mentalese. The mental representation corresponding to an idiom, on this view, must be a Mentalese sentence with well-defined atomic components — though I doubt anyone has a clear picture as to what the composition of *that* sentence would be.

a legitimate response to the anomalies presented in chapter 4. But theorists need to spell out what these representations would contribute to a theory.

5.4.2 The (in)capacity of the evidence to bring theorists along

So far I have suggested that - in order to mediate the gross theoretical disagreement between theorists - we take colloquy as the crucial explanandum of language-science, and use construct validity theory to winkle out the individual sequencing requirements. CSCs are interesting because they seem to support sequencing in a variety of domains. The hope is that further investigation will force theoretical concessions. How much is conceded, however, turns on the details of future research. The basal ganglia are surely not, for example, the source of *all* sequencing in healthy humans (Marsden and Obeso, 1994). It is also likely that the systems underlying motor, linguistic, and cognitive sequencing differ in a multitude of ways.

However, some theorists may simply be *too* insulated from empirical evidence to be brought anywhere by these sorts of observations. Theorists may simply look at relationships between motor, linguistic, and cognitive sequencing, consider the hypothesis that the basal ganglia are involved in all three capacities, and and conclude that, no matter how the research turns out, there will be plenty of room to fit theory to evidence without the loss of cherished assumptions.

So far, I have been distinguishing systems functionally, and have left theorists to disagree about how we specify a system's *function*. It follows that some theorists may be operating at a sufficiently 'high level of abstraction' (see 3.3) that they can re-draw the lines of which neural populations constitute a system, and in doing so accommodate almost any neuroscientific evidence. Perhaps the basal ganglia do not constitute a system, but serve as *component parts* of segregated motor, linguistic, and cognitive systems. On such a view, the language system might support hierarchical sequencing (at *some* level of functional abstraction) even if the motor system did not. Furthermore, PD could disrupt both forms of sequencing by causing dysfunction in crucial neural supports.

For this line of research to have the sort of impact I have in mind it requires something

such as the following; *Cortical-striatal-cortical circuits constitute systems (independently of their cortical targets) which, in virtue of their neural architecture, offer certain kinds of sequencing powers.* In order to force concessions from some LOT theorists, it may have to be added that these circuits constitute *computational* systems. Viewing cortical-striatal-cortical circuits as systems seems to be common practice, and surely cannot be objected to on aprioristic grounds. It should not be a big stretch for LOT theorists to represent them as *computational systems.* Such representations already seem to be present in the literature (see, for example, Ullman, 2006).

I take it that anything learned about the sequencing powers of *this* sort of system would be of interest, even if a specialized 'language-CSC' possessed unique sequencing powers. It would be particularly interesting if some feature of the architecture of CSCs made them helpful for solving specific linguistic sequencing problems. Such a finding would be crucial, as theorists would be prompted immediately to ask how much could be accomplished by such a system, and when (and whence) it could have become available to our ancestors. We return to this question in chapter 6.

LOT theorists must also pay attention to relationships between motor, cognitive, and linguistic sequencing *measures*, even if they completely disregard neuroscience as irrelevant. LOT theorists do not understand the *composition* of Mentalese any better than Linguists can be expected to understand the *composition* of a natural language. They rely on some form of construct validity theory to have these terms explicated through empirical inquiry. Such theorists cannot consistently adopt such a view while deeming the relationships between putative measures of linguistic, motor, and cognitive sequencing to be irrelevant. If they rely on construct validity theory then (trivially) they need to explain observations generated by tests which purport to measure properties of interest. In the process of appealing to empirical evidence in the resolution of foundational disagreements about colloquy structure, I doubt whether we can afford theorists any more *insulation* from data than I already allow.

Chapter 6

Implications for the Evolution of Language

Should theorists wish to revise their representations of language structure (whether it be *English, French,* or *Mentalese*), it would be useful to consider in what pickle their revision would leave relative clauses. On Chomsky's 1963 picture, centre-embedded relative clauses are a mark of natural language complexity. If we lose our enthusiasm for representing language as containing recursive elements, or we admit that hierarchical structure is common to a variety of naturally occurring, non-linguistic, sequences, then we might rightly wonder whether relative clauses are at all interesting. If, however, we are interested in appealing to something such as set-shifting in our model of sentence apprehension, then theories about the *origin* of relative clauses may become increasingly interesting. I provide the following as a conjecture about how relative clauses might evolve in a population of organisms with rudimentary set-switching powers, but without any innate propensity to learn hierarchically complex grammars.[1]

The story —which is little more than a story at this point— begins with the axiom that in biology, variation, most especially incidental variation, drives taxonomic change. Differences that have little taxonomic significance, can eventually form the basis of speciation. The

---

[1] This conjecture was originally formulated in Jennings et al. (2010) and developed in Thompson (2010b).

axiom will be just as secure in a theory of natural language evolution. Steven Pinker, a proponent of the language of thought, mental modules, and universal grammar, recognizes it as much as anyone else.

> The second component of language differentiation is a source of variation. Some person, somewhere, must begin to speak differently from the neighbors, and the innovation must spread and catch on like a contagious disease until it becomes epidemic, at which point children perpetuate it. [. . .] Moreover, one change in a language can cause an imbalance that can trigger a cascade of other changes elsewhere, like falling dominoes. (Pinker, 1994, p. 245-246).

Pinker also notes how incidental errors in transmission account for the word *orange*, which owes its existence to incidental[2] misapprehensions (*naranja* (spanish) ⇒ *a norange* ⇒ *an orange*)(Pinker, 1994, p. 245-246).

It must be emphasized that members of a population need not be conscious of language change. Linguistic innovations are introduced locally and propagate, whether the innovations are recognized by the individuals of the population or not. Pinker's analogy between linguistic engendering and the embellishment and 'reworking' of jokes (Pinker, 1994, p. 246) is therefore misleading. One might also argue that Pinker has missed the most telling examples. More interesting are syntactic misconstruals with the propensity to go unnoticed (Jennings, 2005).

While it is uncontentious that linguistic variation drives taxonomic change, an important caveat is required. Pinker is referring to natural languages such as *English*. He is *not* referring to Mentalese. Natural languages can change, for LOT theorists, without a change to the composition of LOT. As English changes, minds can simply represent, or more accurately re-represent, novel English constructions in the more powerful language of thought. This is presumably why Pinker believes that linguistic evolution is not particularly Darwinian.

> Many linguistic innovations are not like random mutation, drift, erosion, or

---

[2] By *incidental* I do not mean *random*. It may be that such errors are to be expected given our neural architecture.

borrowing. They are more likely legends or jokes that are embellished or proved or reworked with each retelling. That is why, although grammars change quickly through history, they do not degenerate, for reanalysis is an inexhaustible source of new complexity (Pinker, 1994, p.246).

The degree to which linguistic engendering reflects organic evolution is not important for our current discussion, as long as our understanding of linguistic engendering is sufficiently *enlightened* by Darwinian evolution that we are not put off by the taxonomic departure between ancestor and descendent. Such taxonomic departure is necessary if complex languages evolved from comparatively simple ones.

With this in mind, one can ask where we should expect set-shifting to be required in colloquy. One class of especially promising candidates are the *shifts in conversation* that pervade colloquy (Thompson, 2010b).

1. T: i'd like to get a case of blue

2. C: are yuh over nineteen? duh yuh have any id?

3. T: no

4. C: then i can't sell yuh any beer (Turnbull, 2003, p. 101).[3]

C's response to T's utterance in line 1 is not an answer, but a request for clarification that effectively delays C's answer until line 4. Such shifts are obviously rife in conversation. In fact, such shifts may be necessitated by structural features on the scale of linguistic interactions that are on par with the sentence-level features of *noun* and *verb*. Depending on how one carves up the structural features of conversation, such shifts may even generate long distance dependencies, such as the request/denial pair found in lines 1 and 4 (Thompson, 2010b). Similar pairs can be found following *questions, offers, parting words, appeals for*

---

[3] For our purposes here, these transcripts have been grossly simplified. Most notably, rising intonation markers have been substituted with question marks in some places. This is dangerous, as the marking of questions is a theoretical decision. This simplified adaptation is from Thompson (2010b).

justification, and accusations (Turnbull, 2003). As in the above example, these demarche-pairs are sometimes separated by long distances in colloquy. The contents of such gaps themselves tend to contain pairs (like lines 2 and 3), which are sometimes called *insertion sequences* (Turnbull, 2003).

I will not argue that such shifts in colloquy actually constitute long distance dependencies, and I certainly do not have the resources to argue that participants rely on cognitive-set shifting operations to navigate such interactions. A cursory look at actual conversation, however, is suggestive. Participants clearly have the ability to shift the direction of colloquy and then return to where they left off.

1. A: i'm just checking i was gonna ask you i think i did though (0.5 seconds)

2. did i ever get casablanca back from you?

[. . . 14 lines deleted . . .]

17. B: yeah i know i totally forgot about that

18. A: oh that's okay [. . .](Turnbull, 2003, p. 200)

It is tempting to think that partaking in an insertion sequence frequently requires a change of set. Jennings et al. (2010) pointed that such insertions could resemble complete sentences.

A: Senator Vandalprone. . .

B: Who after all does pay our salaries. . .

C: Pittance though it might be. . .

A: has asked whether we might put up some foreign guests for the weekend (Jennings et al., 2010, p. 53).

These *linguistic interruptions* could even occur within a single, prolonged, utterance.

> That one over there
>
> Not that one, more to the left
>
> The dog, I mean, not the handler.
>
> not the pug (that's hardly a dog at all)
>
> the Labrador
>
> Anyway, I saw her with bloody what's-his-face again last night (Jennings et al., 2010, p. 53).

This conjecture, dubbed the I⇒E conjecture (Thompson, 2010b), is that this species of interrupted colloquy, on the scale of linguistic interaction, could be cognate with taxonomically remote multi-clause sentences.

> The following conversation, which took place between the two friends in the pump-room one morning, after an acquaintance of eight or nine days, is given as a specimen of their very warm attachment [. . .] (Austen, 1870).

Of course, one cannot rule out the I⇒E conjecture solely on the grounds that relative clauses are taxonomically remote from interruptions in colloquy. The tasks are obviously different. An entire sentence, or at least some of them, can (arguably) be held in working memory. But remembering an entire conversation, even a short one, is virtually impossible, as evidenced by the very need for detailed transcripts. This is, of course, not a problem for the I⇒E conjecture, which assumes the taxonomic departure of colloquy and relative clauses.

A few caveats are in order. As I have already suggested, the capacity for set-shifting may be required to participate in interrupted colloquy. However, it should be noted that I am uncommitted to the language of *interruption* specifically. While it is fine to speak loosely about 'colloquy's direction' being tentatively interrupted, for instance, theorists will eventually have to say what this means, or replace the notion entirely.[4]

---

[4] I read Turnbull (2003), for example, as providing a much more detailed (but contentious) account of 'colloquy's direction'.

Nor am I suggesting that we 'analyze' relative clauses in terms of interruptions of colloquy. The I⇒E conjecture only proposes an evolutionary story of relative clause origins that (a) makes no reference to an innate capacity for learning grammars with recursive centre embedding and yet (b) endorses the claim that there is something structurally interesting about relative clauses.

A few distinctions are required to make the I⇒E conjecture plausible. I have made the modest, virtually incontrovertible, assumption that relative clauses have ancestors which were not relative clauses, and the two may be taxonomically remote.[5] Thinking about colloquy biologically also requires that we be aware of the variation between relative clauses in *actual conversation*, because such variation is a potential source of change. We may also wish to distinguish those multi-clause structures relying on well established patterns of conversations or, as Pawley and Syder (2000) call them 'linguistic crutches,' from those that do not. Pawley and Syder (2000) present two kinds of linguistic crutch. 'The first is a stock of familiar multi-clause syntactic frames that are memorized independently of lexical content. The second is a store of familiar expressions, consisting of fixed phrases and of more-complex speech formulas, some spanning two or more clauses, whose lexical content is partly fixed and partly variable' (Pawley and Syder, 2000, p. 195). In natural conversation, utterances exploiting such crutches are more common and, unlike their more novel brethren, are far less likely to include dysfluencies such as pauses, 'umms', 'ahhs', and so on (Pawley and Syder, 2000).

It is unclear how employing such crutches would change the neural demands on individual participants. Crutch-or-no, the good Saussurean linguist might represent a sentence as a relative clause — peripheral clumsiness being a mere product of performance error and non-linguistic memory constraints. For most LOT-RCE theorists, understanding a relative clause will ultimately require mentally representing it in a Mentalese sentence with recursive elements, even if the employment of a crutch involves a different computational process. A shortsighted LOT-RCE theorist, then, might disregard such conversational patterns as

---

[5] The only alternative would be an extreme universal grammar hypothesis that argued for the appearance of relative clauses spontaneously, presumably after a genetic change. Surely the recent attacks on universal grammar (Christiansen and Chater, 2008), and on linguistic universals (Evans and Levinson, 2009), are sufficient grounds for casting out such an hypothesis.

uninteresting, since the end product of sentence comprehension must be the same regardless.

But such conversational patterns are inevitably crucial to language change, even on cartoonish reifications of the competence performance distinction in neural architecture. At the very least, producing a sentence fluently or not will have consequences for linguistic engendering, even for someone such as Steven Pinker.

For Lieberman the employment of crutches could be similarly crucial. Relying on such prostheses could allow brains to cope with relative clauses by appealing to mechanisms which, through a process of automatization, have developed to generate these linguistic patterns efficiently. Their development may be a good example of language evolving to fit the brain. Producing relative clauses outside the norm, on the other hand, might require special cognitive resources, and rely heavily on a set-shifting-like capacity.

This is all somewhat tentative, and not yet ready to be cached out as an hypothesis. The studies implicating set-shifting-like operations in relative clause apprehension tend to make use of highly contrived stimuli, which are composed from a pre-defined set of nouns and verbs (*the boy chased the cow, the girl kicked the box* (Hochstadt et al., 2006)). One cannot, then, use the relationship between specific difficulties on the TMS and difficulties on the OMO to argue that novel embedded constructions are involved in set shifting. It is too unclear what the TMS stimuli resemble in a typical conversation.

The I⇒E conjecture essentially posits a transition from interruptions to anticipated interruptions, where the latter more readily fits under the taxonomy of centre-embedding. It must also be emphasized that one needs an argument in order to claim that a distinction between interruptions and anticipated interruptions is significant to relative clause evolution As far as I know, this argument is yet to be made. It is true that this distinction is of *potential* significance in the evolution of relative clauses, but this is just a consequence of population thinking more generally.

6.1 Measuring Anticipation of an Interruption

No one should persist in the illusion that one can simply 'whip up' a research project to study whether general sequencing powers bestowed by ancestral CSCs allow for participation in both interrupted colloquy and a subset of multi-clausal sentences, and thus facilitate the transition towards complex colloquy. It is not even clear how one could acquire the appropriate measures. The promise that such measures might be discovered in the future gives little comfort: the promise might well be an empty one. Nevertheless, this *is* an empirical question, and the reliance on some form of construct validity theory seems inescapable. Lieberman's (2006) discussion of *cognitive respiration* is telling.

> One point needs emphasis because it appears to have been overlooked by many linguists: the respiratory maneuvers that a speaker uses during normal discourse reveal both the *psychological 'reality' of a sentence and the planning involved in discourse.* The data show that a speaker has in his or her mind [*sic*] blocked out the general framework of a sentence before uttering a single sound (Lieberman, 2006, p. 326).

Lieberman is referring to data on respiratory constraints on speech. The production of an utterance requires control over the natural elastic collapse of the lungs, and this control requires an anticipation into the length of an utterance prior to production. Lieberman goes on to note that adults who run out of air provide acoustic cues that a sentence is not over (Lieberman, 2006, p. 325). As Lieberman (2006) seems to suggest, these cues may reflect the neural processing demands of the utterance, and this might reflect what kinds of linguistic crutches are being employed. This is consistent with the work of Goldman-Eisler (Goldman-Eisler, 1958, 1961, 1972) who argued that pauses in speech hint at underlying processing. Pawley and Syder (2000, p. 172-172) employ a similar distinction between planned and unplanned pauses in speech, which are marked by intonational cues. So while it will no doubt be difficult to ascertain the role of CSCs in actual colloquy, there is some reason to think that a research project could craft a useful measure of an anticipated interruption on the basis of conversational markers and dysfluencies in speech.

Of course, anyone who believes in cognitive set-shifting, *including a LOT-RCE theorist*, is likely to accept that cognitive set-shifting is involved in some interruptive colloquy. Sentences and conversations may need to be interrupted given changing environmental demands. The I⇒E conjecture presumes a set of neurological systems that can accommodate both unanticipated interruptions *and a set of systems which can accommodate anticipated interruptions.* Anticipated interruptions in the form of subordinate clauses could never propagate through a community lacking the required neural supports.

The data provided by Lieberman contribute to this story a theory about the neural capacities which would allow such an evolutionary process to occur. Furthermore, this theory does not assume that relative clauses are represented as containing recursively embedded elements. It is plausible (though unconfirmed) that a capacity for set-shifting could permit interruptions. It is also plausible that similar set-shifting capacities support, though perhaps not in sufficiently distant ancestors, apprehension of those especially demanding multi-clause constructions (such as those semantically unconstrained items). Cortical-subcortical-cortical circuits might even play a role in the appropriation[6] of linguistic crutches. For Lieberman, the acquisition of 'syntactic rules,' which I take to include syntactic constraints and conversational patterns, will be the end-product of a process of automatization — a learning process involving the basal ganglia (Lieberman, 2006).

What lends relative clauses their interest, according to the I⇒E conjecture, is that they are a hallmark of our neural capacity to participate in structured colloquy generally. If true, multi-clause sentences are an incidental consequence of participatory requirements for structured colloquy in a way that parallels the way that speech anatomy is ultimately an incidental consequence of bipedalism.

6.2 Ancestral Sequencing Powers

Currently, the I⇒E conjecture *is barely* conjecture. A disciplined approach to testing whether such a proposal *might possibly be true* would require previous knowledge about

---

[6] I say that these patterns of conversation are 'appropriated,' and not 'acquired,' because I want to resist the idea that colloquy's participants must *internalize* some feature of the population.

the sequencing powers of ancestral nervous systems. Even if we accept that the basal ganglia of recent humans constitute a sequencing engine which permits motor, linguistic, and cognitive sequencing, it does not immediately follow that ancestral basal ganglia had similar linguistic sequencing powers. While ancestral humans no doubt had basal ganglia (which even frogs have), it does not follow that ancestral humans possessed the recent human 'sequencing engine.' It may be that a modern linguistic environment (perhaps even one with inscription) is required for cortical-subcortical-cortical circuits to have linguistic sequencing capacities.

But this does not mean theorizing about ancestral sequencing powers is impossible. Our ancestors did not possess precisely the same sequencing systems allowing modern humans to participate efficiently in English relative clauses. Evidently they had *some* sequencing powers. Ancestors may, however, have been able to (or been required to) draw upon a set-shifting-like capacity in their colloquy (or, perhaps, pre-colloquy), especially if this set-shifting-like capacity is granted by the excitatory and inhibitory powers of the direct and indirect cortical-subcortical-cortical pathways.

In fact, Lieberman (2006) already has a well developed hypothesis about the evolution of linguistic sequencing systems. It is the principal subject of his book. He suspects that cortical-subcortical-cortical circuits initially suited to motor sequencing were co-opted for the purposes of cognitive and linguistic sequencing.

The exciting work done on the so called 'language gene,' FOXP2, merits at least a cursory discussion here. As Lieberman (2006) notes, it is grossly misleading to consider FOXP2 a 'language gene,' as it is actually a regulatory gene constraining the expression of other genes, some of which are involved in the production of motor control structures (Lieberman, 2006, p. 218). However, research on the gene retains its capacity to excite, especially for those who believe that cortical-subcortical-cortical circuits are important to language and cognition. FOXP2 is expressed in the basal ganglia (Lai et al., 2003), and a mutation in FOXP2 seems to be responsible for a linguistic and motor sequencing disorder (Vargha-Khadem et al., 1995). Furthermore, the gene seems to have emerged in the last 200,000 years, possibly from positive selection (Enard et al., 2002).

6.3 Independent Support for the I⇒E Conjecture?

Other theories of relative clause origins already exist. Heine and Kuteva (2006) approach the problem indirectly, by providing a theory about the grammaticalization of subordinators. Their proposal is that subordinators (in at least one class of languages) likely evolved from vocabulary marking interrogatives (*who, what, when, where, how, which, etc.*). This proposal, which stems from a study of European languages, obviously has more evidential support than the I⇒E conjecture. One project, then, might be to tentatively evaluate the I⇒E conjecture by seeing how it fits with more established proposals.

On the face of things, the I⇒E conjecture and the proposals of Heine and Kuteva (2006) are at odds. Tentative interruptions in colloquy seem to have little in common with interrogatives. But whether the I⇒E conjecture is actually a competitor to Heine and Kuteva (2006), will turn on the role of interrogatives in natural conversation. The class of interrogatives do, of course, overlap with the class of interrupted linguistic sequences (Thompson, 2010b). Interrogatives are often used to shift the direction of colloquy.

1. A: went to see eraser the other night

2. B: went and saw *what*?

3. A: eraser (emphasis added, Turnbull, 2003, p.167)

Interrogative markers, of course, also appear in non-questions. One example is in an appeal for support.

1. A: there are things that you know anyone of us could do if we had to take out someone's appendix out we probably could figure it out (0.6 second pause) you know? it's not that hard (.) the thing would look like an appendix and (.) would look like a skinny little wormy thing and we'd (.) cut it off (.) you know

2. B: yeah i guess so

3. A: *you know what i mean?*

4. B: *yeah (.) i know what you mean*

5. A: like there's (.) i mean you wouldn't want to try to perform heart surgery or something (emphasis added, Turnbull, 2003, p.159).

Of course, the class of interrogatives and the class of interrupted colloquies are not co-extensive. Student-teacher interactions often, perhaps almost always, begin with an interrogative along the lines of 'may I ask you a question'.[7] So it remains possible that Heine and Kuteva (2006) are correct while the I⇒E conjecture is false.

Another avenue to pursue evidence for the I⇒E conjecture is opened by developmental psychology.[8] Young children also show set-shifting difficulty (Zelazo et al., 2003; Zelazo, 2006), and developmental psychologists have expressed interest in the relationship between executive function and social interaction (Lewis and Carpendale, 2009). Studying how children acquire the capacity to participate in insertion sequences may shed light on the plausibility of the I⇒E conjecture (Thompson, 2010b).[9]

Multi-clause sentences owe much of their previous theoretical interest to the formal complexity of CFGs and the representation of such sentences as containing centre-embedded recursive elements. Such claims about the complexity of linguistic sequences bear upon theories of neural architecture, regardless of whether it is *English* or *Mentalese* (or both) that exhibits hierarchical structure. Notions of hierarchical structure which do not make specific reference to CFGs do less to demarcate the structural complexity of linguistic and motor sequences. They also do less to isolate the study of motor control from the study of language. But if construct validity theory demands that our structural descriptions of sequences in nature be open to empirical revision, then similarities in how the brain accom-

---

[7] See Turnbull's (2003) discussion of *pre-sequences*.

[8] Thanks goes to William Turnbull for this suggestion.

[9] If it turns out that the linguistic difficulties due to Parkinson's disease are the result of a *memory* impairment rather than an impairment in set-shifting, the I⇒E conjecture could still be true. The conjecture only assumes that similar neural resources are required for the participation in structured colloquy, such as colloquy with insertion sequences, and relative clauses. Even if a particular form of working memory underlies both acts, this capacity is essential to understanding linguistic interaction (this point is thanks to Turnbull, personal communication).

plishes motor and linguistic sequencing may prompt us to revise our theories of syntax. I have emphasized in chapter 5 that this revision is by no means *necessary*, given the limited evidence presented here. But even if we drop the sharp demarcation in motor, linguistic, and cognitive sequences, then multi-clause sentences could retain some of their theoretical interest. They may be marked by a more general participatory capacity.

Reference List

Aldridge, J. W., K. Berridge, M. Herman, and L. Zimmer (1993). Neuronal coding of serial order: Syntax of grooming in the neostriatum. *Psychological Science, 4*(6), 391–395.

Alexander, G. and M. Crutcher (1990). Functional architecture of basal ganglia circuits: Neural substrates of parallel processing. *Trends in Neuroscience, 13*(7), 266–271.

Alexander, G. E., M. R. DeLong, and P. L. Strick (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9*, 357–381.

Austen, J. (1870). *Northanger abbey.* London: Spottiswoode and Co.

Baddeley, A. D. and G. Hitch (1974). Working memory. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*, Volume 8. NY: Academic Press Inc.

Bickerton, D. (2009). Recursion: Core of complexity or artifact of analysis? In T. Givon and M. Shibatani (Eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution.* Amsterdam: John Benjamins Publishing Company.

Bloom, P. (1992). Review of lieberman, p. Uniquely human: The evolution of speech, thought, and selfless behaviour. *Language, 68*(2), 328–387.

Carruthers, P. (2006). *The architecture of the mind.* Oxford University Press.

Chomsky, N. (1957). *Syntactic structures.* The Hague, The Netherlands: Mounton and Co., Publishers.

Chomsky, N. (1964). Current issues in linguistic theory. In J. A. Fodor and J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Chomsky, N. and G. A. Miller (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology.* New York and London: John Wiley and Sons, Inc.

Chou, K. L. and H. I. Hurtig (2005). Classical motor features of Parkinson's disease. In M. E. Pfeiffer and R. F (Eds.), *Parkinson's disease*, pp. 171–182. USA: CRC Press.

Christiansen, M. H. and N. Chater (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*, 489–558.

Cowley, S. (2002). Why brains matter: An integrational perspective on 'the symbolic species'. *Language Sciences, 24.*

Cowley, S. (2004). Simulating others: The basis of human cognition? *Language Sciences, 26*(3), 273–299.

Cronbach, L. J. and P. E. Meehl (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 287–302.

Delong, M. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neuroscience, 13*(7), 281–285.

Dennett, D. (1983). *Styles of mental representation.* London: Aristotelian Society.

Devlin, J. (2006). Are we dancing apes? *Science, 314*(10).

Doyon, J., P. Bellec, R. Amsel, V. Penhune, O. Monchi, J. Carrier, S. Lehéricy, and H. Benali. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioral Brain Research, 199*(1), 61–75.

Ebadi, M. S. and R. Pfeiffer (Eds.) (2005). *Parkinson's disease.* CRC press.

Eliasmith, C. (2007). How to build a brain: From function to implementation. *Synthese, 153*(3), 373–388.

Enard, W., M. Przeworski, S. E. Fisher, C. Lai, V.Wiebe, T. Kitano, A. P. Monaco, and S. Pääbo (2002). Molecular evolution of foxp2, a gene involved in speech and language. *Nature, 418.*, 869–872.

Evans, N. and S. C. Levinson (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429–492.

Fitch, W. and M. Hauser (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303.*

Fitch, W. T., M. D. Hauser, and N. Chomsky (2005). The evolution of the language faculty: Clarifications and implications. *Cognition, 97*, 179–210.

Flowers, K. A. and C. Robertson (1985). The effect of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neuroscience, and Psychiatry, 48*, 517–529.

Fodor, H. (2008). *Lot 2: The language of thought revisited.* Oxford, England: Oxford University Press.

Fodor, J. (1975). *The language of thought.* New York: Crowell.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology.* MIT press.

Fodor, J. (1987). *Psychosemantics.* Cambridge: MIT press/Bradford Books.

Fodor, J. (2001). *The mind does'nt work that way: The scope and limits of computational psychology.* MIT press.

Fodor, J. and Z. W. Pylyshyn (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1-2), 3–71.

Fodor, J. A. and B. P. McLaughlin (1990). Connectionism and the problem of systematicity: Why smolensky's solution doesn't work. *Cognition, 35*, 183–204.

Gentner, T., K. Fenn, D. Margoliash, and H. Nusbaum (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(27), 1204–1207.

Goldman-Eisler, F. (1958). The predictability of words in context and the length of pauses in speech. *Language and Speech, 1*(3), 226–231.

Goldman-Eisler, F. (1961). The distribution of pause durations in speech. *Language and Speech, 4*(4), 232–237.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech, 15*(2), 103–113.

Grossman, M., A. Cooke, C. DeVita, C. Lee, D. Alsop, J. Detre, J. Gee, W. Chen, M. B. Stern, and H. I. Hurtig (2003). Grammatical and resource components of sentence processing in Parkinson's disease: An fMRI study. *Neurology, 60*, 775–781.

Grossman, M., J. Kalmanson, N. Bernhardt, J. Morris, M. B. Stern, and H. I. Hurtig (2000). Cognitive resource limitations during sentence comprehension in Parkinson's disease. *Brain and Language, 73*, 1–16.

Grossman, M., C. Lee, J. Morris, M. B. Stern, and H. I. Hurtig (2002). Assessing resource demands during sentence processing in Parkinson's disease. *Brain and Language, 80*, 603–616.

Grossman, M., E. B. Zurif, C. Lee, P. Prather, J. Kalmanson, M. B. Stern, and H. I. Hurtig (2002). Information processing speed and sentence comprehension in Parkinson's disease. *Neuropsychology, 16*, 174–181.

Hadley, R. (1994). Systematicity in connectionist language learning. *Mind and Language., 9*(3).

Hadley, R. (2004). On the proper treatment of semantic systematicity. *Minds and Machines, 14*, 145–172.

Hauser, M. D., N. Chomsky, and W. T. Fitch (2002). the faculty of language, what is it, who has it, and how did it evolve? *Neuroscience, 298*(22), 1569–1679.

Heine, B. and T. Kuteva (2006). *The changing languages of europe.* Oxford University Press.

Henderson, A., F. Goldman-Eisler, and A. Skarbek (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech, 9*, 207–216.

Herbranson, W. and C. Shimp (2008). Artificial grammar learning in pigeons. *Learning and Behavior, 26*(2), 116–137.

Hochmann, J., M. Azadpour, and J. Mehler (2008). Do humans really learn anbn artificial grammars from exemplars? *Cognitive Science, 32*(6), 1021–1036.

Hochstadt, J., H. Nakano, P. Lieberman, and J. Friedman (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language, 97*, 243–257.

Hopcroft, J. and J. Ullman (1979). *Introduction to automata theory, languages, and computation.* Cambridge Massachusetts: Addison-Wesley.

Jackendoff, R. (1994). *Patterns in the mind: Language and human nature.* New York: Basic Books.

Jennings, R., A. Hartline, and J. Thompson (2010). A biological theory of colloquy. unpublished manuscript.

Jennings, R. and J. Thompson (). The biology of language and the epigenesis of recursive embedding. *Submitted for Publication,.*

Jennings, R. E. (1994). *The genealogy of disjunction.* US: Oxford University Press.

Jennings, R. E. (2005). The semantic illusion. In *Mistakes of reason: Essays in honour of john woods*, pp. 269–320. Toronto: University of Toronto Press.

Jennings, R. E. (2007). Language, logic, and the brain. *International Journal of Cognitive Informatics and Natural Intelligence, 1*(1), 66–78.

Jeong, J. H., J. C. Kwon, J. Chin, S. Yoon, and D. Na (2002). Globus pallidus lesions associated with high mountain climbing. *Journal of Korean Medical Science, 17*, 861–863.

Lai, C., D. Gerrelli, A. Monaco, S. Fisher, and A. Copp (2003). foxp2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain, 126*, 2455–2462.

Lashley, K. S. (1951). The problems of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wile.

Lewis, C. and J. Carpendale (2009). Introduction: Links between social interaction and executive function. *New Directions for Child and Adolescent Development, 123.*

Lezak, M. D., D. B. Howieson, and D. W. Loring (2004). *Neuropsychological assessment: Fourth edition.* New York: Oxford University Press.

Lieberman, P. (2000). *Human language and our reptillian brain: The subcortical bases of speech.* Harvard University Press.

Lieberman, P. (2006). *Toward an evolutionary biology of language.* Cambridge Massachusetts: The Belknap Press of Harvard University Press.

Lieberman, P. (2007). Old time linguistic theories. *Cortex, 44*, 218–226.

Lieberman, P., J. Friedman, and L. S. Feldman (1990). Syntax comprehension deficits in Parkinson's disease. *Journal of Nervous and Mental Disease, 178*(360-365).

Lieberman, P., E. T. Kako, J. Friedman, G. Tajchman, L. S. Feldman, and E. B. Jiminez (1992). Speech production, syntax comprehension, and cognitive deficits in Parkinson's disease. *Brain and Language, 43*, 169–189.

Lieberman, P., B. G. Kanki, and A. Protopapas (1995). Speech production and cognitive decrements on mount everest. *Aviation, Space and Environmental Medicine, 66*, 857–864.

Lieberman, P., A. Morey, J. Hochstadt, M. Larson, and S. Mather (2005). mount everest: A space-analogue for speech monitoring of cognitive deficits and stress. *Aviation, Space and environmental Medicine 76*, 198–207.

Maia, A. S. S. F., E. R. Barbosa, P. R. Menezes, and E. C. M. Filho (1999). Relationship between obsessive-compulsive disorders and diseases affecting primarily the basal ganglia. *Rev. Hosp. Clin., 54*(6), 213–221.

Marsden, C. D. and J. A. Obeso (1994). The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain, 117*, 877–897.

Mattay, V. S., A. Tessitore, J. H. Callicott, A. Bertolino, T. E. Goldberg, T. N. Chase, T. M. Hyde, and D. R. Weinberger (2002). Dopaminergic modulation of cortical function in patients with parkinson's disease. *Annals of Neurology, 51* (2), 156–154.

Middleton, F. and P. Strick (2000a). Basal ganglia output and cognition: Evidence from anatomical, behavioural, and clinical studies. *Brain and Cognition, 42*, 183–200.

Middleton, F. and P. L. Strick (2000b). Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews, 31*, 236–250.

Monchi, O., M. Petrides, J. Doyon, R. Postuma, K. Worsley, and A. Dagher (2004). Neural bases of set-shifting deficits in Parkinson's disease. *The Journal of Neuroscience, 24* (3), 702–710.

Monchi, O., M. Petrides, B. Mejia-Constain, and A. Strafella. (2007). Cortical activity in Parkinson's disease during executive processing depends on striatal involvement. *Brain, 130*, 233–244.

Monchi, O. and P. Petrides (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event- related functional magnetic resonance imaging. *Neuroscience 21*, 7733–7741.

Nambu, A., H. Tokuno, and M. Takada (2002). Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. *Neuroscience Research, 43*, 111–117.

Pawley, A. and F. Syder (2000). Perspectives on fluency. The University of Michigan Press.

Pickett, E. (1998). *Language and the cerebellum.* Ph. D. thesis, Brown University.

Pinker, S. (1994). *The language instinct: How the mind creates lanugage.* New York.: Williams Morrow and Company.

Pinker, S. and R. Jackendoff (2005). the faculty of language: What's special about it? *Cognition, 95*, 201–236.

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realizability. *Cognition, 61* (1-2), 127–159.

Saussure, F. d. (1959). *Course in general linguistics.* New York: Philosophical Liberary.

Siegel, A. and H. N. Sapru (2006). *Essential neuroscience: Revised first edition.* Baltimore: Lippincott Williams & Wilkins.

Skinner, B. (1981). Selection by consequences. *Science, 213*(4507).

Stewart, T. and C. Eliasmith (forthcoming). Compositionality and biologically plausible models. In W. Hinzen, E. Machery, and M. Werning (Eds.), *Oxford handbook of compositionality.* Oxford University Press.

Stowe, L., A. Paans, A. Wijers, and F. Zwarts. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language, 89*(2), 290–299.

Taylor, A. E. and J. A. Saint-Cyr (1995). The neuropsychology of Parkinson's disease. *Brain and Cognition, 28*, 281–296.

Thompson, J. (2010a). The empirical investigation of mental representations and their composition.

Thompson, J. (2010b). Motor control theories of changing colloquy.

Tomasello, M. (1998). Introduction: A cognitive-functional perspective on language structure. In M. Tomasello (Ed.), *The new psychology of language.* New Jersey.: Lawrence Erlbaum Associates, Inc.

Tomasello, M. (2005). *Constructing a language: A usage-based theory of langauge acquisition.* Harvard University Press.

Tomasello, M. (2008). *Origins of human communication.* Cambridge, MA.: MIT Press.

Turnbull, W. (2003). *Language in action: Psychological models of conversation.* East Sussex: Psychology Press.

Ullman, M. (2006). Special issue position paper: Is broca's area part of a basal ganglia thalamocortical circuit? *Cortex, 42*, 480–485.

Vargha-Khadem, F., K. Watkins, K. Alcock, P. Fletcher, and R. Passingham (1995). Cognitive deficits in a large family wtih a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences, 92*, 930–933.

Yin, H. N. and B. J. Knowlton (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7*, 464–476.

Zelazo, P. D. (2006). The dimensional change card sort (dccs): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297–301.

Zelazo, P. D., U. Müller, D. Frye, S. Marcovitch, G. Argitis, J. Boseovski, J. K. Chiang, D. Hongwanishkul, B. V. Schuster, A. Sutherland, and S. M. Carlson (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3).

Index