A Change Vector Method to Study Behavioral Development

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

David L. Cooper
Master of Arts
University of Southern California, 1978
Artium Baccalaureus
Princeton University, 1975

Director: James L. Olds, Professor
Department of Molecular Neuroscience

Spring Semester 2010
George Mason University
Fairfax, VA

Copyright: 2010, David L. Cooper
All Rights Reserved

DEDICATION

This is dedicated to my wife Eileen Conway, who has observed this dissertation in the making with the bemused support and tolerance that I needed; and *in memoriam* to my father and very first thesis advisor, William G. Moulton,who both encouraged me to say what I thought.

ACKNOWLEDGEMENTS

I would like to thank Dr. Jim Olds, who provided the opportunity for me to try for this degree, and guided me throughout the process. I would also like to thank my committee, Dr. Giorgio Ascoli, Dr. Layne Kalbfleisch, and Dr. Jim Gentle, who all provided useful advice and timely encouragement. Dr. Gentle was especially helpful in getting me to think statistically. In addition, I would like to thank Dr. Ernie Barreto for collaborating on articles and for keeping my equations interpretable. Finally, I would like to thank Dr. Ann Butler for advice on numerous anatomical questions, and for "just doing her job."

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A CHANGE VECTOR METHOD TO STUDY BEHAVIORAL DEVELOPMENT

David L. Cooper, PhD

George Mason University, 2010

Dissertation Director: Dr. James L. Olds

Broca's Area was the first region in the human cortex to be tied definitively to a specific behavior—language. However, structural, cytological and molecular peculiarities identified in Broca's Area are not unique to humans, and thus language appears to have emerged from other traits that were advantageous in the evolution of primates in general, such as fine motor control for gestures and vocalizations, and the so-called mirror system. One potential source of insight into the emergence of language is to study the correlation of brain structures with behavioral function. This work capitalizes on the existence of a unique resource to undertake that study: eight detailed cytological studies of the developing human cortex from birth through six years of age, accomplished by JL Conel from 1939 to 1967. Conel's atlases provide a consistent methodology applied to 37 cortical areas at each of the observations ages for neurons and 42 cortical areas for myelinated fibers, which further enables a quantitative comparison of change patterns during human cortical development. The change vector method that was developed to

conduct that investigation normalizes measures for the change steps that occur in the seven change intervals that occur in the Conel data, measures the statistical significance of any of those change steps, and permits the direct comparison of change trajectories using k-means cluster analysis. This analysis reveals significantly correlated synchronized changes at different ages, linked to specific "core" area/layer addresses that imply a clock-like coordination mechanism that appears to support sensorimotor developmental functions at the appropriate age. Neither functional cores nor statistically significant change steps emerge when a similar analysis is applied to verbal behavior, where such a clock-like mechanism is unlikely. In general for the cortex, and for language behavior related to Broca's Area in particular, the analysis supports innate architectural mechanisms that facilitate specific address level accommodation to external activity, whereas external behavioral evidence, as from language change, merely supports rapid evolution of the behavior itself to accommodate those same cortical mechanisms. That is, language evolves to enable the speaker to speak, just as writing evolves to enable the child to read.

This dissertation research has complied with all George Mason University standards for the ethical conduct of research and for the appropriate use of human and animal subjects.

1. Introduction: Broca's Area, a neuroanatomical localization of a complex cognitive function

Paul Broca's 1861 demonstration of linguistic specialization in the left hemisphere of the human brain (Broca, 1861a-d) provided the first synthesis of the clinical description of a pathology with localization of brain function and neuroanatomy with sufficient precision to attract the attention of both scientists and the lay public (Kolb and Whishaw, 1990). His association of stroke patient Tan's difficulties with the posterior half of the second and third left frontal gyri (*circonvolutions*) corresponds to Brodmann area 44 and 45. His clear and compelling correlation of specific symptoms to identifiable structures has likewise been refined and expanded considerably, but remains justifiably famous as a critical tipping point in the history of neuroscience (Gazzaniga, et al., 2002).

Broca first encountered the victim as a 51-year old man named Leborgne, who was dying of cellulitis and gangrene, who had suffered epilepsy in his youth and had been hospitalized since losing the ability to speak at 31. His speech capabilities were restricted to the syllable 'tan' , which became his nickname, and various curse words (Finger, 2004). Broca called this pathology an apemia, meaning "lack of speech" (Broca, 1861c). By 1863, Broca had identified a total of eight similar cases, all with left-frontal lesions (Finger, 2004).

Patients suffering from what is now called Broca's aphasiagenerally produce speech telegraphically, in very short bursts, and delivered with a great deal of effort. In terms of comprehension, they have difficulty with repeating or understanding phrases with unusual or complex word order, such as the English passive, and they also have difficulty repeating long and complex words accurately (Gazzaniga et al., 2002). Broca ascribed these problems to "the motor image of the word." (Kolb and Whishaw, 1990: 580)

While evidence from patients with Broca's aphasia as well as evidence from brain activations indicate that Broca's Area is important for processing syntactic information (Caplan, et al., 2000), other areas in the brain, such as Wernicke's area, including portions of Brodmann areas 22, 41 and 42 (Just et al., 1996), or the anterior portion of the superior temporal gyrus (Dronkers, 1996), are also implicated. These areas lie along the perisylvian area of the brain, which is a highly conserved area in terms of human genetic expression (Thompson, et al. 2000). Their linkage overall has been taken as support for the classic Lichtheim-Geschwind triangle model with a motor processing area (Broca's), an auditory processing area (Wernicke's), and an unlocated conceptual area (Lichtheim, 1885; Geschwind, 1967). However, their common association with the task of processing syntactic information may also indicate that language processing is a very complex and multi-faceted task, and that the triangle model is too simple.

In fact, despite being the first such area identified and one of the most widely studied, the precise correlation of Broca's Area to language and language processing is still a matter of some controversy. This is exacerbated considerably by the fact that

Broca's Area in humans corresponds to similar brain regions in primates. Moreover, the famous foxp2 mutation in humans that reproduces symptoms very similar to Broca's aphasia also affects much wider areas in the human brain; while FOXP2 has nearly identical homologs not only in primates, but in vertebrates generally (Varga-Khadem, et al., 2005).

1.1 Comparative anatomy and linguistic function

1.1.1 Human anatomy and language function

In humans, Broca's Area includes portions of the *pars triangularis* (BA 45) and *pars opercularis* (BA 44) of the inferior frontal gyrus of the brain. This lies across the anterior ascending ramus of the lateral, or Sylvian fissure opposite the temporal lobe, which is both an auditory and visual association area. It is adjacent to the premotor cortex, proximal to the areas that control oral and facial features, particularly the lips and tongue. Many descriptions refer to the perisylvian area of the brain, a shorthand term for the areas astride the lateral fissure that includes Broca's Area rostrally and Wernicke's Area caudally. When studying genetic expression in fraternal and identical twins, Thompson et al. (2001) showed very clearly that the perisylvian area is highly conserved in both groups, whereas many other regions co-expressed in identical twins show divergent genetic expression in the fraternal twins.

Wernicke himself noted the direct connection of Wernicke's area to Broca's Area by way of the angular gyrus and the *arcuate fasciculus* (Wernicke, 1874). Because

Wernicke's area is adjacent to the auditory cortex in Heschl's gyri, he ascribed an auditory memory role to it. Wernicke's aphasia describes a condition where the patient is fluent, but produces nonsensical insertions of word forms (Wernicke, 1874). The two kinds of aphasia are contrasted here in examples drawn from Gazzaniga, et al. (2002: 385-387):

Broca's aphasia:

Spontaneous speech—"Son ... university ... smart ... boy ... good ... good"

Listening— Prompt: "The boy was hit by the girl. Who hit whom?"

Response: "Boy hit girl"

Repeating— Prompt: "Chrysanthemum"

Response: "Chrysa...mum...mum..."

Wernicke's aphasia:

"I called my mother on the television and did not understand the door"

Thus, we have physically connected areas in the brain, where damage to the endpoints corresponds to polar opposite dysfunction. At one end, Broca's, the patient makes sense but has much difficulty with even short sequences, while at the other end, Wernicke's, the patient is completely fluent but makes no sense.

Using voxel-based lesion-symptom mapping, Bates et al. (2003) provides strong support for a functional connection along this neuroanatomical pathway. Summing across 101 patients with lesions on the left side of their brains, they found that those who had difficulty with fluency in language showed a very high correlation between the

symptoms and their associated lesions along the gray matter connections in the *insula* and the *arcuate fasciculus* lying between Wernicke's and Broca's Areas. Similarly, they found the highest correlation of symptoms to lesions for language comprehension in the middle temporal gyrus, in the vicinity of Wernicke's area. These observations essentially confirm the association of Broca's aphasia to fluency and Wernicke's aphasia to word comprehension, but the highest symptom-location correlations for fluency lie on the fiber pathway itself, and posterior to Broca's Area in the cortex, while the correlation of cortical lesions to Wernicke's aphasia is much stronger.

This fiber pathway is also implicated in learning to read, which Dehaene (2009) attributes to the "recycling" of neurons in the left occipito-temporal fissure to specialize in the distinction of letter and character shapes with outputs to language areas thereafter. Dahaene (2009) argues that reading is critically dependent on the ability to associate graphemes from the occipito-temporal fissure with phonemes in Wernicke's area, which is supported by the early maturation of activity in the intervening left superior temporal sulcus area in juvenile readers (Turkeltaub, et al., 2003). In expert, adult readers, both the superior temporal sulcus and Broca's Area are active. Dehaene (2009: 207) ascribes this pattern to the conscious ability to manipulate elementary speech sounds and describes these two brain regions as "the anchor points for reading."

As for Broca's Area itself, using an observer-independent method for measuring cell densities, Amunts et al. (1999) focused on the detailed structure of the cortical layers in ten human brains, taking thousands of profiles across five male and five female subjects. They confirmed the left-biased lateralization that is normally imputed to

language functions. However, while all five male subjects showed a left over right asymmetry for area 44, which is adjacent to area 6 in the motor cortex, only three of the female subjects did. Neither group showed a similar asymmetry for area 45. Their other findings also show clear structural differences between areas 44 and 45 at the cellular level, but without large-scale landmarks marking their limits. There are also significant differences between subjects. Thus, the size of area 44 on the left could vary by a factor of ten between subjects, while the differences between subjects for both areas were greater than the differences between hemispheres in a single subject.

1.1.2 Primate anatomy and function

Despite the association of Broca's Area with human language, there are homologous regions in other primates. An area equivalent to area 45 is present in monkeys (Preuss and Goldman-Rakic, 1991). Chimpanzees have equivalents to areas 44 and 45 (Carroll, 2003), and a recent cytoarchitectural and electrostimulation study revealed an equivalent to area 44 in macaques as well (Petrides, et al., 2005). Moreover, the Broca homolog in great apes shows the same kind of left over right asymmetry found in area 44 in humans (Cantalupo and Hopkins, 2001).

Consequently, this left-biased asymmetry clearly cannot be related just to human language. In fact, there is evidence for left-hemispheric dominance in all the early hominins, as well as modern chimapanzees, bonobos and gorillas (Hopkins and Leavens, 1998). This dominance corresponds to right-handedness, pointing gestures, and vocalizations. Similarly, there is a parallel left-asymmetrical extension of the *planum*

temporale, adjacent to Wernicke's area in humans, found in *Homo habilis*, *Homo erectus*, and *Homo neanderthalensis* (Holloway, 1980). This left-asymmetric extension is also found in chimpanzees (Gannon et al., 1998; Hopkins et al., 1998). Thus, the asymmetry probably existed throughout the perisylvian region prior to the appearance of hominids.

However, these left-biased asymmetries are not identical across primates. For example, left lateralization in Broca's Area in great apes is evident down to the level of fine structure (Amunts et al., 1999), yet the lateralization in gross structure of the *planum temporale* does not correspond to a similar asymmetry in minicolumn size and connectivity in chimpanzees (Buxhoeveden et al., 2001). Moreover, Fritz, et al (2005) reported that monkeys had deficits in responding to auditory stimuli when short-term memory areas in rostral superior temporal gyrus or the medial temporal lobe were lesioned, but no loss in performance when the lesions occurred in rhinal cortex. Thus, they appear to be unable to form long-term memory of auditory stimuli.

This produces an overall pattern that implies that the perisylvian area has been important to primates for millions of years, based on gross structure. Homologous fine-structure asymmetry appears in great apes for Broca's Area, while it does not at the *planum temporale*. This would make the fine-structure changes near Wernicke's area, the *planum temporale* and the auditory cortex in general more likely locations for at least some language-specific changes such as long-term auditory memory than Broca's Area, where fine structure changes already took place before the appearance of language.

This leaves the corollary issue of what non-linguistic functions might be associated with Broca's Area. Kohler and colleagues shed some important light on this

question by examining the firing patterns of individual visuo-motor "mirror neurons" in area F5 of macaques, the homolog to Broca's Area in humans (Kohler et al., 2002). These mirror neurons were already associated with action-related perception that required viewing both an agent, such as a hand or a mouth, and an object manipulated by the agent (Gallese et al., 1996). These neurons are thought to be important in planning and execution of movement. Kohler and colleagues showed that these neurons are multi-functional, and also react to the sounds produced by objects on which the monkey performed an action.

1.1.3 Mirror neurons

Mirror neurons may have an even closer tie to linguistic performance than multi-modal correlation. These neurons in area F5 appear to code goal-oriented movement of the hand and mouth (Rizzolatti and Camarda et al., 1988; Murata et al., 1997; Rizzolatti and Fogassi et al., 2000). Some of these mirror neurons are highly specific, coding particular types of grasping movements, for example, but most of them are active under much broader sets of stimuli, and appear to generalize across classes of particular instances (Rizzolatti, Fogassi, and Gallese, 2001). Nelissen, et al., (2005) provided further functional magnetic resonance imaging (fMRI) evidence for the macaque homologs for BA 44 and BA 45. These areas favor gestural information over object identification, but also distinguish their focus between hand gestures in the homolog to BA 45 and actions taken by an acting person in the homolog to BA 44, representing the action and its context, respectively. Understanding of these gestures may very well be

accomplished by mapping the visual representations onto their motor representations, creating a type of "motor knowledge."

There is both direct and indirect evidence that humans have mirror neurons, and that Broca's Area responds similarly to area F5 when humans undergo experiments on arm and hand actions (Rizzolatti and Fadiga et al., 1996; Grafton et al., 1996; Decety, 1997; Grèzes, 1998). These experiments show the left over right asymmetry associated with primate anatomy already noted, and associate Broca's Area very clearly with "meaningful" rather than "meaningless" gestures. In assessing the evidence for and against competing hypotheses related to the primate mirror system, Rizzolatti, Fogassi and Gallese (2001) point out that the main weakness in the "visual hypothesis"--whereby actions are understood solely on the basis of their visual inputs, without reference to motor representations--is that there is no mechanism for validation of the meaning of the observed action. By contrast, "motor knowledge" provides the mechanism for validating and understanding gestures under the "direct matching hypothesis."

Mirror system signals to Broca's Area from the parietal cortex (Rizzolatti, et al., 2001) are associated both with language and non-linguistic motor function. For example, BA 39 is related to language by lesion and cognitive deficit correlations, including deficits in active voice, agentless passives, subject relatives, object clefting, negative passives and object relatives (Dronkers, et al., 2004). In reading tasks, BA 39 is correlated with BA 18 ($r = 0.60$) and BA 20 ($r = 0.63$), and is implicated in developmental dyslexia as well as acquired alexia (Horwitz, et al., 1998). Similarly, BA 7, which is generally associated with mental rotations, observation of tools, and

recognition of motor actions (Binkofski, et al, 1998), is also implicated in imitative learning (Iacoboni, et al., 1999; Iacoboni, 2005), and imitative learning is crucial to language acquisition (Tomasello, 2003). In AIP, a macaque homolog to BA 7, neurons distinguish types of handgrips and particular shapes (Sakata, et al., 1992). The co-activations near Broca's Area for these functions include BA6 as well (Rizzolatti, et al., 2001). Neurons in macaque area PF, homologous to BA 7 and 40, are active for hand and eye movements (Nishitani, et al., 1999).

1.1.4 Non-linguistic functions and memory in humans

Besides language, meaningful patterns and sequences related to music also trigger responses to areas in human brains that are generally associated with language. In an fMRI study, which provides spatial resolution on the order of 1-2 mm and temporal resolution on the order of a few seconds (Logothetis, et al., 2001), subjects exposed to musical sequences ending in discordant notes reacted significantly more throughout the perisylvian region than when exposed to note sequences that behaved according to rules of tonality with which they were familiar (Koelsch et al., 2002). Magnetic encephalography (MEG), which has temporal resolution of a millisecond (Murakami and Okada, 2006), provides good evidence that Broca's Area is involved in processing this "musical syntax" at different timescales. For example, "in key" and discordant tones produce distinctly different levels of activity, and Broca's Area is particularly active in the case of discordant tones, with the reaction occurring approximately 200 msec after the tones are heard (Maess et al., 2001). These are consistent with data on linguistic

perception, where gaps in a "tone group" are perceived from 80 to about 240 msec, and ignored otherwise (Butcher, 1981). In sequential processing, when confronted with a key phrase that was grammatical in a simple sentence, grammatical in an embedded sentence, and inserted ungrammatically into a third sentence, subjects showed a similar electrical activity (a P600 event-related potential/ERP) beginning at 200 to 300 msec, and reaching a maximum amplitude at 800 to 900 msec for the ungrammatical structure. Incongruous tones in musical sequences showed a virtually identical pattern (Patel et al., 1998). Reaction times and amplitudes tend to be proportional to the "distance" from anticipated musical values, similar to the hierarchy of reaction times in language as phrases become difficult or impossible to interpret (Patel, 2003; Koelsch et al., 2000).

Broca's Area also appears to have an important role in memory. For language, bilingual subjects show distinctly different fMRI activation patterns for a given language depending on whether their second language was acquired simultaneously with their first one, or later when they were adults. In the case of subjects who learned two languages at the same time, activation patterns in Broca's Area overlap considerably. When the second language was acquired in adulthood, the two areas are distinct (Kim et al., 1997).

Broca's Area is also involved with a number of memory tasks apart from language. This returns to the multi-modal associations with the area noted earlier in monkeys. In humans, Broca's Area is involved in both spatial and object memory, in storage tasks and in executive tasks (Smith and Jonides, 1999).

1.2 Genetic evidence and a "clock" hypothesis

1.2.1 Foxp2

The foxp2 gene was isolated thanks to a point mutation in the KE family in which afflicted members have problems with fluency and grammar (Varga-Khadem, et al., 2005). However, just as the parallel anatomical patterns and behaviors related to Broca's Area in great apes and area F5 in monkeys rule out an exclusive tie between Broca's Area and human language, the evidence also rejects an exclusive tie between foxp2 and language. The gene is highly conserved, among the five percent most highly conserved genes in human-rodent pairings, for example (Enard, et al., 2002). Moreover, the human form of the gene differs from the chimpanzee version at only two amino acids. The two changes occurred in the last 4 to 6 million years, after the branching of the hominid line from its common ancestor with the chimpanzee. However, this is twice the expected mutation rate, providing evidence for intense evolutionary pressure on human ancestors at that time. The two changes, in association with co-located alleles on the seventh chromosome, show evidence of an evolutionary "sweep" no more than 200,000 years ago. (Enard et al., 2002). That is, the evidence points to a small, important change related to Broca's Area, language, and to the brain.

FOXP2 belongs to a family of "forkhead box" proteins, which regulate the expression of their respective DNA sequences by means of a three-winged helical structure (Carlsson and Mahlapouu, 2002). FOXP2 normally has arginine at site 553, while mutant FOXP2, as found in the KE family, substitutes histidine, which is adjacent

to another histidine in the third helix (Lai et al., 2001). The analogous mutation at that site in FOXC1 causes a critical loss of function (Saleem et al., 2003).

There are only two differences between humans on the one hand and chimpanzees and gorillas on the other (threonine to asparagine at site 303 and asparagine to serine at site 325), three between humans and orangutans, and five between humans and mice (Zhang et al., 2002). Of particular interest when we take up comparisons of functions at Broca's Area to similar functions in songbirds later, there are only eight differences between humans and zebra finches, making the protein 98 percent identical (Haesler et al., 2004).

The human-specific change at site 325 probably created a substrate for phosphorylation by protein kinase C (Enard et al., 2002). This particular prediction is based on an artificial neural network method that can estimate a protein's structure with an accuracy that exceeds 70 percent, and generally lies between 75 and 82 percent (Rost, 1996; Sun, 1997). This new substrate may well be the small but important change that enabled Broca's Area and related areas, particularly in the limbic system and cerebellum, to function as they do in language. On the other hand, a new phosphorylation substrate may simply be an example of the general upregulation of gene expression in the brain that characterizes the differences between humans and chimpanzees (Preuss et al., 2004).

Knockout mice lacking a form of protein kinase C (PKCγ) that occurs in climbing fiber cells are capable of learning simple motor skills, but are impaired in the smooth coordination of those same skills, such as in walking or balancing on a narrow object (Chen et al., 1995). With these mice, the usual paring back of synapses between

climbing fiber cells and Purkinje cells does not take place during development, so that more than one climbing fiber cell will have synapses with the same Purkinje cell in adults. In normal mature mice, there is a one to one correlation. Mice lacking foxp2 do not reach maturity. However, mice missing one copy of foxp2 experience altered ultrasonic vocalization, as well as cerebellar abnormalities associated with Purkinje cells (Shu, et al., 2005). These results support the correlation of FOXP2, protein kinase C, and the smooth coordination of compound motor activity, such as that required for fluent speech, albeit outside the neocortex.

1.2.2 Other fox genes

FOXP2 is one of a subfamily of FOXP proteins that has at least three other members. Proteins in the subfamily contain four signature domains: a DNA-binding winged-helix, a leucine zipper, a zinc finger, and a polyglutamine tract (Wang et al., 2003). All of them appear to be highly conserved, with the winged-helix the most divergent in structure. Generally all seem to function by repressing genetic transcription, especially during development. The leucine zipper appears to foster dimerization of FOXP proteins, and both DNA binding and dimerization may be required for these proteins to function. Possibly as a result, FOXP1 and FOXP2 are expressed in mice in different ways in the epithelial tissue of airway branches in lungs, in motor pathways in the brain, in the outer mesoderm of the intestines, and in the outflow tract of the atria of the heart (Shu et al., 2001). FOXP1 and FOXP3 are co-expressed in lymphoid cells.

FOXP4 overlaps the FOXP1/FOXP2 pattern in lung, intestine, and neural tissues as well, which further supports a complex pattern for transcription control (Lu et al., 2002).

The zinc finger domain that appears in FOX proteins in all species with fox genes can modify expression in either direction. For example, while a zinc finger domain normally provides one of the mechanisms for repression, when tested in yeast cells, this domain seems to do the opposite: FOXP2 transcription activity triples when fused to the GAL4 binding domain (Li et al., 2004). The polyglutamine region, larger in FOXP2 than FOXP1, seems to modulate repression activity as well. The phosphorylation site in human FOXP2 provides yet more capacity for varying the activity of the forkhead domain. In general, this set of complex interactions, whether by various combinations of dimerization, or by other means of cooperative or antagonistic control, is representative of the forkhead box family of genes. In fact, the number of types of fox genes in an organism is directly correlated with that organism's complexity: humans have more than 40 kinds of fox genes overall. They are related to a wide variety of developmental disorders, including the linguistic difficulties of the KE family stemming from a mutation in foxp2 (Carlsson and Mahlapuu, 2002).

1.2.3 Foxp2 and anatomy in humans and songbirds

Much like humans, many songbirds learn their vocal patterns by copying the vocal cues they hear. Unlike humans, birds do not have a six-layered cortex, so there is no question of an exact homolog to Broca's Area, but they do express FOXP1 and FOXP2 in a manner strikingly similar to human fetuses (Teramitsu et al., 2004) The

parallel structures for bird song are located in avian pallial and sub-pallial areas, as well as the homologous sets of nuclei within the dorsal thalamus. These avian brain areas help provide sensorimotor integration, as well as skilled, coordinated movement, strikingly similar to the cognitive and motor capacities related to Broca's Area.

In humans, there are both cognitive and motor pathways associating the cortex with the basal ganglia and the cerebellum (Middleton and Strick, 2000). Both are implicated in KE family sufferers from the foxp2 mutation (Varga-Khadem et al., 2005). The principal subcortical areas implicated in human cognitive and motor circuits by FOXP2 expression are the caudate nucleus and putamen, the *substantia nigra pars reticulata* and *globus pallidus* internal segment, as well as the medial dorsal, ventral anterior, and other nuclei of the thalamus on the cognitive loop; while the cerebellum (lobules VIIB, VIIIB, as well as the inferior olivary complex and red nucleus), the dentate nucleus and the medial dorsal, ventral lateral, and other nuclei of the thalamus are on the motor loop.

As for songbirds, there are evident correspondences between these subcortical areas implicated by FOXP2 and the sub-pallial areas involved in the avian song cycle (Jarvis et al., 2005). For example, the anterior (cognitive) loop in both involves the striatum and thalamus. In birds, it contains the lateral Area X (LAreaX) in the striatum, which passes signals to the dorsal lateral nucleus of the medial thalamus (DLM). In zebra finches, FOXP2 is expressed in Area X during the critical period for song learning. In adult canaries, it is expressed in Area X seasonally, when song production is unstable. Its expression in other birds varies similarly, indicating its association with vocal

plasticity (Haesler et al., 2004). Similarly, the motor loop also contains the thalamus, in this case the *nucleus uvaeformis*.

Other correspondences are more notional, but the sub-cortical analogies to avian sub-pallial areas seem very strong. To extend the analogies into the equivalent of the cortex for songbirds, there are also areas in the avian pallium that participate in both the auditory and motor pathways—the higher vocal center (HVC), and the robust nucleus of the arcopallium (RA). Another key area in the cerebrum on the cognitive loop is the lateral magnocellular nucleus of the anterior nidopallium (LMAN). These may play the computational role of Broca's Area. They are involved in the moment-to-moment modulation of syllables in the songs of zebra finches (Kao et al., 2005).

Birds are also capable of acquiring "syntax." White-crowned sparrows, when exposed to their native song in two syllable phrases, were able to learn the entire song sequence despite never hearing the entire song from end to end. Exposed to the song when the syllable pairs were in reverse order, the sparrows learned the song backwards (Rose et al., 2004).

1.2.4 FOXP2 expression in mammals

FOXP2 very likely plays a part in cell-to-cell interaction during development, and persists at a reduced levels in the adult, both in birds and in mammals. Takahashi, et al., (2003) studied FOXP2 expression in rats, reporting initial expression by E13 in the lateral ganglionic eminence (LGE) and subsequent alignment in striosomal compartments of the striatum, in patches with high levels of mu-opioid receptors, and devoid of calbindin-

D28K. LGE is the source of migrating GABAergic interneurons (Nadarajah and Parnavelas, 2002). This implies a role for FOXP2 in interneurons as well as the striatum.

Embryonic neocortical development is "inside-out" with lower layers preceding the upper for layers II – VI (Sidman and Rakic, 1973; Marin-Padilla, 1999). FOXP2 appears only in subgranular layers (Maviel, 2004), so it is expressed only during the first half of that process, and thus affects the layers that communicate across and outside the cortex. In Broca's Area, Layer V contains numerous and noticeably large pyramidal cells (Amunts et al., 1999). If FOXP2 is always associated with GABAergic interneurons, however, it only accounts for modulation of the internal processing and outputs from Layers V and VI.

Broca's Area is associated with language and non-linguistic memory functions as outlined in the previous section. It thus makes associations over times less than a second, and stores patterns acquired over a lifetime. This implies a function requiring both short- and (very) long-term memory. Recent and remote memories have different processes associated with them. In the case of episodic memories, these are related to cortical-hippocampal networks (Frankland and Bontempi, 2005). The expression of C-FOS, which is activity dependent, in cortical layers demonstrates these different processes, and shows distinct differences in storing spatial memories in the parietal cortex of mice (Maviel at el., 2004). That pattern is essentially identical to the expression of FOXP2 with respect to cortical layers: FOXP2 is expressed predominantly in Layers V and VI, which are also the layers showing the greatest c-fos activation, and implying a tie to recent memories; FOXP2 is not expressed in Layers II and III, and hardly expressed in

Layer IV, where the lack of c-fos activity implied a significant relationship to remote memories.

On the other hand, the association of Broca's Area with long term memory (Smith and Jonides, 1999) and especially differential storage of first and second languages (Kim, et al., 1997) as well as the association of Layers II and III with remote memories (Frankland and Bontempi, 2005) implies that other factors than foxp2 affect Broca's Area, even for language functions.

1.2.5 Additional interacting change factors

Just as the differing evidence of left-biased asymmetry across humans and great apes for Broca's Area and the *planum temporale* indicates that more than one set of changes underlies the emergence of language, evidence of development in hominids and great apes also shows that other human traits emerged from more than just one change, no matter how important. Compared to chimpanzees, for example, humans have an immature skull shape and size. This accounts for its relatively large size in humans, but it is unlikely that these differences stem from a single source. For example, human and chimpanzee growth rates differ substantially through adolescence (Gould, 1977). Humans likewise developed more slowly and had more immature features than the earlier hominids (Dean et al., 2001; Rice, 2001), and differences between humans and Neanderthals also arose early in child development (Ponce de León and Zollikofer, 2001). Generally, comparative evidence indicates a mosaic pattern of developmental

traits, and not a simple change of rates or the acquisition of a single new trait (Moggi-Cecchi, 2001

Developmental changes of any kind, much less acquisition of complex traits such as language are thus very likely to involve the interaction of multiple factors. During corticogenesis, known developmental patterns involve interacting factors, such as cellular precursors and cell cycle duration (Lukaszewicz, et al., 2005), or short and long-term control factors (Donoghue and Rakic, 1999). Generally, mammalian neocortical development patterns conserve GABAergic transverse migration patterns, the transformation of cortical germinal zones and the striatocortical junction, and the amplification of the REELIN signal pathway (Molnár, et al., 2006).

In addition to reelin, the fate of cortical projection neurons, and thus of neocortical development is controlled by a number of other factors as well, such as emx2, pax6, lhx2, and foxg1. Without emx2 or pax6, for example, the neocortex fails to form at the expense of the basal ganglia (Muzio, et al., 2002). Similarly, without lhx2, the neocortex fails to form at the expense of the cortical hem and choroid plexus (Monuki, et al., 2001). Suppression of foxg1 likewise suppresses the formation of both basal ganglia and neocortex at the expense of the cortical hem and the archicortex (Muzio and Mallamaci , 2005; Hanashima, et al., 2004). Molyneaux, et al., (2007) have proposed a model whereby these four factors provide neocortical potential, which sets the stage for the generation of glutamatergic projection neurons of varying types.

FOXP2 exhibits region and layer-specific patterns of expression and function. FOXP2 is expressed in Broca's Area and Brodmann area 6 in the motor cortex, as well as

the caudate/putamen and inferior lobes of the cerebellum. The KE-point mutation results in reduced gray matter connecting these areas to other areas in the brain, especially in the caudate nucleus, the cerebellum, and the left and right inferior frontal gyrus consistent with the KE family problems with fluency and fine motor movement in the face and mouth. Conversely, it is also associated with increased gray matter in the *planum temporale* (Belton et al., 2003).

This pattern is further consistent with expression of FOXP2 in mammal fetuses, where the mRNA signal appears on the inner cortical plate and is limited to the tissue below the granule cells in Layer IV, and especially in Layer VI (Lai et al., 2003; Ferland et al., 2003; Takahashi et al., 2003). Layer VI is principally involved in projections to the striatum and thalamus (Creuzfeldt, 2005), and mutant foxp2 disruption corresponds to the same areas. KE family members with the point mutation show significant lack of activity in Broca's Area and the putamen between Broca's Area and the limbic system in verbal generation tasks. Possibly in partial compensation, they show heightened activity in verbal tasks in the *planum temporale*, and in both hemispheres (Liégeois et al., 2003), consistent with added gray matter.

1.2.6 Control of neocortical development in mammals

Neocortical development has unique features in primates, particularly in the development of the upper layers, where regulation of the G-1 portion of the cell cycle is especially important in determining the numbers and types of neurons in specific areas by controlling the rates of neuron production (Dehay and Kennedy, 2007). At least two

dozen other proteins act as substrates to control the activity of cdk5, which regulates the cytoarchitecture of the central nervous system (Dhavan, et al., 2001).

Postnatal development responds to comparable control. Lamina-specific developmental control of thalamocortical projections, with thalamic projections to the cortex preceding projections back to thalamus, has been observed in rat and mouse neocortex. Yamamoto, et al. (1997) and (2000) describe developmentally-regulated and lamina-specific expression of molecules in rat cortex that control axonal branching, particularly in Layer IV. Jacobs, et al. (2007) reveals a post-natal pause at 10-14 days before early corticothalamic projections begin to invade their target nuclei in the thalamus. Laminar distribution of NR1, NR2A and NR2B NMDA-receptor molecules differs systematically as ocular dominance columns develop in cats (Mower and Chen, 2003). All three types had low expressions in Layer IV and VI postnatally, but a complementary distribution between NR1 on the one hand and NR2A/NR2B emerged during the critical period, when NR1 appeared preferentially in Layer IV and the supragranular layers, while NR2A/NR2B did the converse. NR1 declined in expression in adult cats after the formation of permanent dominance columns.

Experience-dependent synaptogenesis is also subject to internal genetic control. For example, neurogenin2, lmo4 and neurod2 mediate the activity-dependent refinement of thalamocortical axon terminals (Polleux, et al., 2007). Thus, internal control factors operate at all spatial scales from subcellular to tissue formations, and all ages from conception. Penttonen and Buzsáki (2003) describe a hierarchy of cerebral oscillators related to cortical functions. Buzsáki (2006) is a comprehensive synthesis and review of

related research that makes the case that oscillations provide the unifying mechanism for linking physically separate neural operations. In this view, sensory inputs provide perturbations to self-organizing oscillatory signals. These and subsequent down-stream perturbations are kept coherent by the underlying oscillatory patterns that ultimately enable learning and adaptive behaviors. Llinás (2001) makes a similar argument.

These oscillations are embedded in slower cellular processes, particularly during development. For example, Katz and Shatz (1996) demonstrate that patterned visual stimuli can transform intrinsic activity in cell assemblies into new patterns correlated with the stimuli. Feldman et al. (1999) describe similar results for rat somatosensory cortex. Khazipov, et al. (2004) further establish the correlation between local spindle activity in neonatal rat somatosensory cortex with episodic long-range Layer V and thalamocortical activity linked to spontaneous limb movements.

1.2.7 Noise and differential genetic expression

Gene transcription is subject to two principal sources of noise that can affect the gene regulation function—noise internal to the cell, and external noise. Quantitative assessment of these two sources indicates that the internal noise is subject to very rapid decay, so that the principal influences on single gene regulation are the biochemical factors that trigger gene transcription, together with external noise and only those internal factors that change slowly relative to an entire cell cycle (Rosenfeld et al., 2005). In gene networks, including the gene cascades that proceed as a cell develops, even the small perturbations at the local noise level, however, can have major effects downstream

(Pedraza and van Oudenaarden, 2005; Elowitz et al., 2002). Translational noise is also a source of variation in phenotypes, so the combination of these factors probably led to selection pressure in favor of inefficient translation of genes even when efficient translation would consume less energy, since inefficient translation would lead to lower fluctuations (noise) in protein concentrations in the cell. One well-known example of this inefficient translation is cyclic AMP (Ozbudak et al., 2002).

At the same time, cell-to-cell variations can depend on noise at the transcription level, so that cell populations demonstrate extended bistable states in gene expression. Thus, noise in genetic cascades may very well play a significant role in cell phenotype variation, or cell differentiation as well (Blake et al., 2003). Taken further, this means that incorporation of noise in regulation is an evolvable trait that can help maintain a balance between the fidelity of gene expression and in creating cellular diversity (Raser and O'Shea, 2004). Autocatalysis, or positive feedback, demonstrably contributes to control of cellular function in the context of these bistable states (Becskei et al. 2001), while negative feedback tends to function in homeostatic adjustments (Becskei and Serrano, 2000), and can produce effects five times faster than when negative autoregulation is absent (Rosenfeld et al., 2002). Bistability of this sort essentially creates an all-or-nothing switching mechanism, illustrated in another well-known example by the activity of MAP Kinase in the cell cycle (Ferrell and Machleder, 1998).

1.2.8 A "clock" hypothesis

The remaining chapters in this thesis will focus on a unique set of data on the development of the neocortex in humans (Conel, 1939-1967). It includes observations at eight ages from birth to six years old (N = 54) of neuron somal sizes and dimensions, layer depths, neuron densities from cresyl violet staining, and fiber densities from Golgi-Cox and Weigert staining. It is possible to calculate changes to local neuron populations for the left hemisphere in 37 neocortical areas, including BA 44 and 45. As this review of observations on just Broca's Area illustrates, the underlying factors that produced these population changes in Conel's subjects are legion and only partially known. Nevertheless, it is possible to search for regularities across the neocortex, and then to assess some of the local peculiarities that underlie the postnatal development of human cognitive function.

Given the interweaving of small-to-large scale spatial and temporal changes that determine individual cell fates and the gross organization of the brain, and given the evident utility of noise in the evolution and regulation of genetic cascades, one can postulate 1) a self-organizing and potentially hierarchical developmental "clock" that paces the organization of neocortical regions (Brodmann Areas) needed for specific cognitive processes; and 2) differential heterogeneous response to the "clock" at the level of area/layer combinations that is subject to evolutionary pressure. Generalizing from the examples of left-over-right asymmetry in great ape and hominid perisylvian cortical areas and the extreme level of conservation in foxp2 in vertebrates, one can also posit 3) the conservation of proven hierarchical developmental sequences in the neocortex.

The remainder of this chapter will examine some of the concepts that will be used to probe this hypothesis in light of Conel's data.

1.3 Bistability, attractor regimes, and network dynamics

1.3.1 Multiple time scales

While genes, by virtue of their control of cell differentiation and migration, set down a framework for data processing that specifies what data types interact with each other, the neural constituents of the brain must take that input, use it, and learn from it. Development sets the conditions, while the components of the active brain must then follow through to create an individual's actual competence at any given task. Thus, genetic, epigenetic and environmental factors all interact during postnatal development.

The timescale for this activity spans short and long-term processes over several orders of magnitude. Migrating cerebellar neurons have average velocities of 3-4 μm per minute (Maviel et al., 2004). Thus, migrating cells in rat neocortex could achieve their final position in several days (Bai et al., 2003). By contrast, the actual processing based on the results of all this movement and positioning takes place in 10s and 100s of milliseconds, while learning processes set in over different time courses from minutes to days, months and years for complex behaviors such as language.

1.3.2 Bistability at multiple levels

As described previously, noisy regulation of gene expression can lead to bistable regimes. The processes for information processing and learning are very different from genetic regulation, but they also are essentially stochastic and lead to very similar patterns of stability. For example, sensory changes that affect potassium conductances can shift Purkinje cells from a bistable spiking mode to either of two single stable spiking states (Loewenstein et al., 2005). Bistable behavior is also easy to find in assemblies of neurons. Networks with NMDA-mediated recurrent synapses, consisting of pyramidal cells and interneurons that provide the feedback show bistable behavior that shifts to a single stable state either when GABA conductance prompted by the interneurons shifts above a critical threshold, or when AMPA-mediated conductance shifts the state of the network in the other direction (Lisman et al., 1998).

Models of working memory consisting of recurrent excitatory networks with simpler leaky integration neurons also produce this behavior (Durstewitz et al., 2000). Other studies of bistability in networks show that the hysteresis loops that create these bistable reactive regions, also allow sensitivity by the network to its recent history on input patterns (Pouget and Latham, 2002). Maximization of Gaussian mutual information in the presence of noise turns out to provide the stable computational behavior and sparse coding found in biological neural networks as well (Linsker, 1993).

Activity-dependent processes in excitatory neurons show just this kind of bistable behavior in experimental studies of CA1 hippocampal neurons (Lisman et al., 2002). While the AMPA and NMDA receptors in these neurons function at time scales of far

less than a second, NMDA receptors, anchored by PSD95 connected to AMPA anchored by assemblies containing Actinin, Actin and SAP97, all phosphorylated by calcium/calmodulin-related kinase II (CaMKII) provide an energy-efficient, bistable "switch" that persists on the order of days. At normal local concentrations of Ca^{2+}, this configuration, which relies on an autophosphorylated state of CaMKII, shows only 10 percent dephosphorylation of the kinase after 45 hours (Lisman and Zhabotinsky, 2001).

In the case of Broca's Area, this NMDA-AMPA/CaMKII mechanism would be a reasonable one to impute within a larger framework for synpotogenesis to the numerous and prominently large pyramidal cells in Layer III (Amunts, et al., 1999). Generally, Layer III is also predominantly involved in remote, rather than recent memories (Frankland and Bontempi, 2005), so activation of these neurons in language tasks would probably account for the fMRI patterns that can distinguish native from learned second languages (Kim, et al., 1997). The fact that FOXP2 is subgranular may thus help explain why the KE family members with the foxp2 mutation show fewer difficulties with comprehension than with speech production (Belton et al., 2003).

1.3.3 Dimensions and structure of neuron state and change spaces

Interlinked processes that operated across spatial and temporal scales can be examined productively using techniques from fractal geometry and from attractor systems in dynamics (Buzsáki G, 2006). One such technique is the use of correlation integrals to probe the dimensions and structure of a "state space" (Grassberger and Procaccia, 1983). While correlation integrals are most frequently used in the study of

time series, they have also been applied to examine temporal and spatial synchronization (Vasconcelos, et al., 2004). Cooper (1999) applies them to language change.

Dimensions and structures of spaces become meaningful if one can associate the idea of a distance to the data. Shankle, et al, (1998) attempted to do this for the entire Conel neocortical data set, using an ordination technique called correspondence analysis to organize all of the data into profiles and then measuring concerted upward or downward changes across profiles. They concluded first that Conel's data is remarkably consistent, and thus free from observational noise. They also conclude that these combined profiles of neuron density, myelinated and non-myelinated fiber density, layer depth, somatal size, and somatal dimensions demonstrate a unified developmental process across the neocortex. This already supports the first premise of the clock hypothesis that such a unified process exists. However, correspondence analysis, which is essentially a scale-free procedure, does not produce sufficient differences between neocortical areas to allow an examination of differential patterns within that process. Correlation integrals also rely upon scale-free behavior and are similarly limited.

Microarray analysis provides the basis for an alternative analytic model that retains the scale of pairwise distinctions. While there does not appear to be a rigid process for evaluating gene and protein expression (Ewens and Grant, 2001), the analysis of the conservation of regional genetic expression in human and mouse brains in Strand, et al. (2007) provides a good example for differential analysis of normalized genetic expression by brain region. In comparing expressions of human and murine genes in motor cortex, striatum and cerebellum, they found that amino acid sequences for genes

with non-uniform, patterned distinctions between the three regions are conserved, while they are not for genes with uniform expression.

As to method, after screening for significant regional expression levels, the relevant steps for comparing mouse to man consisted of creating "probesets" of log ratios by area, correction for the variance of human or mouse data, as appropriate, and calculation of the resulting Euclidean distances to examine for clustering. Since Conel provided eight sets of observations at standard locations for all of his subjects, comparable steps can be used as described in the following chapter to process Conel's data. These sets provide an eight-dimensional state space that can be used in turn to create a seven-dimensional change space, where each neocortical area has its own vector, representing its trajectory during development. It is then possible to use the distance metric, this time the correlation distance, to assess statistically significant clustering among the vectors. Consequently, the correlation distance measures correlations between the vectors with greater precision than using correlation integrals. These measures provide considerable information about the change vectors. Each of these concepts will receive a precise definition in the next chapter.

This thesis will employ two of Conel's sets of observations—layer depths and neuron density—to calculate changes to postnatal human neuron populations from birth to six years of age. Neuron population changes probably correlate with general changes to cognitive processing. For example, in plausible noisy neural networks with population recruitment and heterogeneous populations, increased neuron density can confer a learning advantage (Samengo and Treves, 2001; Sompolinsky, et al., 2001; Shamir and

Sompolinsky, 2006). Poisson-like covariances in the firing rates of neuron populations appear to provide nearly optimal Bayes inference (Ma, et al., 2006). Moreover, the stable neuron population at a given age appears to correlate well with generalized environmental demands. Consequently, the results in this thesis could potentially be tested by imagery where structural and fMRI results are contrasted. For example, the fMRI distinction in language areas in bilinguals in Kim, et al. (1997) provides a location for an on-going language task. By contrast, Lee, et al., (2007) reveals that vocabulary size correlates with gray matter density, which could be estimated by structural MRI, but it is not correlated with fMRI activity.

Figure 1-1 provides a sketch of the "clock" hypothesis. As outlined above, developmental inputs are quite complex and include genetic, epigenetic and environmental factors, all of which affect neuron population changes. Those changes will be measured by population *change vectors*, and the population changes will reflect multistable conditions at each layer/area *address*. The methods developed in the following chapter will be used to identify any significantly correlated developmental changes to neuron population.

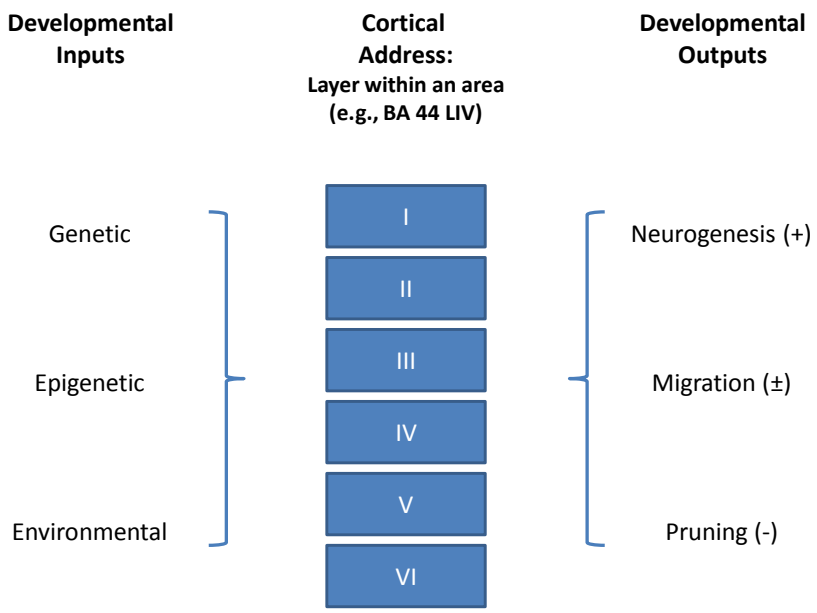


Figure 1-1. Sketch of the "clock" hypothesis for neurons. Conel's data permits tests of developmental models at the level of neuron population changes. We assume genetic, epigenetic and environmental factors act as inputs that effect cortical addresses differentially. We will test emerging change patterns against the null hypothesis that the relative share of neurons at a given address at a given observation age is the same as at the prior observation age, going back to birth. We propose that correlated and statistically significant simultaneous change has a functional basis.

2. A change vector method to assess longitudinal changes

The clock hypothesis in the previous chapter has three components: 1) a self-organizing and potentially hierarchical developmental clock that paces the neurodevelopment of neocortical regions (Brodmann Areas) needed for specific cognitive processes; and 2) differential heterogeneous response to the clock at the level of area/layer combinations that is subject to evolutionary pressure and thus leads to 3) the conservation of hierarchical developmental sequences in the neocortex that promote fitness. The first step in applying this hypothesis to the specific cognitive functions related to Broca's Area, as well as to the human brain in general, is to find a method that can detect developmental changes at the level of area and layer combinations (defined as "addresses" for the remainder of this thesis). In this chapter, we will accordingly derive the normalization steps and analytic methods needed to detect these developmental changes in the Conel (1939-1967) data set mentioned at the end of Chapter 1. In Chapter 3, this method will then be applied to neuron population changes in the developing human neocortex. In Chapter 4, it will likewise be applied to changes to myelinated fiber densities during development. In Chapter 5, the method will be extended to a second body of data related to language change—a complex behavior linked in part to Broca's Area.

2.1. The change vector method

Change vectors consist of normalized changes (change steps) derived from sequential observations. The change vector method introduces a right-tailed test for statistical significance of these change steps, and applies k-means cluster analysis to determine significant correlations between change vectors. The method provides a parametric statistical method for measurements of a single non-negative variable that can be ranked for each iteration of the sequence of observations. In Chapter 3, the ranking will be of neuron populations evaluated across 37 Brodmann Areas and the six cortical layers (222 addresses). In Chapter 4, the ranking will of myelinated fiber densities evaluated across 42 Brodmann Areas in ten layers (420 addresses). In Chapter 5, the ranking will be of usage frequencies of English verbs that happen to use non-'ed' forms for the past tense in different centuries. For each data set, we first transform the rankable observations into normalized change steps to create the respective change vectors for each chapter. Then we compare the change vectors by measuring their correlation with each other as described below. This allows us to isolate which changes are statistically significant, and which sequence of changes are significantly correlated.

2.1.1 Deriving change vector data

Begin with a table *T(r, c)* of *r* cases arranged as rows and *c* columns that record sequential observations of some variable for each of the *r* cases at common times t_1 to t_c. For Chapter 3, *T(r, c)* will be a 222 x 8 matrix of neuron populations where t_1 to t_8 correspond to the eight observations ages in Conel's data described later. Equation 1 normalizes the value of the variable (neuron population for Chapter 3) at each time from t_1 to t_c and defines each member of the log ratio table *R(r, c)*.

Equation 1

$$R(r_i, c_j) \equiv \log T(r_i, c_j) - \left\langle \log T(r_i, c_j) \right\rangle_i$$

We next normalize *R(r, c)* between the times of observation by dividing out the pooled variance across each change step. The sample variance in each column c_j is given by

Equation 2

$$s^2(r, c_j) = \frac{1}{r} \sum_{i=1}^{r} \left[\log T(r_i, c_j) - \left\langle \log T(r_i, c_j) \right\rangle_i \right]^2$$

In Equation 3 we then define the change *C(r, c-1)* for each of the *r* cases and each of the *c-1* change steps defined by *T(r, c)* as the squared difference between $R(r, c_j)$ and $R(r, c_{j+1})$ divided by the pooled variance as defined in Equation 2.

Equation 3

$$C(r_i, c_j) = \frac{\left[R(r_i, c_j) - \overline{R(r_i, c_{j+1})} \right]^2}{s^2(r, c_j) + s^2(r, c_{j+1})}$$

As noted, Equation 3 is defined for *c-1* change steps over a series of *c-1* time intervals (or change intervals) pertaining to the initial observation table *T(r, c)*. For

consistency and simplicity, the three corresponding tables that contain all observations, normalized log ratios and normalized changes in a data set will be referred to as *T(r, c)*, *R(r, c)* and *C(r, c-1)*, respectively. The functions in Equations 1 and 3 will be abbreviated as *R* for $R(r_i, c_j)$ and *C* for $C(r_i, c_j)$. Times of observation t_1 to t_c correspond to the times in table *T(r, c)*. Change intervals will refer to the *c-1* times at which the respective change steps end. Thus, in the Conel data later, the 72 months change interval will be the period between the 48 and 72 month observation ages. In the verbal data, the 10^{th} century would be the change interval between 901 and 1000 AD, which is consistent with current usage.

The normalized change function *C* has some useful properties. Let us begin with the null hypothesis that $R(r_i,c_j) = R(r_i,c_{j+1})$, abbreviated below as *R0 = R1*. This means that the relative magnitude of the variable in *T(r, c)* for each case does not change with respect to all the other cases during any change step. Define the likelihood ratio λ

Equation 4

$$\lambda = \frac{L_{\max}(\omega)}{L_{\max}(\Omega)}$$

where $L_{max}(\omega)$ is the maximum likelihood for the null hypothesis, and where $L_{max}(\Omega)$ is the maximum likelihood of *R(r, c)* in general. This is a case of a nested hypothesis as required by Wilks (1962: 419-421). Then,

Equation 5

$$-2\log\lambda \approx -(R_1 - R_0)^2 \frac{d^2}{dR^2}\log L$$

where L is the likelihood of $R(r, c)$. Using the Cramér-Rao bound

Equation 6

$$Var(R) \sim \frac{-1}{E(\frac{d^2}{dR^2}\log L)}$$

Equation 5 simplifies to

Equation 7

$$-2\log\lambda \approx \frac{(R_1 - R_0)^2}{Var(R_1)}$$

This is asymptotically chi-square with one degree of freedom with increasing n, where $n = 222$ for Conel for each change step and $n = 1554$ for the entire Conel neuron data set in Chapter 3, and $n = 420$ and $n = 2940$, respectively, for Conel's myelinated fiber data set in Chapter 4 (for a general discussion of the $-2 \log \lambda$ statistic see Ewens and Grant, 2001: section 8.4.2). Using the pooled variance for each change step as defined in Equation 2, the $-2 \log \lambda$ statistic is approximately the function C in Equation 3, which is thus also approximately chi-square with one degree of freedom. Since values of $C > 3.9$ have $p < 0.05$ in a right-tailed test of significance against a chi-square distribution with one degree of freedom, it is possible to establish a threshold of statistical significance for changes embedded in any table of non-negative observations $T(r, c)$.

Equation 8 then defines a normalized $(c-1)$-dimensional "change vector" that defines the successive change intervals in $T(r, c)$ for each of the r cases.

Equation 8

$$\vec{P}(r_i) = \left[C(r_i,1) \;\; C(r_i,2) \;\; \cdots \;\; C(r_i, c-1) \right]$$

These vectors are used in k-means cluster analysis as outlined below in section 2.1.4. First, however, we will examine the estimated right-tailed error associated with the normalized function *C*, as well as some implications to be drawn from the null hypothesis that $R_0 = R_1$.

2.1.2 Estimated error of the C(r, c - 1) approximation to a chi-square distribution with one degree of freedom

To estimate the difference between *C(r, c - 1)* and the chi-square distribution with one degree of freedom we use a Monte Carlo simulation. For the simulation, the mean cumulative distribution function (cdf) output from 100 iterations of Equation 3 for increasing numbers of standard normal random input variables with zero mean and standard deviation = 1 (i.e., N(0,1)) was compared to the cdf of the chi-square distribution with one degree of freedom.

Figure 2-1 shows the difference between the two cdfs as a function of *C(r, c – 1)* (panel A), and the logarithm of the difference (base 10) (panel B). N increases from 5 to 37 (the number of Brodmann Areas or subareas described in Chapter 3 (neurons)), to 42 (the number of Brodmann Areas or subareas described in Chapter 4 (myelinated fibers)), to 222 (the total number of addresses discussed in Chapter 3), and finally to 420 (the total number of addresses discussed in Chapter 4). Generally, the difference between the *C(r, c – 1)* and chi-square cdfs decreases with increasing N and with increasing *C(r, c – 1)*

value, which means that the approximation is most valid for right-tail tests. The difference, or error, is less than one percent when the *C* and chi-square statistics are greater than 5. Convergence is generally faster with increasing N.

Figure 2-2 shows the estimated discrepancy in p-value that results from using Equation 3 as an approximation for the chi-square distribution with one degree of freedom. These are smaller than the differences in cdf values, with the greatest discrepancy in p-value at 8.1 percent for $C \approx 0$. For N > 5, the greatest discrepancy is less than one percent for $C = 2.3$. For $C = 3.9$, used as the threshold for statistical significance in this thesis, the discrepancy is less than 7.6×10^{-4} for N = 222 and less than 1.6×10^{-3} for N = 420.

The data sets analyzed in Chapters 3, 4 and 5 have total n = 1554 (neurons), n = (2940 myelinated fibers), and n = 2280 (irregular verbs), respectively. Figure 2-3 depicts the p-discrepancy for datasets of this size. For $C > 3.9$, the discrepancy is approximately 0.001. The discrepancy decreases with increasing values of *C* to approximately 10^{-5} for the largest observed values.

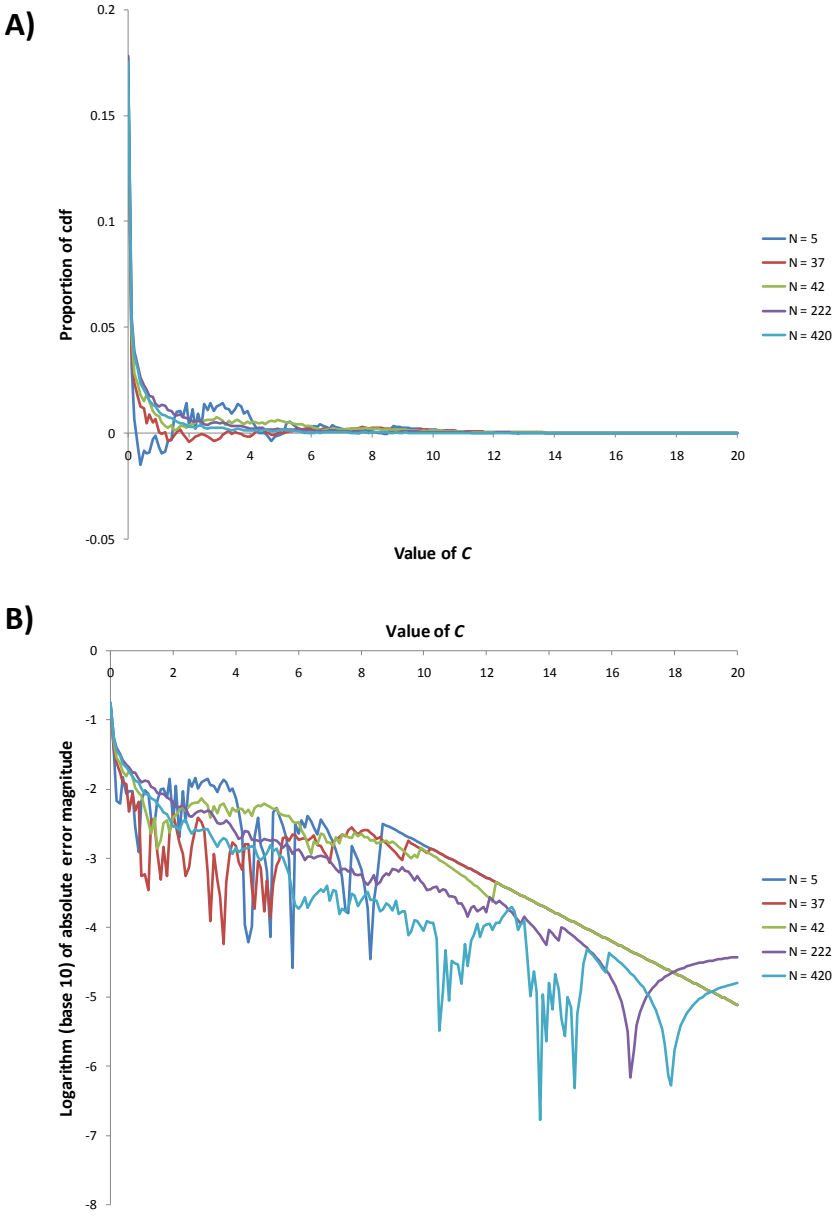


Figure 2-1. Discrepancy in cdf value between Equation 3 and the chi-square distribution with one degree of freedom. A) Discrepancy for increasing N and increasing C. B) Logarithm (base 10) of the discrepancy for increasing N and

increasing C.

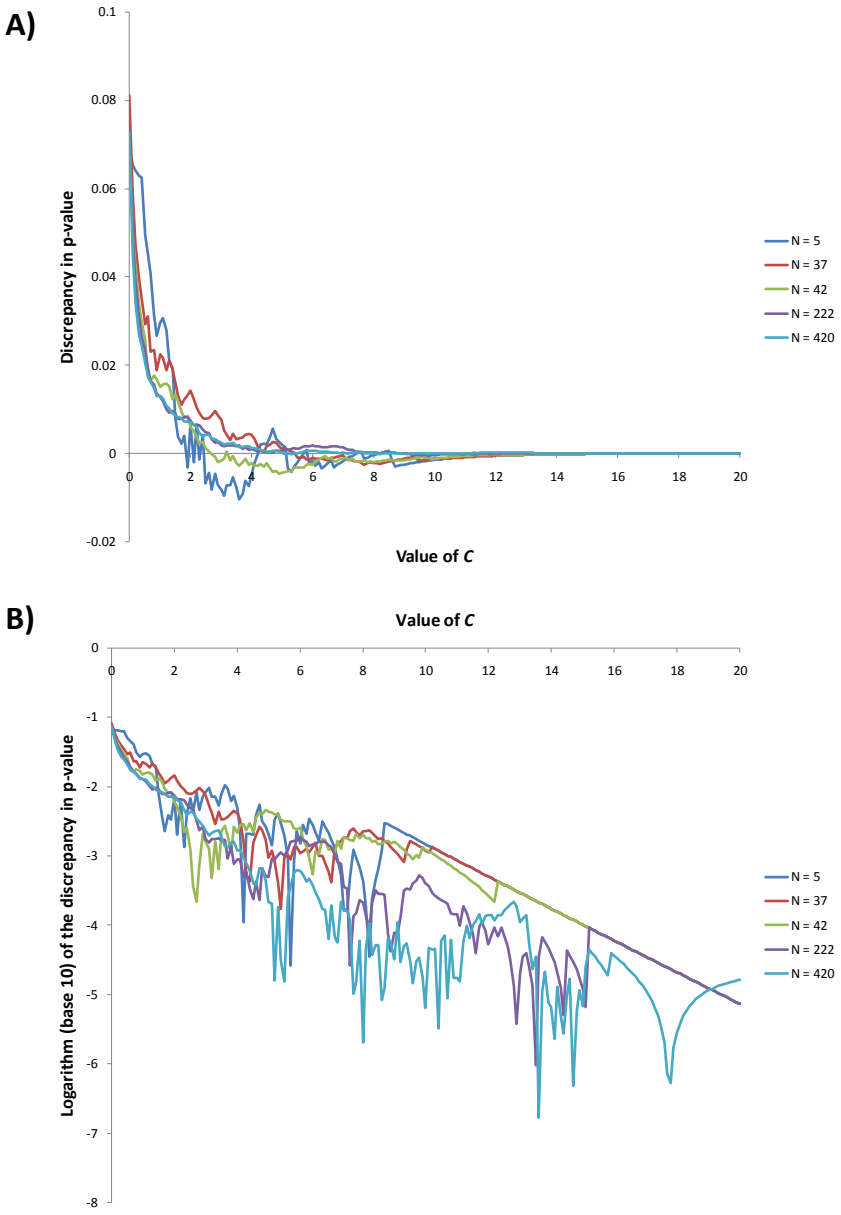


Figure 2-2. Discrepancy in p-value between Equation 3 and the chi-square distribution with one degree of freedom. A) Discrepancy for increasing N and

increasing C. B) Logarithm (base 10) of the discrepancy for increasing N and increasing C.

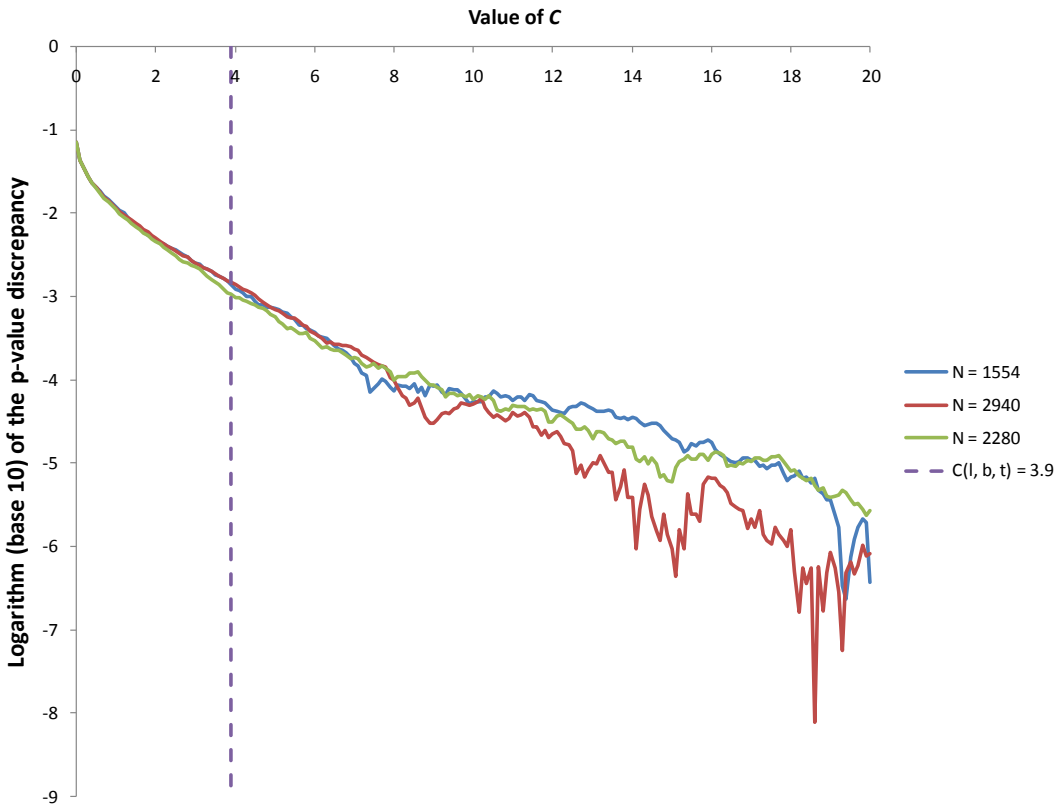


Figure 2-3. Discrepancy in p-value between Equation 3 and the chi-square distribution with one degree of freedom for the datasets in Chapters 3, 4 and 5. The estimated p-value for $C > 3.9$ is $p < 0.05$, based on the chi-square distribution with one degree of freedom. The dashed line shows $C = 3.9$. The discrepancy for the p-value at that point is approximately 0.001.

2.1.3 Interpretation of the $R_0 = R_1$ null hypothesis

In Section 2.1.2 the null hypothesis that $R_0 = R_1$ leads to the statistic C having a distribution closely approximated by the chi-square distribution with one degree of freedom, especially for large n and for large values of C. Here, Equation 1 will be used to show some useful situations where $R_1 = R_0$. In addition, surrogate data will be used to show situations where R_1 is approximately equal to R_0 and where clearly $R_1 \neq R_0$.

For $T(r, c)$ with values drawn from one distribution (analogous to Friedman (1937) for ranks), the value of the standard deviation for any row is the same by hypothesis, with the differences among the values arising solely from sampling fluctuations. The relative magnitude for a particular address in a column of $T(r, c)$ is then a matter of chance; repeated samples will eventually result in the same distribution at each address. The values of $R(r, c)$ will likewise have the same transformed distribution within each column and the distribution $C(r, c - 1)$ is approximately chi-square with one degree of freedom.

The next case of dependence on initial conditions is more interesting. Consider $T(r, c)$ where c_1 is randomly distributed according to some distribution and $c_2 = ac_1$, $c_3 = bc_2$, and so on for each row. That is, for each change step the end state is defined as the same as the initial state multiplied by a constant. In Equation 1 the logratio is dependent on the ratio v_x/v_i, where v_x is $v(x, c_j)$ and v_i is $mean[v(x_i,c_j)]$, where i ranges over all the values in column c_j. Clearly, $v_x/v_i = av_x/av_i = bv_x/bv_i = \ldots$ Thus, $R_1 = R_0$ in all cases and the distribution of $C(r, c)$ is approximately chi-square with one degree of freedom. Exponential growth and decay are two biologically significant special cases of this kind

of dependence: When written in the form $N_0 exp(\gamma t)$ for growth and $N_0 exp(-\lambda t)$ for decay of an initial population N_0, $exp(\gamma t)$ and $exp(-\lambda t)$ are multiplicative constants at each change step.

For the next case, assume that the columns of $T(r, c)$ are normally distributed with mean μ_i and standard deviation σ_i for columns $i = 1, \ldots N$. Thus, the overall distribution of $T(r, c)$ constitutes a mixture of N normal distributions arrayed in N columns. When the columns are independent and identically distributed, $R_1 = R_0$ as already discussed. When the columns contain different normal distributions, Equation 1 transforms the respective means to zero so that $R_1 - R_0 = 0$, but the standard deviation varies, as illustrated in Figure 2-4.

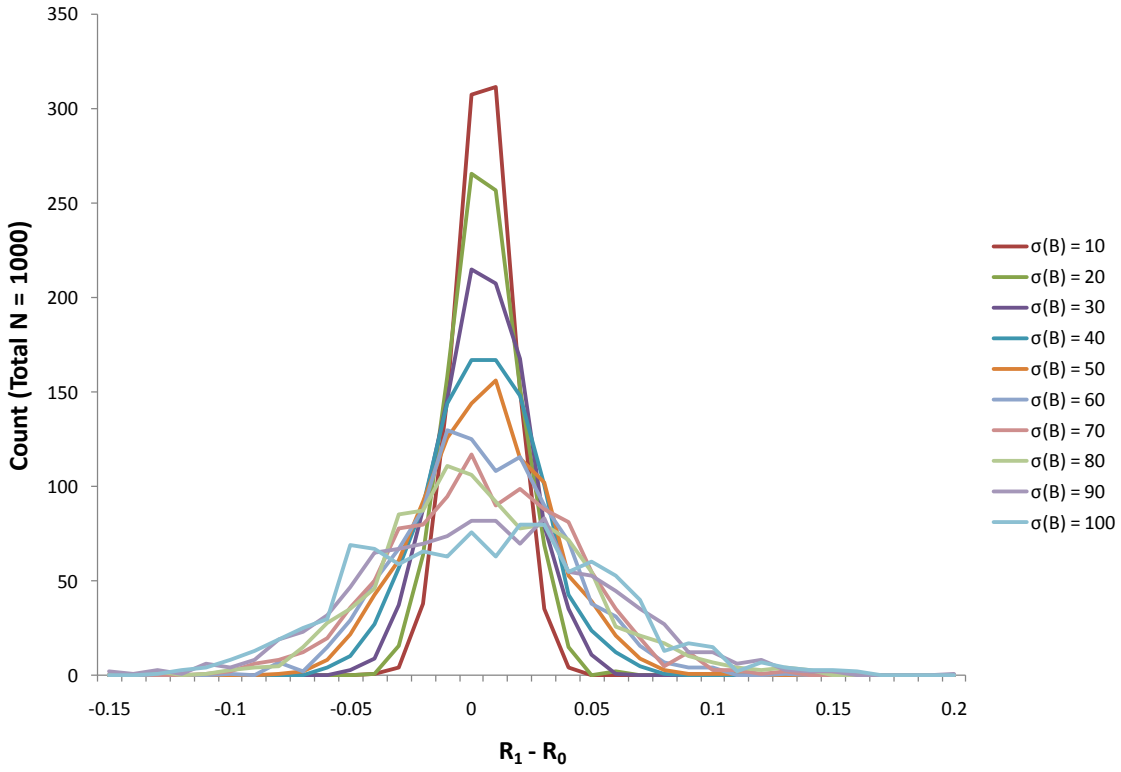


Figure 2-4. Histogram of $R_1 - R_0$ when T_0 and T_1 are normally distributed with differing means and standard deviations. Surrogate data consists of the mean of 100 iterations of two sets of 1000 draws arranged in two columns, A and B. Mean(A) = mean(T_0) = 100, σ_0 = 10. Mean(B) = mean(T_1)= 200. The standard deviation of B = T_1 varies from σ = 10 to σ = 100. When T_0 and T_1 are different normal distributions, but each of the 1000 elements in A and B are drawn from the same normal distribution respectively, R_0 and R_1 are almost the same, with the discrepancy increasing with variance.

Thus, when only one normal distribution exists at each observation age, change steps are only sensitive to different normal distributions with increasing variance. Figure

2-5 shows that Equation 3 does not distinguish these cases with much sensitivity either. A KS test for goodness-of-fit for the cumulative distribution function of *C(r, c - 1)* compared to a chi-square distribution with one degree of freedom fails to reject the null hypothesis that the two are the same until the number of observations *n* is large or until the variance of the underlying distributions is large.

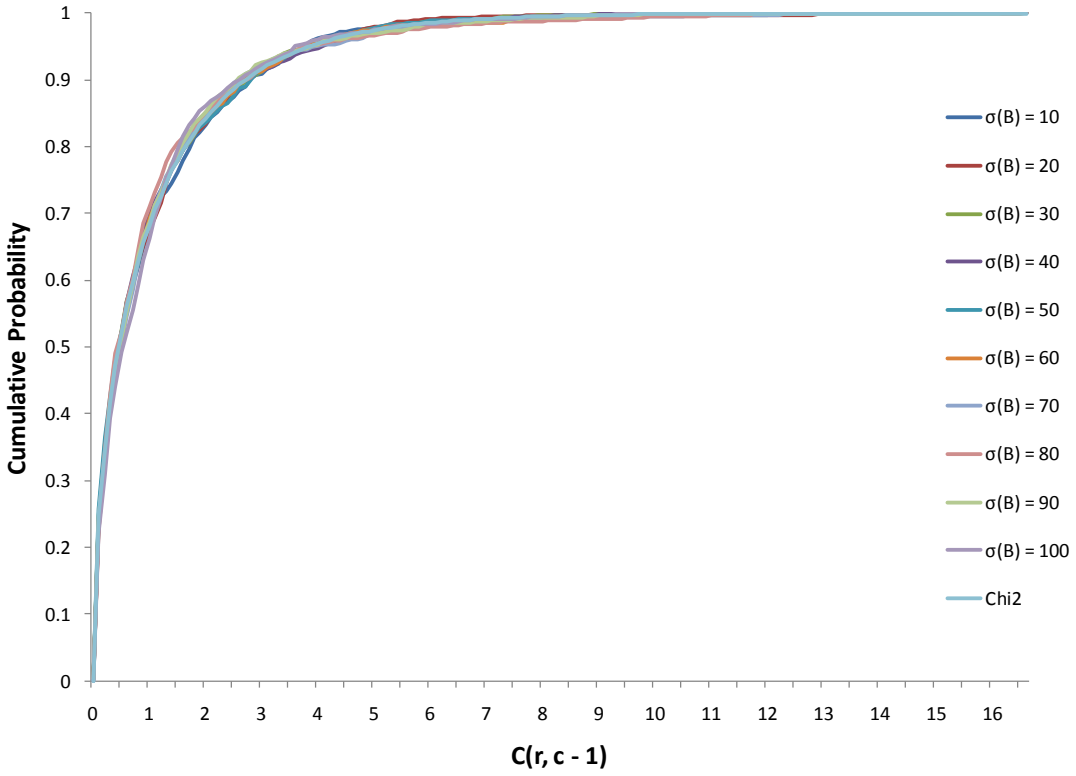


Figure 2-5. Cumulative probability distributions of C(r, c - 1) for the cases in Figure 2-4 compared to a chi-square distribution with one degree of freedom.

Although the distributions for T_0 and T_1 are different, a KS test fails to reject the null hypothesis that *C* for the change step from T_0 to T_1 is the same as a chi-square distribution with one degree of freedom until values for $C > 9.5$ for 100 iterations of 1000 observations. At that point the simulation for $\sigma(B) = 80$ has a cumulative

probability of 1 and $p < 0.05$ for the KS test. When all cumulative probability distributions have reached approximately 1, p for the KS test ranges from $p < 0.05$ for $\sigma(B) = 20$ to $p < 10^{-12}$ for $\sigma(B) = 100$.

Thus, a KS test of change steps does eventually distinguish between two different normal distributions when there is a single distribution in each column in $T(r, c)$. However, when there is a mixture of normal distributions *within* a single column in $T(r, c)$, the KS test is far more sensitive. Figure 2-6 shows $R_1 - R_0$ for the same surrogate set A as used above, but this time the first 500 elements in a new surrogate D have increased from 100 to 150, and the second 500 elements in D have increased from 100 to 200. The standard deviation is allowed to vary as before. In this case, $R_1 - R_0$ is bi-modal. If a third distribution is introduced into a surrogate set, the result is likewise tri-modal (not shown).

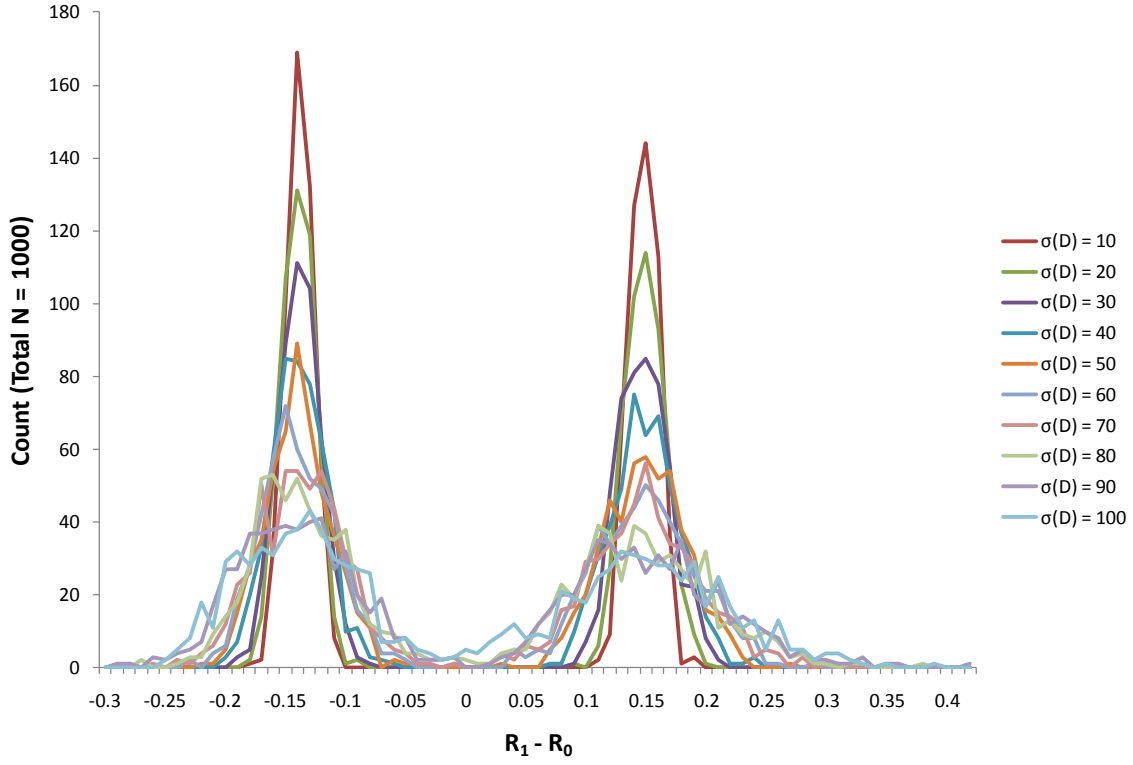


Figure 2-6. Histogram of R_1-R_0 for a mixture of two normal distributions in T_1. The surrogate data setup is identical to the previous figures, except that set D is used for T_1. D consists of 500 elements where the count increases from a mean of 100 in A (or T_0) to a mean of 150, and a further 500 elements where the count increases from a mean of 100 in A to a mean of 200. Standard deviation of D is as shown.

As expected, the cumulative distribution function of $C(r, c_1)$ for this change step is obviously not the same as a chi-square distribution with one degree of freedom as shown in Figure 2-7.

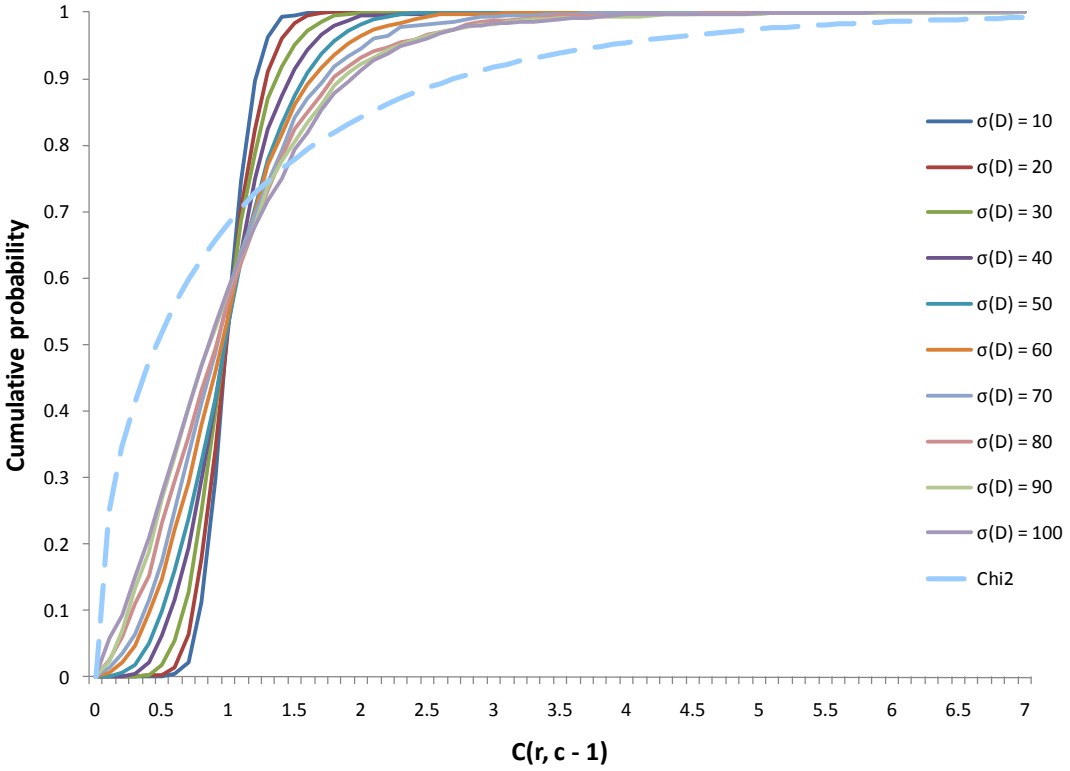


Figure 2-7. Cumulative probability distributions for *C(r, c - 1)* for the case of a mixture of two normal distributions at T_1, compared to a chi-square distribution with one degree of freedom. The KS test in this case is much more sensitive than the previous one where there are no mixtures within the columns in T(r, c). Although the effect of multiple mixtures diminishes with increased variance, p ranges from $p < 10^{-6}$ to $p < 10^{-12}$ for the portions of the cdfs depicted on the graphs, where a comparable KS test in Figure 2-5 would fail to reject the null hypothesis in any of the cases.

To summarize briefly, assume a table of observations *T(r, c)* with r rows and c columns. It has a corresponding logratio table *R(r, c)* derived from Equation 1 and a table of change steps *C(r, c-1)* from Equation 3. For any given change step, to a good

approximation the value of C(r, c-1) will follow a chi-square distribution with one degree of freedom when R_0 at the beginning of the change step and R_1 at the end of the change step are the same. These demonstrations with surrogate data show that the null hypothesis that R_0 and R_1 are the same is most easily rejected using the KS test when the distribution of c_i in *T(r, c)* consists of a mixture of multiple different normal distributions. Conversely, if for each row c_i is a multiple of c_{i-1}, R_0 and R_1 are automatically the same by Equation 1 and the *C(r, c-1)* statistic will follow the chi-square distribution with one degree of freedom. This includes cases of exponential growth or decay where the relationship can be written in the form $N_0 exp(\gamma t)$ or $N_0 exp(-\lambda t)$, respectively. For cases of linear dependence of the form $y = ax + b$, b will contribute to the variance of *R(r, c)*, but this contribution from b would only be detectable for large b and with a large numbers of observations. Thus, when *C(r, c-1)* does not have a chi-square distribution with one degree of freedom, the most likely reason is that the columns of *T(r, c)* contain mixtures of distributions within them. In that case, the k-means cluster analysis can assist in isolating and identifying simultaneous changes among these mixtures.

2.1.4 K-means cluster analysis

The purpose of k-means cluster analysis is to separate a group of data points into k clusters, referred to as k-clusters, where the value of k is optimized given the definition of an appropriate separation distance. The value of k is optimal when the mean k-cluster separation is maximal. To associate correlated developmental trajectories in the clusters, the separation distance was calculated using the correlation distance between the $\vec{P}$

vectors defined by Equation 8. The correlation distance between $\vec{P}_x$ and $\vec{P}_y$ is defined as 1 minus the sample correlation, *r*, between the points treated as sequences of values and is made explicit below in Equation 9. By definition, when $\vec{P}_x$ and $\vec{P}_y$ are completely uncorrelated, $r = 0$ and $D(\vec{P}_x, \vec{P}_y) = 1$, abbreviated as $D = 1$. When $r = 1$, $D = 0$. When $r = -1$, $D = 2$.

Equation 9

$$D(\vec{P}_x, \vec{P}_y) = 1 - \frac{(c-1)\sum_{i=1}^{c-1} C_x(r_x,i)C_y(r_y,i) - \sum_{i=1}^{c-1} C_x(r_x,i)\sum_{i=1}^{c-1} C_y(r_y,i)}{\sqrt{(c-1)\sum_{i=1}^{c-1} C_x(r_x,i)^2 - \left(\sum_{i=1}^{c-1} C_x(r_x,i)\right)^2}\sqrt{(c-1)\sum_{i=1}^{c-1} C_y(r_y,i)^2 - \left(\sum_{i=1}^{c-1} C_y(r_y,i)\right)^2}}$$

In Equation 9, the sums are taken over $c - 1$ change intervals corresponding to $T(r, c)$. For Conel's data in Chapters 3 and 4, *i* ranges from 1 to 7. For the verb data in Chapter 5, *i* ranges from 1 to 10.

Figure 2-8 shows that the optimal value for k for both neurons and myelinated fibers is 7. This graph was calculated using 1000 replications for both data sets to avoid local minima. The effect of replications and increasing k will be shown in Figure 2-9.

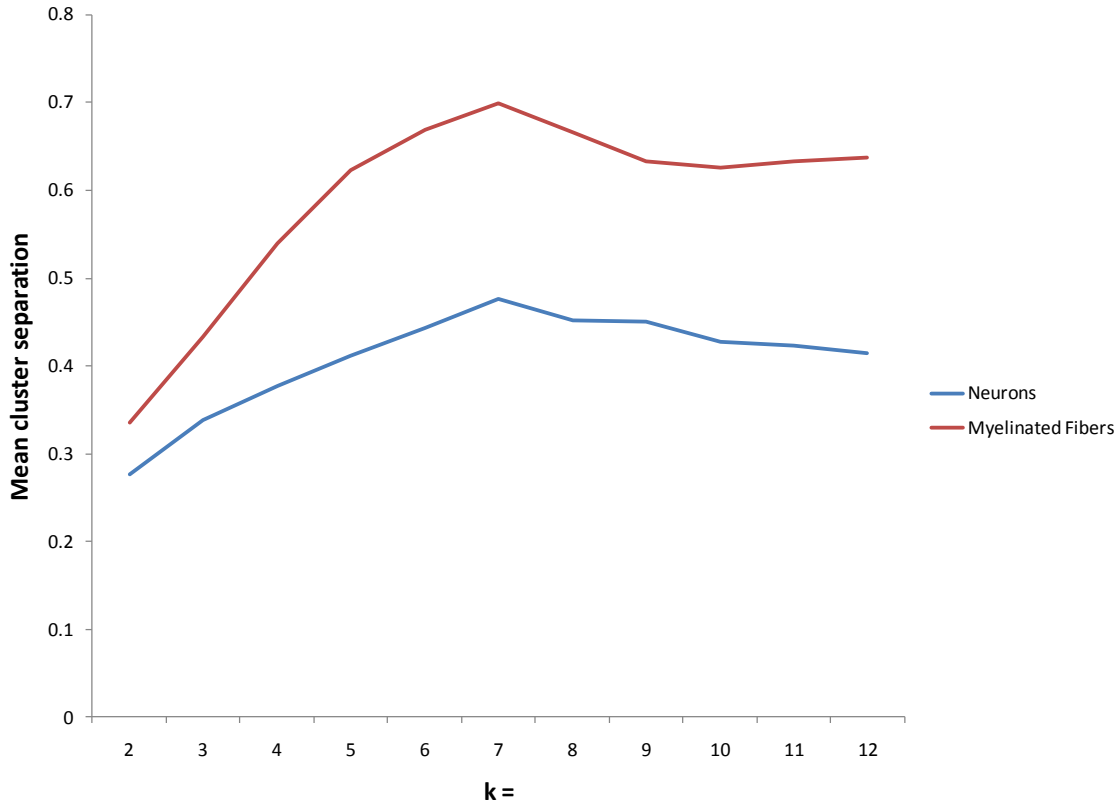


Figure 2-8. Estimation of the optimum value of *k* for k-means analysis of neurons and myelinated fibers. The optimum value of k has the greatest mean separation between clusters. In both cases, the optimum value is 7. The separation distance is calculated using the correlation metric from Equation 9.

The comparable figure for the verbal data in Chapter 5 appears in Figure 2-9. It illustrates the effect of increasing the number of replications beyond the optimum value for *k*. It also illustrates a case where the optimum value is not the same as the number of change intervals in the data (7 for neurons and myelinated fibers compared to 10 for the verbal data).

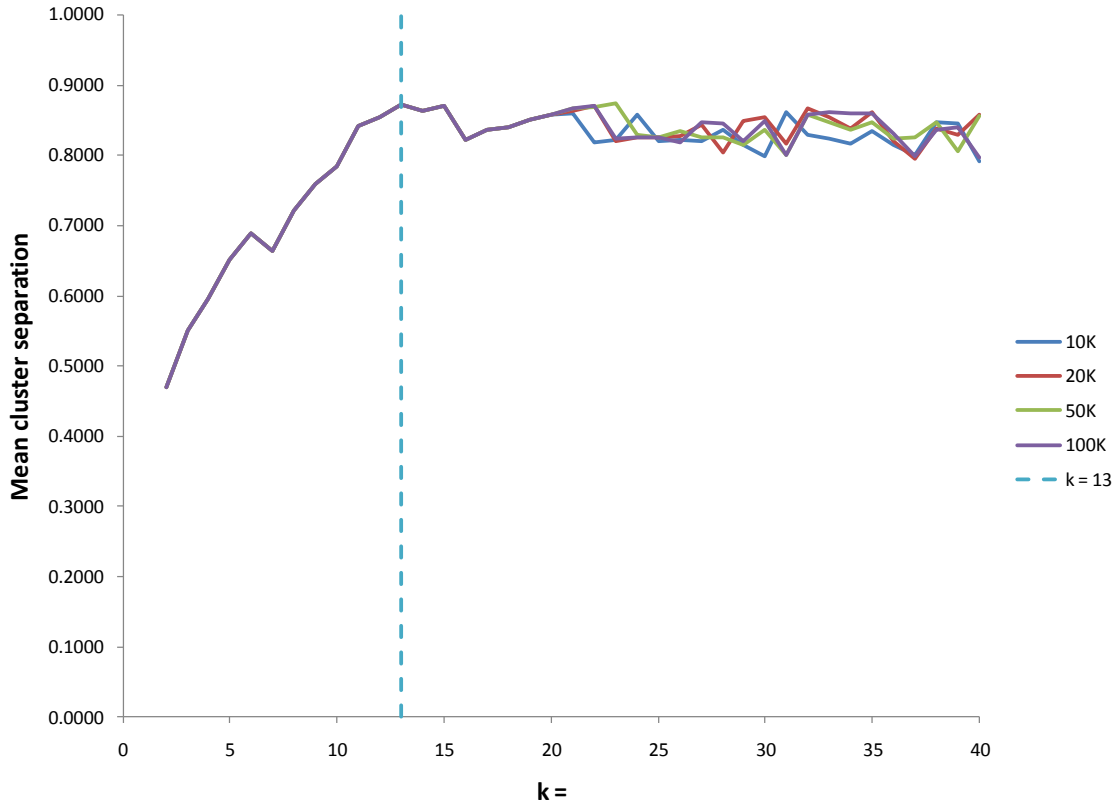


Figure 2-9. Estimated optimal value for *k* for verbal data in Chapter 5. The local maximum of 0.8722 for $k = 13$ emerged with 1000 replications. To test whether the global maximum was the same, the number of replications was increased from 10,000 to 100,000 as indicated in the legend. The global maximum for 50,000 replications (0.8734) exceeded the local maximum at 1000 replications, but this occurred at $k = 23$, in a portion of the curve that was no longer stable. Comparing the cluster separation values for the curves with increasing replications, the curves are identical for *k* up to 20 for 10,000, 20,000, and 50,000 replications, and for k up to 21 for 50,000 compared to 100,000 replications.

The set of k-clusters accounts for all developmental trajectories for the 222 addresses in Conel's neuron data, as well as all 420 addresses in Conel's myelinated fiber data. To analyze each k-cluster for potential biological function, we then examined the relational networks defined for $D < 0.1$ (or $r > 0.9$ within each cluster) in terms of the $C(r, c - 1)$ components of their respective change vectors. For the seven-dimensional change vectors derived from Conel's data, this relationship is statistically significant ($p < 0.01$, two-tailed t-test). Within these highly correlated relational networks, we further isolated the maximal "cliques," or totally connected networks that are subsets of the undirected network graph for $D < 0.1$ (Godsil and Royle, 2004). In the neuron data but not in the myelinated fiber data, several k-clusters had more than one clique of maximal size. In the following chapters, we refer to the addresses in the one or more maximal cliques as the "core" for each respective k-cluster.

2.2 Conel's data

Ideally, any developmental hypothesis would be tested using large longitudinal data sets at the level of granularity relevant to the hypothesis, in this case at the level of individual neuron somata and myelinated fibers within a cortical column. As such a data set does not yet exist for human development, the neuroinformatics analysis in this thesis depends instead on the massive dataset of post mortem histological observations collected by JL Conel (1939-1967). Conel remains one of the few sources for histological data on a systematic, structural basis (Azevedo, et al., 2009). However, his

data set is massive only in the sense of detail at each observation age, not in the sense of large numbers of subjects at each age.

Notwithstanding his reputation as an accomplished clinical neuroanatomist, for his study of human neocortical development, Conel did not report the precise cause of death for any of the 54 subjects. However, he did explicitly state that the cause of each subject's death was unrelated to neurodevelopment.

> Only those brains are accepted as normal which approximate the average weight for the age, which present no malformations or evidence of pathology, and are removed from well-formed, well-nourished infants who died from some disease which, in all probability, would not affect the nerve cells in the brain.
> Conel, 1939, p. 3

While the field of human clinical neuroanatomy has changed since Conel collected his data, for the purposes of this dissertation we assume Conel's assertion to be valid.

2.2.1. Neuron somata

Using classic histological techniques, Conel recorded a mean neuronal density value based on 30 separate cresyl violet-stained tissue samples from each subject within specific addresses. In addition, he recorded the corresponding mean across individuals at each observation age ($5 \leq n \leq 9$; total $N = 54$), for 37 neocortical areas in the left cerebral hemisphere. Conel's procedure avoids possible stereological error from a single soma contributing to adjacent slices (Abercrombie, 1946). As has been reported, the number of cresyl violet-stained somata provides an accurate marker for the neuron population

(Pilati, et al., 2008). Azevedo, et al. (2009) uses cresyl violet to calibrate an automated method for counting neurons using fluorescent DNA-tagged brain tissue. Chapter 3 uses the Conel cresyl violet-stained human neocortical tissue to reveal information about neuron densities during the neurodevelopmental series of observations from birth to 72 months. Conel reported this data as densities for unit volumes of 100 μm^3 for each of the six neocortical layers for a total of 222 addresses.

For neuron populations, *T(r, c)* becomes *v(l, b, t)*, for neocortical layer *l*, Brodmann Area *b* (listing all 222 addresses) and observation age index *t = 1, ..., 8*, corresponding to observation ages 0, 1, 3, 6, 15, 24, 48 and 72 months, respectively. The reported variable is Conel's reported mean density for that address multiplied by the layer depth (in 0.1 mm increments). Thus, *v* represents the number of neurons in a "sample count" with base 100 μm^2 that spans the neocortical layer reported in that measurement. The six sample counts for the addresses in each Brodmann Area, when summed, constitute a *sample column.* For each fixed *t*, Conel's data include observations on each of several individuals for various values of *l* and *b*. The data at the different values of *t*, however, were from different subjects.

Input values for *v(l, b, t)* were carefully cross-checked for accuracy with the raw tabular values for neuron density, layer depth, and myelinated fiber density in the CYBERCHILD data base (Shankle, et al., 2000). As noted in Cooper, et al., (2010), input neuron population values differ from CYBERCHILD for 72 month values, as CYBERCHILD lists an average of left and right hemisphere values for each address, whereas this thesis uses left hemisphere data exclusively because Conel did not make

observations on neuron density in the right hemisphere except at the 72 month time point in neurodevelopment. In addition, as noted in Cooper, et al., (2010), calculated populations also differ from CYBERCHILD for Layer V because CYBERCHILD interpreted Conel's Layer V notations distributing small, large, and extra-large neurons to sublayers Va and Vb as creating separate sublayers for each cell size, whereas Cooper, et al., (2010) totals the cell somata of different sizes within two respective sub-layers.

2.2.2. Myelinated fibers

Conel took his neuron counts from samples in the left hemisphere. He took his Weigert-stained, myelinated fiber samples from the right hemisphere (Conel, 1939). Weigert staining is effective at highlighting extremely small-scale fiber structures (Sammet, 2008). It is used as the standard by which to evaluate alternative methods for imaging neuron fibers. For example, Richter and Warner (1974) compares the accuracy of assessments of unstained tissue to results from Weigert staining.

Conel reported his myelinated fiber results as counts taken from a sample 50 x 50 x 25 μm in volume. These are listed on a table for each observation age ($5 \leq n \leq 9$) as a range with a low and high value. Whereas Conel normally listed neuron counts for 37 areas and six neocortical layers, he listed the myelinated fiber counts in 42 areas for 10 layers (using Arabic rather than Roman numerals to distinguish the list from the six cortical layers recorded for neurons): 1, 2, 3a, 3b, 3c-4, 5a, 5b-6a, 6b, subcortical and vertical. The Brodmann areas for myelinated fibers omit BA 4 for the paracentral lobule that is included in Conel's counts for neurons, but differentiates BA 1 and BA 3 into

separate counts for leg, trunk, hand and head, while these are reduced to a single count for BA 1 and BA 3 for neurons. As a consequence, myelinated fibers are counted at 420 addresses at each observation age, rather than the 222 addresses Conel used for neuron somata.

Given these differences in how Conel reported his results, and because Conel did not report layer depths for the ten layers used for myelinated fibers, fiber populations were not estimated. Instead, for myelinated fibers *v(l, b, t)* uses the mean fiber counts in a 50 x 50 x 25 μm sample volume. Thus, these calculations pertain to changes revealed by samples of fiber density at each address, not fiber population within a volume that spans the respective layer.

Just as neuron population calculations were cross-checked with CYBERCHILD, *v(l, b, t)* for myelinated fiber densities were also checked. As no assumptions about sublayers in Layer 5 occur with the myelinated fiber data, no adjustments were needed to verify myelinated fiber density accuracy. There was missing data at 72 months for myelinated fibers for BA 29 and BA 30. In these cases, the change vectors were calculated using the density data from 48 months.

Note that for Conel's data on neuron populations and myelinated fiber densities, it has been convenient to differentiate the rows in *T(r, c)* into layer and Brodmann Area components, where $l \times b = r$. Likewise, it has been convenient to use *t* for the identity of the respective column, as a reminder that the columns reflect increasing observation age, or time. Thus, in Chapters 3 and 4, *T(r, c)*, *R(r, c)* and *C(r, c-1)* will be discussed as *v(l, b, t)*, *R(l, b, t)*, and *C(l, b, t)* respectively, where *t* for the *C* function is taken as an index

for the endpoint of the seven defined change intervals. In Chapter 5, the discussion will revolve around verbs on a numbered list and centuries, so mnemonics for these functions will be *T(n, c)*, *R(n, c)* and *C(n, c)* respectively, where *c* now stands for century, and is interpreted for the *C* function as *t* is for the neurons and fibers. When all three datasets are described together in Chaper 6, we will revert to the *C(r, c-1)*, etc., nomenclature, where *r* is again a row index and *c* is again a column index.

2.2.3 Conel's nomenclature

Conel (1939) stated that he followed "the method of von Economo," which included his nomenclature for regions of the cortex. Since both Conel and von Economo (and Koskinas, 1925) made extensive use of standard Latin terminology, it is possible to correlate these locations exactly with the Brodmann Areas used throughout this thesis. Table 2-1 provides a crosswalk of Conel's terminology with Brodmann Areas and sub-areas.

Table 2-1. Correlation of Brodmann Areas with Conel's terminology.

BA	von Economo	Nomenclature from von Economo and Conel
4L	FAγ – L	Area praecentralis gigantopyramidalis - leg
4P	FAγ – P	Area praecentralis gigantopyramidalis - paracentral lobule
4T	FAγ – T	Area praecentralis gigantopyramidalis - trunk
4Hn	FAγ – Hn	Area praecentralis gigantopyramidalis - hand
4Hd	FAγ – Hd	Area praecentralis gigantopyramidalis - head
6	FB GFS p	Area frontalis agranularis-gyrus frontalis superior posterior
6 8 9	FB GFM p	Area frontalis agranularis-gyrus frontalis medialis posterior
44	FCBm GFI p	Area frontalis intermedia- gyrus frontalis inferior posterior/pars opercularis
8	FC GFS m	Area frontalis intermedia-gyrus frontalis superior medialis
8 9	FDm GFM m	Area frontalis granularis, macrocellular part-gyrus frontalis medialis medialis

45	FDΓ GFI m	Area frontalis granularis - gyrus frontalis medialis inferior/pars triangularis
9	FDm GFS a	Area frontalis granularis, macrocellular part-gyrus frontalis superior anterior
46	FDΔ GFM a	Area frontalis granularis media-gyrus frontalis medialis anterior
45r	FDp GFI a	Area frontalis granularis, microcellular part-gyrus frontalis inferior anterior
47	FF G Or	Area orbitalis
10	FE P F	Area frontopolaris
3	PB – T	Area postcentralis oralis
1	PC – T	Area postcentralis intermedia
7	PE s	Area parietalis superior
40	PF supra marginalis	Area supramarginalis
39	PG angularis	Area angularis
37	PH basalis	Area parietalis basalis sive temporooccipitalis
19	OA peristriata	Area peristriata
18	OB parastriata	Area parastriata
17	OC striata	Area striata (calcarine cortex)
22	TA	Area temporalis superior
42	TB	Area supratemporalis simplex
41	TC	Area supratemporalis granulose
21	TE	Area temporalis propria
36	TF	Area fusiformis
38	TG	Area temporopolaris
14	IA	Area praecentralis insulae
13	IB	Area postcentralis insulae
24	LA	Gyrus cinguli - area limbica anterior agranularis
23/31	LC	Gyrus cinguli - area limbica superior posterior, granularis
30	LD	Area retrosplenialis agranularis
29	LE	Area retrosplenialis granularis

This table lists the areas used in Conel's neuron data. The myelinated fiber data added information for BA 1 and 3 for the hand, head, and leg, and did not contain any information for BA 4 for the paracentral lobule. Abbreviations: BA = Brodmann Area, Hd = head, Hn = hand, L = leg, P = paracentral lobule, T = trunk.

3. Synchronized changes in neuron populations in the human neocortex

The role of Layer I and reelin in the prenatal formation of mammalian neocortical laminae (Bar and Goffinet, 2000) is well known, as is the pattern of lateral migration of interneurons targeted to specific neocortical areas (Letinic, et al., 2002). These and other control factors, such as bmp, shh, and the homeobox genes appear to follow a basic developmental template shared by vertebrates to which the mammalian modifications have been added (Gallego-Diaz, et al., 2002; Kolpak, et al., 2005; Holland and Takahashi, 2005). Furthermore, the consensus is that the template employs highly conserved genetic and epigenetic factors to provide the necessary coordination of neurodevelopment (Noden, 1991). Human neocortical development is consistent with this pattern to the extent that neuron vs. non-neuron cell counts in human brains have the same relationship to body weight as for primates in general, whereas the great apes such as gorillas apparently have over-sized bodies (Azevedo, et al., 2009).

Developmental studies in animals with high temporal resolution reveal that a neuron's birth cohort is the best predictor of cell-fate during prenatal development (Takahashi, et al, 1999). During postnatal development, neuron death (pruning) putatively replaces neurogenesis as the major factor defining the structure of cortical

neuronal populations (Stiles, 2008). Postnatal histogenic cell death does not recapitulate the prenatal sequence of laminar development (Verney, et al., 2000), but rather seems to be a neurotrophin- and activity-dependent process in which neuron survival hinges on successful synapse formation with target cells (Huang and Reichardt, 2001). This pruning, in conjunction with synapse elimination, might be an efficient developmental process that matches projection cells to target cell populations, but cell death may also serve other functions, such as the elimination of transient structures or error elimination (Stiles, 2008).

Studies of animal postnatal development provide examples of so-called "sensitive periods" when functions such as ocular dominance must be established (DiCristo, et al., 2007). Sensitive periods for complex cognitive functions such as human language are also well-reported (Ross and Bever, 2004; Marcotte and Morere, 1990). In these examples, the animal or human subject interacts with external stimuli as cognitive resources (neurons) become committed to a particular function. Insufficient stimulation can result in the diversion of cognitive resources to alternative functions (Michel and Tyler, 2005).

In this chapter, the neuroinformatics analysis using the methods in Chapter 2 supports the notion that functional, activity-based postnatal alignment of cortical areas occurs at the level of synchronized changes at specific addresses. By address, we mean a specific layer within a particular Brodmann area, e.g., BA 44 LV. We suggest that both the area and layer component of the address are important. Given the clear relationship between function and Brodmann Area (Rakic, 1988; 2001), one would hypothesize that a

Brodmann Area would be predictive of neuron population dynamics during postnatal cortical development. At the same time, Nomura, et al. (2008), showed in a comparative study of the mammalian and avian pallium that most control mechanisms are highly conserved in evolution, including reelin-dependent neuronal migration in mammalian corticogenesis. Thus, if the conserved neurogenetic and neuromigration factors that lead to laminar fate (Takahashi, et al., 1999) as well as to corollary cell death mechanisms (Gohlke, et al., (2007) continue to operate after birth, one would also hypothesize that the neocortical laminar identity would be a predictor of the course of postnatal neocortical neurodevelopment as well.

Here, the human developmental data in Conel (1939-1967) described in Chapter 2 provides evidence that synchronized neurodevelopmental changes occur at the address level in cortical neuron populations. These results support the hypothesis that for human subjects postnatal activity-dependent mechanisms (laminar-neocortical area interactions) are superimposed on a combination of the previously established animal model pre-natal control mechanisms (such as laminar development) to coordinate neocortical neuronal populations over the neurodevelopmental time course from birth to six years. Moreover, the emergence of functionally related neuron changes prior to much of the overt cognitive activity related to those functions implies that these synchronized changes have both an internal, clock-like, basis as well as an activity basis.

3.1 Neuron population results from change vector analysis

3.1.1 Raw neuron population data from Conel

Figure 3-1A shows the mean sample count, or mean neuron population by layer and observation age in Conel, uncorrected for shrinkage. Figure 3-1B shows the corresponding mean neuron population for each layer as calculated from values for each of the 37 Brodmann Areas in the dataset. A sample count is calculated for a volume with a base of 100 μm^2 that spans the respective cortical layer (see Chapter 2).

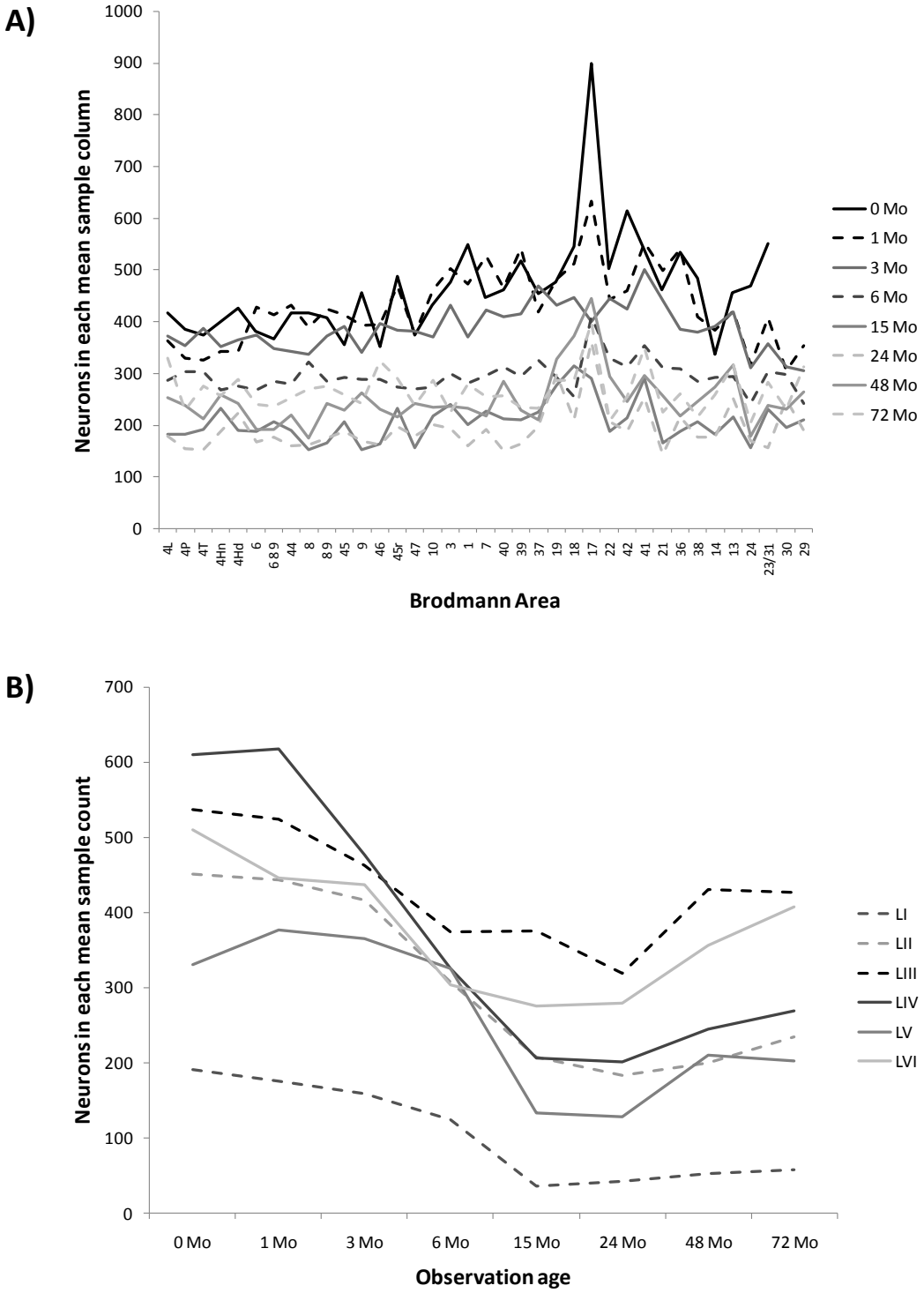


Figure 3-1. Mean raw neuron population by area, layer and age in Conel. Raw

population values have not been adjusted for shrinkage in either panel. A) Mean sample column neuron population by Brodmann Area by observation age, with the mean taken over the six layer samples in the respective Brodmann Area. The curve for zero months reflects missing data, since Conel did not report values for BA 29 and 30 at that age. The raw mean neuron population over all sample columns declines from 463.5 at birth to 249.5 at 48 months. At 72 months it is 266.3. The greatest absolute decline in mean neuron population, 95.9 per sample column, takes place between 3 and 6 months. The greatest relative decline in mean neuron population, 24.55 percent, takes place between 6 and 15 months. B) Mean raw sample counts averaged over each neocortical layer by observation age, with the mean taken over the samples of all Brodmann Areas within a given layer. The mean declines from 438.4 at birth to 192.3 at 48 months. At 72 months it is 266.5. The greatest absolute decline in mean neuron sample count, 92.9 per layer, takes place between 3 and 6 months. The greatest relative decline in mean neuron sample count, 29.91 percent, takes place between 6 and 15 months.

3.1.2 Neuron population change intervals

Figure 3-2 shows the distribution resulting from the normalization steps in Equations 1-3 in Chapter 2. It depicts the distribution of all $C(l, b, t)$ values except nine values greater than 5.5 that are included in Table 1, which lists all 24 statistically significant addresses with $C(l, b, t) > 3.9$. The distribution of $C(l, b, t)$ in Figure 2 is not the expected chi-square distribution with one degree of freedom ($p = 0.0112$, KS test).

As noted in Chapter 2, this supports the alternative hypothesis that *C(l, b, t + 1)* is different from *C(l, b, t)* for some addresses *(l, b)* for at least one change interval.

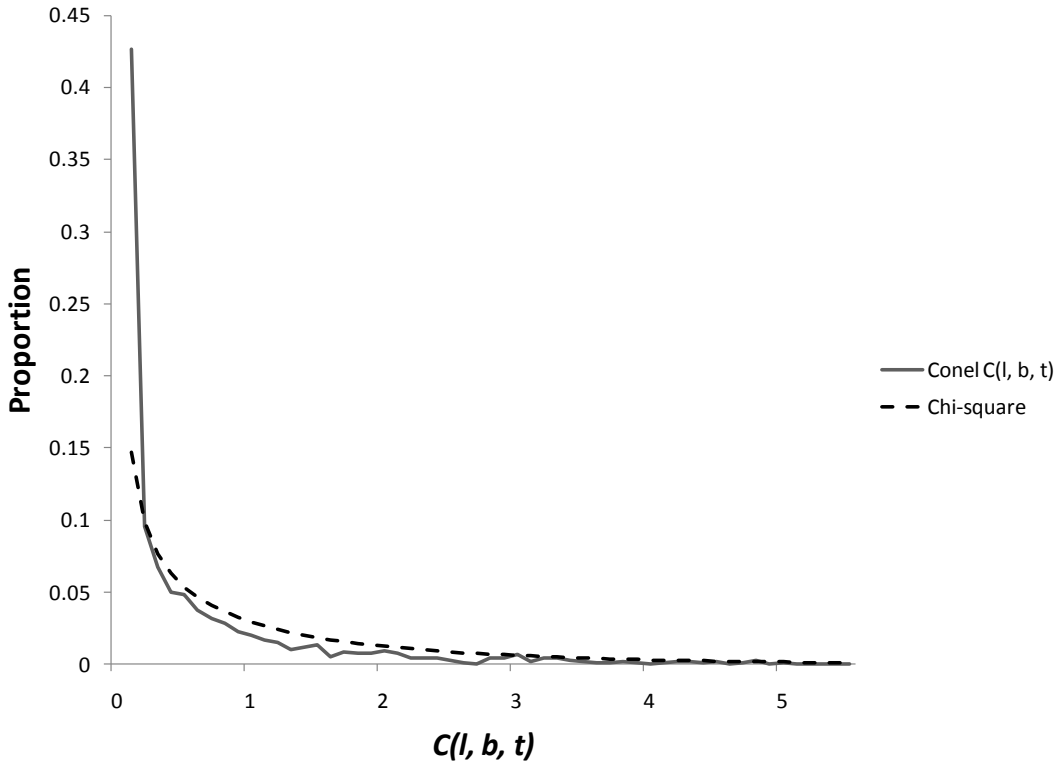


Figure 3-2. Distribution of normalized change interval magnitudes. Distribution of *C(l, b, t)* values for all seven change intervals. *N = 1554* overall, *N = 9* for *C(l, b, t) > 5.5* (not shown). Histogram bins are in increments of 0.1 units. Rather than the predicted chi-square distribution with one degree of freedom (also shown), the histogram in this figure is well-described by a gamma distribution (*p = 0.99997*, Kolmogorov-Smirnov test; parameters $\alpha = 0.3638$, $\beta = 1.6570$ and *mean = 0.6028*).

3.1.3 K-means cluster analysis of neuron population change vectors

Chapter 2 showed that neuron population change vectors that result from respective sequences of change intervals for each address *(l, b)*, are optimally separated into seven k-clusters. Figure 3-3 shows the distribution of all 222 addresses into these seven k-clusters. The number of each cluster is arbitrary, assigned by the k-means algorithm during 1000 replications as described in Chapter 2.

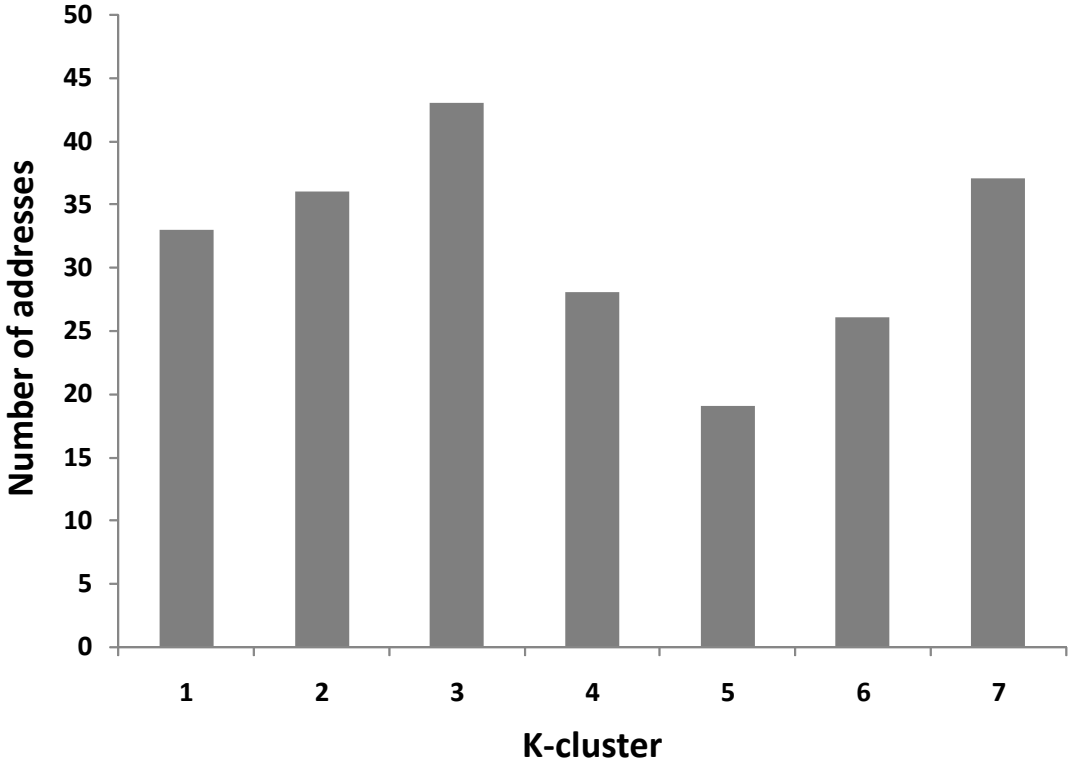


Figure 3-3. Distribution of neocortical addresses into seven k-clusters (N = 222).

Chapter 2 showed that the optimal number of clusters for the k-means algorithm was 7. The distribution of the 222 change vectors in Conel's data from Equation 8 was then

sorted into seven k-clusters by the algorithm using Equation 9 as the distance metric for cluster membership and separation of the clusters. Addresses with the most highly correlated changes in their respective values for *C(l, b, t)* were assigned to the same k-cluster. Cluster number was assigned arbitrarily by the k-means algorithm.

A maximal clique in a k-cluster is the largest set of totally connected addresses in the cluster. Defining a connection as a relationship between addresses *x* and *y*, where $D(x, y) < 0.1$ in Equation 9 in Chapter 2 (or correlation $r_{xy} > 0.9$), 11 of the 24 outliers in Table 3-1 belong to the maximal clique in their respective k-cluster, indicated by an asterisk (sample correlation $r = 0.8192$, $p = 0.0242$, two-tailed t-test against the null hypothesis that the distribution by change interval of the 11 maximal clique members and the 24 significant *C(l, b, t)* values are from different distributions).

Table 3-1. Statistically significant values of *C(l, b, t)*.

1 Mo	29 LI	29 LIV						
	4.4938	4.1803						
3 Mo	17 LIII	17 LV	**38 LI***	4T LV				
	10.8246	7.0145	4.9739	4.239				
6 Mo	**29 LVI***	17 LIII	**17 LI***					
	16.6004	6.4275	6.254					
15 Mo	**4P LVI***	**29 LV***	**18 LII***	41 LI	**8 LIII***	**4Hd LVI***	10 LI	**21 LII***
	7.1081	6.6642	6.5119	4.8475	4.8088	4.7902	4.4962	4.3299
24 Mo	**14 LI***	**42 LVI***						
	5.7625	4.442						
48 Mo	45r LI							
	4.7949							
72 Mo	45r LI	10 LI	22 LVI	23/31 LIII				
	4.7137	4.4902	4.267	4.0839				

Change interval, address (Brodmann Area and layer) and *C(l, b, t)* value are indicated. Asterisks denote the addresses that belong to a maximal clique in a k-cluster. Nine addresses with $C(l, b, t) > 5.5$ were not included in Figure 3-2. Altogether, 24 values of *C(l, b, t)* were statistically significant ($p < 0.05$, one-tailed test against chi-square cdf, one degree of freedom). Abbreviations: T = trunk, P = paracentral lobule, Hd = head, r = rostral.

Figure 3-4 shows the mean value of *C(l, b, t)* for each of the seven k-clusters as a radar plot graph. There is a one-to-one correspondence between a maximum change vector component in a k-cluster and each of the seven change intervals. To test whether the single spike per k-cluster pattern in the radar plot was an artifact of the k-means method applied to seven change intervals, an empirical cumulative distribution function (cdf) was constructed using a Monte Carlo process using 1000 iterations, which is more than sufficient (Gentle, 2002). Each iteration began with a 222 x 8 table of random numbers drawn from the unit interval, corresponding to a null hypothesis that neuron populations are equiprobable. Equations 1 to 3 from Chapter 2 were then applied to the table and the resulting change vectors were clustered using the k-means algorithm with k = 7 using Equation 9 to define the distance metric. Let $m(k,t) = mean[C(l, b, t)]$ for each of the k-clusters in a given change interval resulting from these calculations. After sorting, the statistic $s(t) = max[m(k,t)]/mean[m(k,t)]$ characterizes each change interval in the graph, and *mean(s)* to characterizes each graph across all change intervals. Using the

results from 1000 iterations, the surrogate cdf then allowed us to estimate a p-value for the actual Conel data ($p \approx 0.001$).

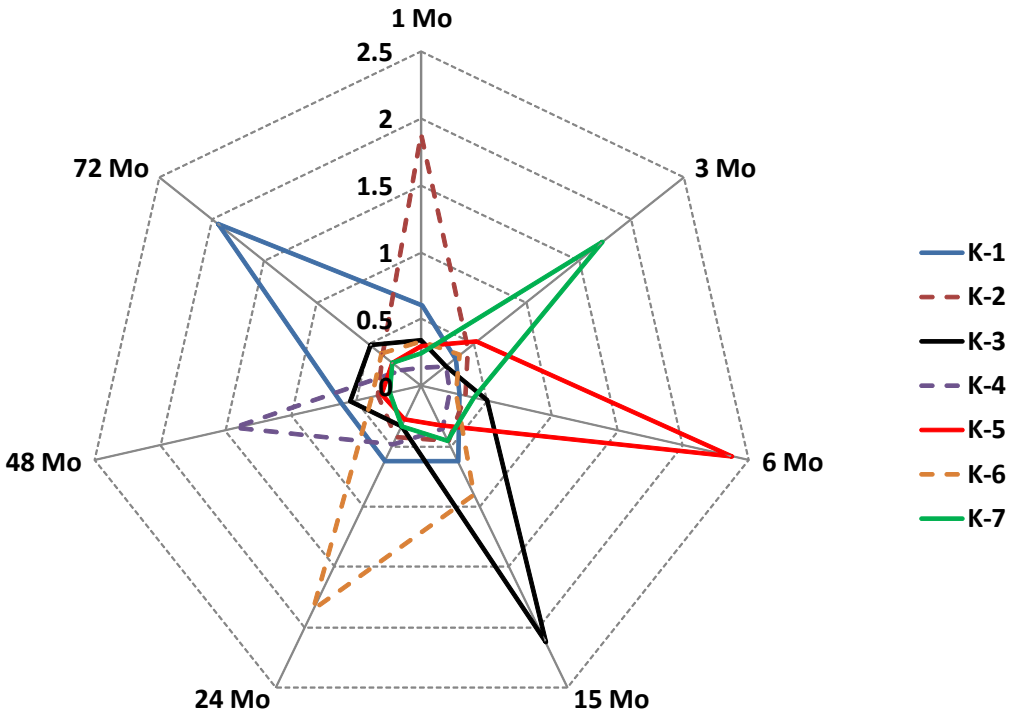


Figure 3-4. Correspondence of k-cluster to change interval. Radar plot axes indicate mean *C(l, b, t)* values for the respective k-clusters. Radial axes are the seven change intervals. Each of the k-clusters has a maximum on a unique change interval axis. Using a Monte Carlo process to construct an empirical cumulative probability distribution function (cdf) for randomly generated graphs of seven k-clusters using a "mean max-over-mean" statistic for each graph, the relative size of the maxima and their one-to-one distribution across the change intervals in this figure is statistically significant ($p \approx 0.001$ from the empirical cdf; the sample mean Monte Carlo graph statistic was 2.845 ± 0.104

(sample standard deviation for $n = 1000$). The same statistic for the Conel data was 3.180).

3.1.4. Relational network diagrams and maximal cliques for the seven neuron population k-clusters

Figure 3-5 shows the relational network for the first k-cluster (K-1), where D < 0.1. The figure depicts all members of K-1. Defining the D < 0.1 criterion is a relationship between two addresses, Figure 3-5 also illustrates each of those relationships. There is only one totally connected maximal clique (with nomenclature from Conel): BA 6 LVI (posterior supplemental motor area of the gyrus frontalis superior), BA 45 LIV (inferior gyrus frontalis medialis /pars triangularis), BA 22 LI (area temporalis superior), and BA 24 LIII (area limbica anterior agranularis of the gyrus cinguli).

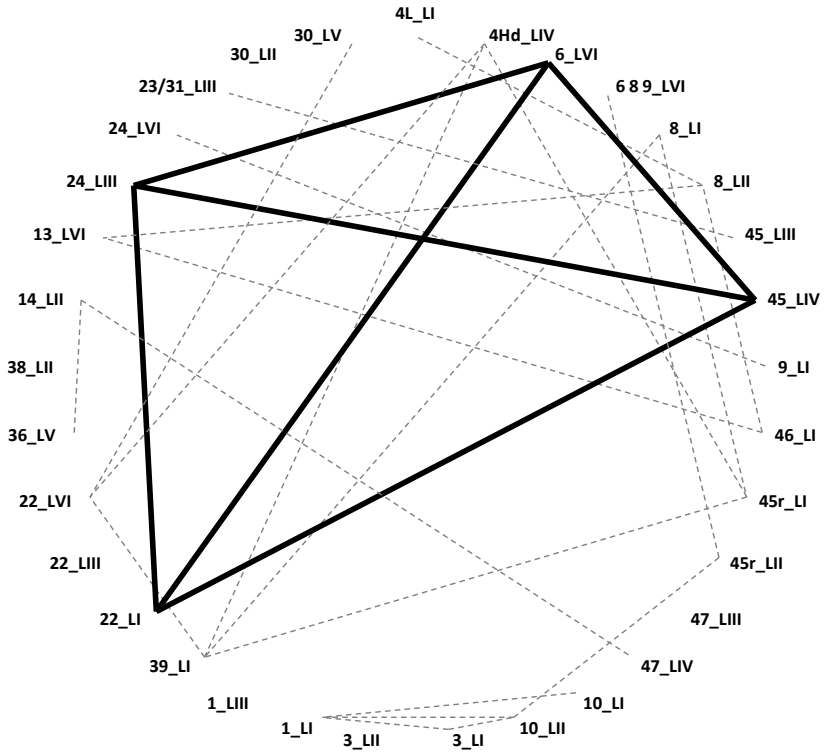


Figure 3-5. Relational graph and maximal clique of the first k-cluster (K-1) for *D* < *0.1*. The 33 addresses in K-1 are listed in clockwise order from BA 4L Layer I at the top to BA 30 Layer V. The order is the one used by Conel, generally proceeding from frontal lobe to the parietal, occipital, temporal lobes, and then medial cortical areas. Lines show all instances where $D < 0.1$. The bolded solid lines show the maximal clique for the cluster, which has size four. Only K-1 and K-4 have single maximal cliques of size 4. K-5 and K-6 have more maximal cliques of size 4. K-2 and K-7 have two overlapping

maximal cliques of size 6, and K-3 has two overlapping maximal cliques of size 9. Abbreviations: L = leg; Hd = head; r = rostral.

The "core" of a k-cluster is defined to be the set of addresses contained in its maximal cliques. In the case of K-1, the core consists of a single maximal clique of size 4. Figure 3-6 reveals the *C(l, b, t)* values for the core, as well as the mean core values compared to the entire cluster. The mean cluster *C(l, b, t)* values are the same as those as in Figure 3-4, where the maximal values for *C(l, b, t)* occur in the 72 month change interval.

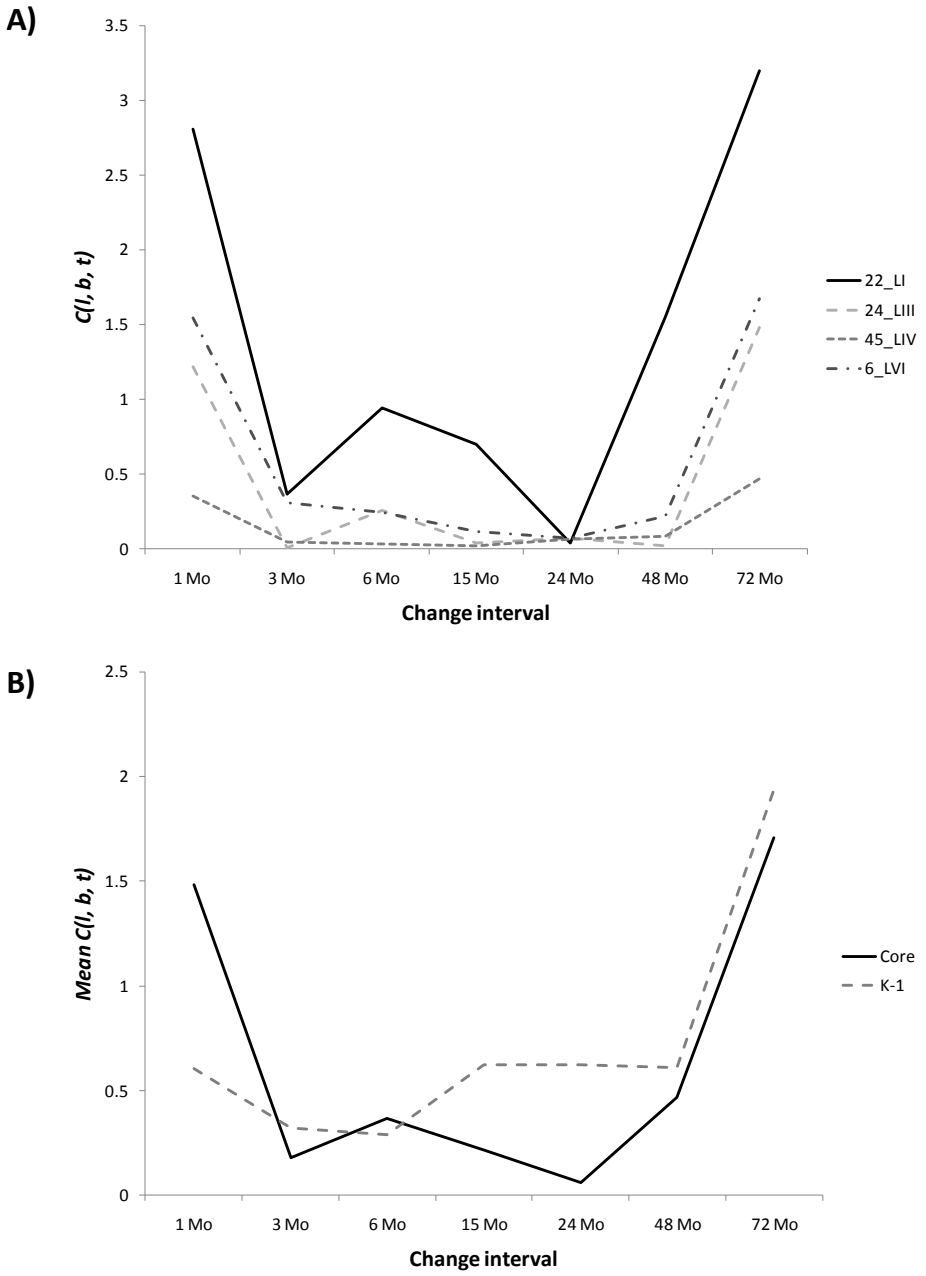


Figure 3-6. *C(l, b, t)* values for K-1 and the K-1 core. A) C(l, b, t) for the four members of the single maximal clique. This is a graph of the entire trajectory for each of the four addresses in the K-1 core, by *C(l, b, t)* and change interval. The secondary peak

C(l, b, t) values for 1 month provide a relatively complex overall trajectory for this group of addresses. B) Mean *C(l, b, t)* for the core compared to mean *C(l, b, t)* for the entire cluster. The mean trajectory for the entire k-cluster has a much lower value for 1 month compared to the mean for the core addresses, making the core addresses a relatively distinct sub-set of the k-cluster that has correlated maximal *C(l, b, t)* values at 72 months.

The largest k-cluster in the Conel data is K-3. It also has the largest core, with two maximal cliques of size 9. Eight of the addresses in these maximal cliques overlap, as shown in Figure 3-7. Thus, the core contains 10 addresses. The "unique" nodes on this graph are BA 29 LV (area retrosplenialis granulosa) and BA 8 LIII (medial gyrus frontalis superior in von Economo's area frontalis intermedia). They are "unique" in the sense of being totally connected to eight other addresses, but not to each other.

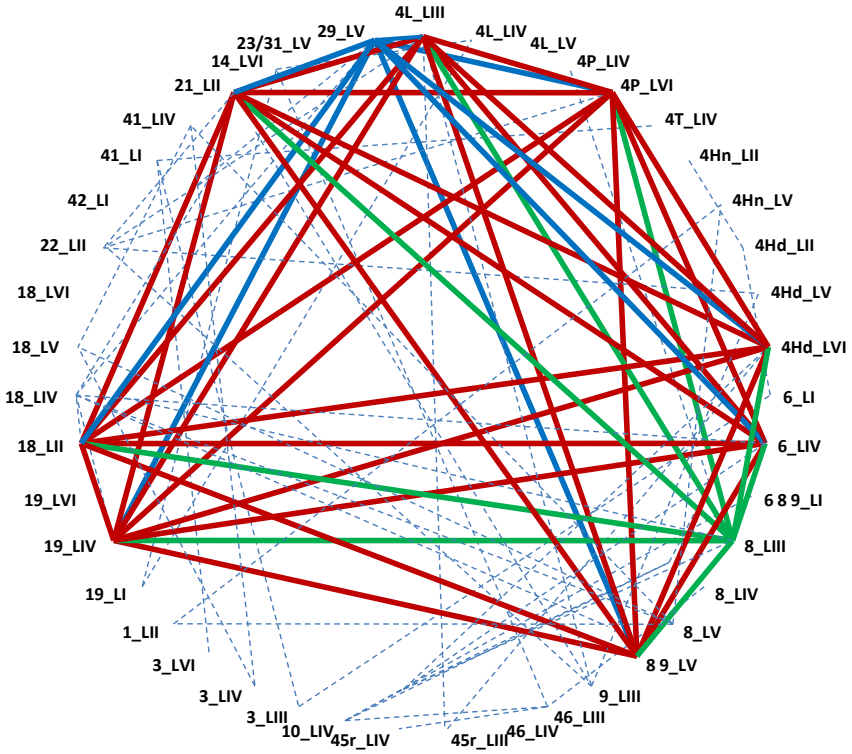
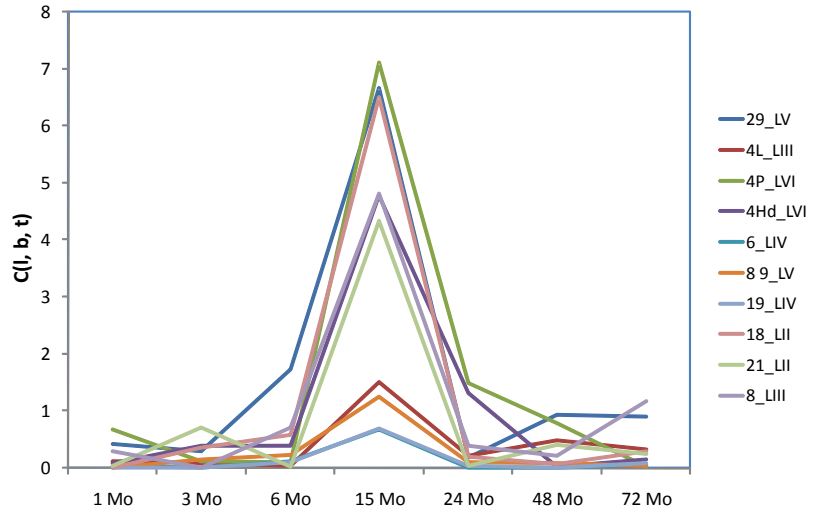


Figure 3-7. Relational graph and maximal cliques for K-3 for *D* < *0.1*. The 43 addresses in K-3 are listed in clockwise order from BA 4L Layer III at the top to BA 29 LV. The order is the one used by Conel. Solid red lines indicate relationships common to both of the two maximal cliques in the K-3 core. Solid blue lines show the relational links for BA 29 Layer V, which is the unique address in one of the two maximal cliques. Solid green lines show the relational links for BA 8 Layer III, the unique address for the second of the two maximal cliques. Including common and unique nodes, the two maximal cliques have size nine. Dashed blue lines indicate the remaining correlation distances with D < 0.1. Abbreviations: P = paracentral lobule, T = trunk, Hn = hand.

A)



B)

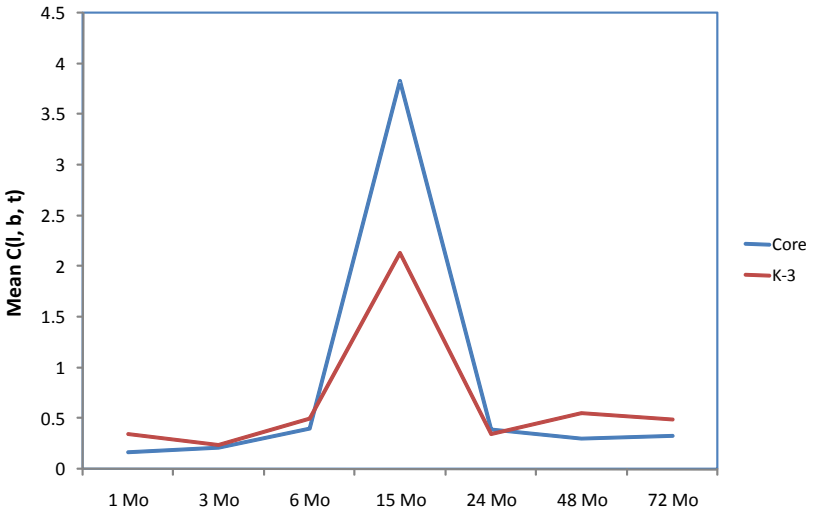


Figure 3-8. *C(l, b, t)* values for K-3 and the K-3 core. A) *C(l, b, t)* by change interval for the ten change vector trajectories of the core addresses in K-3. BA 29 LV and BA 8 LIII are the unique nodes in Figure 7. The eight common addresses in the core are listed

between them on the right. The correlated maximum *C(l, b, t)* values at 15 months occur at multiple scales. B) Mean *C(l, b, t)* for the core addresses compared to the mean *C(l, b, t)* values for the entire k-cluster by change interval.

Figure 3-8 shows the core *C(l, b, t)* and mean *C(l, b, t)* values for K-3. The two maximum values in Figure 3-8A, BA 4P LVI (the paracentral lobule portion of BA 4) and BA 29 LV are both outliers in Table 3-1. However the correlated trajectories in the K-3 core also include BA 19 LIV (area peristriata) which also has a maximum of 0.7003 at 15 months. The remaining core areas have *C(l, b, t)* maxima at 15 months between 0.7003 and 7.1076, reflecting correlations on multiple change scales. The comparison of mean *C(l, b, t)* values for the core and the k-cluster as a whole in Figure 3-8B reflects greater homogeneity among the change vector trajectories than in K-1.

Maximal clique sizes and descriptions for all clusters are summarized in Table 3-2. There are seven maximal cliques of size four distributed among four of the k-clusters. There are four maximal cliques of size six, two in K-2 and two in K-7. There are two maximal cliques of size nine in K-3. Thus, there are 47 addresses in the seven k-cluster cores. They belong to 27 Brodmann Areas. BA 6 has four addresses in different cores. No other Brodmann Area has more than two addresses in any core.

Table 3-2 Summary of k-cluster cores.

Cluster	Maximum clique size	Number	Overlap	Partial overlap	Non-overlap	Maximum C(l, b, t) in change interval:
1	4	1	6 LVI 45 LIV 22 LI 24 LIII	--	--	72 Mo
2	6	2	46 LV 45r LV 40 LIII 37 LI 42 LIII	--	41 LIII 10 LIII	1 Mo
3	9	2	4L LIII ***4P LVI**** ***4Hd LVI**** 6 LIV 8 9 LV 19 LIV ***18 LII**** ***21 LII****	--	***29 LV**** ***8 LIII****	15 Mo
4	4	1	7 LIII 40 LIV 22 LIV 21 LVI	--	--	48 Mo
5	4	3	***29 LVI (gp 1/2/3)****	17 LIV (gp 1/3) ***17 LI (gp 1/3)****	37 LII (gp 1) 14 LIII (gp 2)	6 Mo

				45r LVI (gp 2/3)	7 LII (gp 2)	
6	4	2	--	--	44 LV (gp 1)	24 Mo
					45 LVI (gp 1)	
					42 LVI (gp 1)*	
					13 LI (gp 1)	
					4P LI (gp 2)	
					41 LV (gp 2)	
					38 LVI (gp 2)	
					14 LI (gp 2)*	
7	6	2	6 LIII	--	44 LII	3 Mo
			6 LV		39 LVI	
			39 LIV			
			38 LI*			
			13 LIV			

Addresses in bold and marked by an asterisk indicate C(l, b, t) > 3.9 ($p < 0.05$, chi-square test). For K-5, there are three maximal cliques of size four, each partially intersecting the others. The three components are grouped as "gp 1" to "gp 3." K-6 has two maximal cliques of size four that do not intersect, grouped as "gp 1" and "gp 2." Abbreviation: L = leg.

3.2 Discussion of results from neuron population change vector analysis

The principal result in this chapter is the one-to-one correlation between large change steps in successive neuron populations and observation age in human neocortical development to age six. Change steps are normalized measures of relative neuron population shares at individual area/layer addresses. The pattern of synchronized change in Figure 3-4 emerges when 222 neurodevelopmental trajectories for each of the six cortical layers in 37 Brodmann areas or subareas are parsed into optimal k-clusters using k-means cluster analysis with the correlation distance as the metric. In this section, we will discuss the members of these k-clusters in terms of two distinctive types of members: those individual addresses that have a statistically significant magnitude of change as defined by the expected chi-square distribution of change steps from random mixtures (i.e., $C(l, b, t) > 3.9$), and those addresses whose developmental trajectories define them as part of a totally connected, maximal clique in a given k-cluster (Table 3-2). As noted at Table 3-1, maximal clique membership and large magnitude change steps are significantly correlated.

Previous neuroinformatic analysis of Conel's data (Shankle and Romney, et al., 1998, Shankle, et al., 2002) employed the statistical methodology known as correspondence analysis across all six types of data reported in Conel: layer width and neuron density (also used in this chapter) as well as somal height and width, large fiber density resulting from Cajal staining and myelinated fiber density resulting from Weigert staining (Shankle and Romney, et al., 1998; used in the next chapter). Their method is

scale-free and shows a regular developmental pattern from birth to 72 months, with 99-percent confidence intervals for each cortical area in neocortex overlapping almost completely (Shankle and Romney, et al., 1998). By contrast to Shankle et al., the *C(l, b, t)* transformed measurement in the change vector method is normalized across cortical areas and layers and between ages, but is not scale-free.

The evidence for synchronized change in the Conel data comes from three arguments. First, the relative magnitude of the radial spikes in Figure 4 is not likely an artifact of randomly distributed neuron populations. As described in section 3.1.3 above, we tested the contrary by making 1000 random draws from the unit interval for 222 surrogate addresses and 8 surrogate observation ages, used Equations 1 to 9 to construct 7 k-clusters, and measured the relative size of the spike for each k-cluster to the average change step magnitude for that k-cluster. Comparing the Conel result to the empirical cdf from this process, the estimated probability of an artifact is rejected with $p \approx 0.001$. Second, we expect a chi-squared distribution with one degree of freedom to emerge from Equations 1 and 3 based on the reasoning and error estimates in Chapter 2. This contrasts with Conel's data in Figure 3-2, where the distribution mean for *C(l, b, t)* across all ages is 0.6028 rather than 1 expected from the chi-square distribution, and the KS test for goodness-of-fit to the chi-square is rejected with $p = 0.0112$. Finally, The KS test was performed against the null hypothesis that the normalized neuron population *C(l, b, t)* is drawn from the same distribution as the previous *C(l, b, t – 1)* at each address. Occurrences of *C(l, b, t)* that negate this null hypothesis are the statistically significant change step magnitudes that co-occur within the same change step. That is, the null

hypothesis is negated by statistically significant synchronized changes to relative neuron population for successive groups of addresses during development. While neuron counts alone are silent as to a potential mechanism, these synchronized changes visible in Figure 3-4 at specific addresses that are significantly correlated with k-cluster cores ($r = 0.8192$, $p = 0.0242$, two-tailed t-test) imply a clock-like process that produces the synchronization.

The observed synchrony in Figure 3-4 is not between selected areas and age, nor selected layer and age, but rather between specific addresses and age. Thus, only BA 6 has four addresses that appear in the maximal clique catalog at Table 3-2, distributed across three k-clusters. Only BA 17 and BA 29 have four addresses, likewise distributed across three k-clusters, in Table 3-1 where large magnitude change steps and maximal clique membership are crosswalked. Thus, in addition to area-to-area adjustments known to be significant in postnatal development (Verney, et al., 2000), the data also indicate the continued importance of the layer-dominated mammalian prenatal pattern (Butler and Hodos, 2005; Noden, 1991; Takahashi, et al., 1999, etc.), revealing a highly conserved evolutionary pattern (Finlay, et al., 2001).

BA 6 is supplementary motor cortex , BA 17 is primary visual cortex (Gazzaniga, et al., 2002), and BA 29 (granular retrosplenial cortex) may be involved with the integration of egocentric inputs from the parietal lobe and allocentric representations from the hippocampus and medial temporal lobe (Burgess, 2008). This suggests a potential functional role for maximal cliques and the set of unusually large change step addresses that we will take up next.

As noted with Figure 3-1, the largest absolute change in raw neuron populations in Conel's data is the drop of 95.9 neurons per sample column that takes place during the change step from 3 to 6 months. The largest relative drop across the cortex (24.55 % per sample column) takes place during the next change step from 6 to 15 months. After normalization, the largest values for *C(l, b, t)* likewise occur during these change steps, as revealed in Table 3-1. K-3, the largest k-cluster with the largest maximal cliques has its maximal changes during the 6 to 15 month change step. Taking neuron cell death as symptomatic of assimilation and commitment (Stiles, 2008), these changes should be consistent with corresponding sensitive periods (e.g., Marcotte and Morere, 1990; Ross and Bever, 2004; DiCristo, et al., 2007).

In humans, these sensitive periods would reasonably be correlated with Piaget's sensorimotor stage, which extends from birth to 2 years (Piaget, 1952), particularly since this is the Piaget stage that is least subject to environmental factors such as family income or education (Renner, et al., 1976). In this context, the change step for 6 months is not surprising. Its three maximal cliques all involve BA29 LVI which has the largest magnitude for *C(l, b, t)* of any address at any change step. Two of the three maximal cliques contain primary vision address BA17 LI, which also has a large magnitude change value, as well as BA17 LIV. BA17 LIII, associated with color vision processing (Gazzaniga, et al., 2002) has statistically significant *C(l, b, t)* magnitudes in this change step as well as the prior change step at 3 months, where it is a member of the k-cluster. According to Piaget (1952), vision is not especially important during the reflex stage immediately after birth, but infants begin to "really look" at things after 1 month, and

actively track objects by 2 months. These remaining associations to 6 months remain consistent with continued adaptation and assimilation as called for in Piaget (1952).

This sensorimotor theme continues into the 15 month change interval, where the maximal cliques include addresses from primary motor cortex (4L LIII, 4P LVI, 4Hd LVI; Gazzaniga, et al., 2002), supplementary motor cortex (BA6 LIV), secondary vision (BA19 LIV and BA18 LII), and the area associated with the frontal eye fields (BA8 LIII; Berman, et al., 1999). Retrosplenial cortex also appears in this set (BA29 LV). Primary audition (BA41 LI; Gazzaniga, et al., 2002) appears among the large magnitude *C(l, b, t)* set as well. Six of the ten maximal clique members also have large magnitude change steps.

Mandler (1999) contends that such pre-linguistic primitive steps, particularly kinetic "image-schemas" provide a sufficient conceptual basis on which to build language proper, including such abstract notions as agents, patients, permanent objects, and causation. Tracing backwards from childhood and adult imaging studies to identify key cortical areas for language functions, the k-cluster evidence from Conel indicates that these foundations are sometimes laid extremely early during development. Broca's Area (BA 44 and 45), associated with language (Broca, 1861a), appears in three cores with peak *C(l, b, t)* ranging from 3 months (K-7), very early in language acquisition, to 24 (K-6) and 72 months (K-1), when children can speak increasingly sophisticated sentences (Tomasello, 2003). The earliest of these cores associates BA 44 LII with BA 6 LIII and LV, BA 13 LIV, BA 39 LIV and LVI (area angularis) and BA 38 LI (area temporopolaris). These additional parietal and temporal association areas also show

significant activations in fMRI during verb generation tasks (Le Bihan, et al., 1993; Cao, et al., 1999). BA 39 is associated with Broca's Area in word-picture matching among children from 5 to 18 years old (Schmithorst, et al., 2007), and semantic processing (Binder, et al., 1996). Moreover, BA 39 and superior frontal gyrus, from which the BA 6 samples were drawn in Conel's subjects, form part of a complementary network to the human mirror system, which focuses on the "end" or goal of an action, while the human mirror system appears to focus on the "how" of an action (Hesse, et al. , 2009). In K-7, BA 44 LII is one of the two non-overlapping addresses, and BA 39 LVI is the other. BA 44 is a recipient of mirror neuron projections (Cadoret, et al., 2005). The portion of BA 6 that would have been part of the "classical" human mirror system (Hesse, et al., 2009) is von Economo's area FBop, which was not included in Conel's data. In all of these cases, the timing of the maximum change for the change vectors in K-7 is the 3 month change interval. This is well before most overt manifestations of language or mirror functions, and in agreement with Mandler (1999).

In the remaining cores associated with Broca's Area, at 24 months BA 44 LV and BA 45 LVI co-occur with secondary auditory cortex (BA 42, Gazzaniga, et al., 2002) and BA 13 (area postcentralis insulae). Direct electrostimulation connects Broca's Area with BA 42 (Matsumoto, et al., 2004), as does Bayesian analysis of BOLD signals from fMRI for speaker and sentence distinction tasks (Patel, et al., 2006). Broca's Area and BA 13 are associated in the retention phase of a delayed match-to-sample working memory task (Habeck, et al., 2005). Similarly, at 72 months, BA 45 LIV is associated with Wernicke's Area (BA 22 LI; Wernicke, 1874), supplementary motor (BA 6 LVI,

Gazzaniga, et al., 2002) and anterior cingulate cortex (BA 24). Broca's Area and supplementary motor areas are associated in functional magnetic resonance imaging (fMRI) for noun generation (Cuneod, et al., 1995; McCarthy, et al, 1993), and orthogonal lexical retrieval (Blacker, et al., 2006). BA 6 and anterior cingulate cortex are associated for the stimulus and retention phases of delayed match-to-sample working memory (Habeck, et al., 2005). BA 22 is also involved in the retention phase (Habeck, et al., 2005; Woodward, et al., 2006). Broca's and Wernicke's Area are both involved in word-picture matching in children between 5 and 18 years old (Schmithorst, et al., 2007). Broca's area is associated with supplementary motor and anterior cingulate cortex in morphological tasks related to irregular verbs as well (Sahin, et al., 2006). In direct electrostimulation of the cortex in epileptic patients, Broca's and Wernicke's Areas demonstrate bidirectional connectivity (Matsumoto, et al., 2004).

All of these cited studies associate additional areas outside the k-cluster cores, and none has the resolution of an individual cortical address. However, these functional circuits in adult and older juvenile subjects do correspond to the present correlated changes in Conel's neuron population data. We suggest that these patterns in human postnatal neocortical development reflect conserved mechanisms for synchronized functional change steps in postnatal mammalian brain development that are later manifested in behavior.

Keeping in mind the small sample sizes in Conel's work at each observation age (see Chapter 2) despite the detail he reported at each age, we suggest that similar results can be obtained from animal models. For example, paternally expressed gene 3 (Peg3)

mutant mice have a range of deficits, including olfactory function. Inactivation of this gene reduces the incidence of caspase 3 positive cells at 4 to 6 days postpartum, indicating a reduction in neuron apoptosis. Affected areas in these mice include the bed nucleus of the stria terminalis, nucleus accumbens, caudate putamen, medial pre-optic area, arcuate nucleus, medial amygdala, anterior cortical and posteriodorsal amygdaloid nuclei (Broad, et al., 2009). We hypothesize that at least some of these areas would constitute a maximal clique in mice between P4 and P6 if the analogous murine values for *C(l, b, t)* were calculated from birth to an observation age greater than P6. The associated deficits would emerge later as the mice develop.

4. Synchronized change to myelinated fiber density during human neurodevelopment

Chapter 3 showed that human neocortical neuron population changes from birth to six years demonstrate a series of synchronized changes consistent with sensorimotor developmental stages. Functionally-relevant synchronized change is a necessary condition for some form of clock mechanism. This is bolstered by the evidence in Renner, et al., (1976) that environmental conditions have the least effect on sensorimotor development, leaving genetic and epigenetic factors under the clock hypothesis introduced in Chapter 1. A similar demonstration of synchronized change in Conel's myelinated fiber data would provide more support for an underlying genetic/epigenetic developmental clock, since environment effects would be identical for both sets of data.

As discussed in Chapter 3, neuron population change is likely to be neurotrophin and activity-based (Huang and Reichardt, 2001). While the molecular mechanisms differ, the common locus for genetic/epigenetic and activity-based signals for synchronized change to myelinated cortical axons remains within the neuron. First, the progress of axon myelination in the cortex depends on oligodendrocyte development, controlled in turn by neuron secretions such as insulin-like growth factor and fibroblast growth factor, which play different roles in oligodendrocyte differentiation, survival, and

development (Butt and Berry, 2000). In addition, cell adhesion molecules differentially suppress or enhance myelination at different times during development (Coman, et al., 2005). Finally, myelination is promoted by electrical activity in the neuron (Sanchez, et al., 1998; Sirevaag and Greenough, 1987; Szeligo and Leblond, 1977).

Generally, the completion of myelination is a signal of the onset of full functional capacity (Flechsig, 1920; Yakovlev and Lecours, 1967). Collectively, these factors interact so that different cortical areas begin to myelinate at different times during development and have different characteristic myelination rates (Yakovlev and Lecours, 1967; Sampaio and Truwit, 2001). Thus, Conel's myelinated fiber data is an appropriate set for the change vector method, as well as a useful test for the existence of a functionally-related clock mechanism.

As discussed in Chapter 2, the neuron population and myelinated fiber data are not directly comparable. Conel's neuron data come from the left hemisphere, while the myelinated fiber data come from the right. Moreover, Conel provided layer depth information so that neuron population could be calculated for each of the six layers in the neocortex, whereas Conel reported his myelinated fiber data in ten layers where there is no layer depth information available so that only density information is available. Nevertheless, the change vector method can be applied to the myelinated fiber data. This enables a comparison of the two data sets at the level of change step populations, as well as a comparison of prospective functions when the respective neuron and myelinated fiber k-clusters are compared at each developmental age. This will reveal further

evidence for a functionally-related mechanism to synchronize cognitive development that is located within the neuron.

4.1 Myelinated fiber results

4.1.1 Density increases monotonically with age, rank-ordered by layer

Figure 4-1 shows a general increase in myelinated fiber density with age in human cortical development. Figure 4-1A reveals that increases in mean myelinated fiber density vary by Brodmann Area and that the highest densities at one month are not necessarily the same areas with the highest densities at 72 months. By contrast, Figure 4-1B shows that with the exception of Layer 1 and Layer 5a, the onset of myelination and increase in mean density is greatest progressing from vertical fibers that exit the cortex, followed by subcortical connections, and then progressively from inferior to superior layers up to Layer 2 in the cortex. Including the exceptions, the rank order of mean density by layer is essentially retained in the order in which the respective layers begin to show an increase.

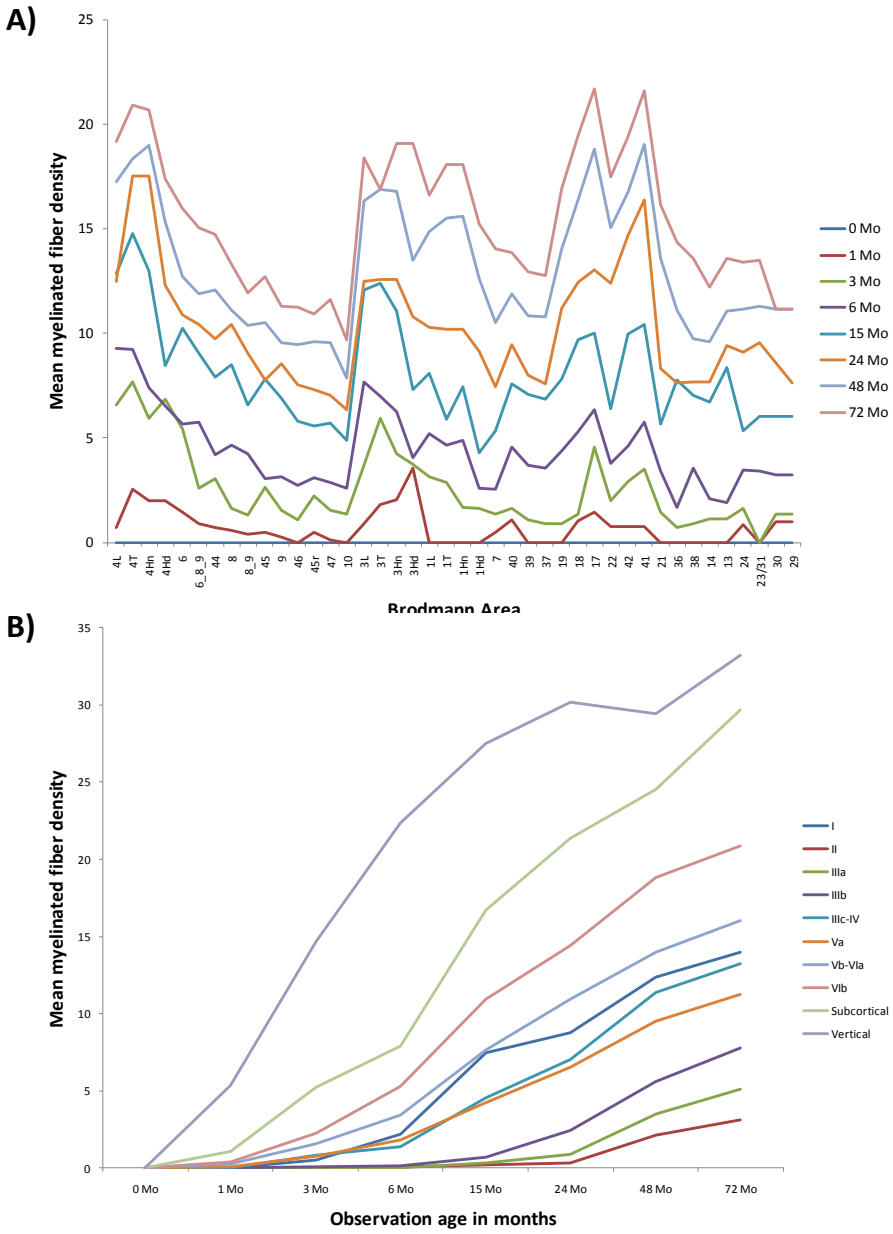


Figure 4-1. Mean myelinated fiber density by layer during human neurodevelopment. Myelinated fiber density is Conel's mean count of Weigert-stained axons present in 50 x 50 x 25 µm samples at each observation age. At birth, no

myelination is observed. A) Mean myelinated fiber density by Brodmann Area and observation age. At 1 month, the highest mean density occurs at BA 3Hd, followed by BA 4T, BA 3Hn, and BA 4 Hn and Hd with the fourth highest. BA 17 has the 7th highest mean density. By 72 months, BA 17 has the highest mean density in the cortex, followed by BA 41, BA 4T, BA 4Hn and BA 19. BA 3Hd has the 8th highest mean density, tied with BA 3Hn; BA 4 Hd has the 14th highest density. B) Mean myelinated fiber density by layer and observation age. Generally, myelination onset and myelination rate are greatest for vertical axons that exit the cortex ranging to the latest onset and slowest rate for Layer 2. The onset and myelination rate for Layer 1 and Layer 5a do not fit this sequence. Layer 1 increases faster and falls between Layer 3c-4 and Layer 5b-6a. Layer 5a falls between Layer 3c-4 and Layer 3b. Abbreviations: Hd = head, Hn = hand, L = leg, T = trunk, r = rostral.

4.1.2 Mean C(l, b, t) for myelinated fiber density decreases exponentially with observation age

Chapter 2 showed that the optimum number of k-clusters for neuron populations and myelinated fiber densities was the same: seven. Figure 4-2 compares mean *C(l, b, t)* for neuron population k-clusters and the k-clusters for myelinated fibers. Whereas the mean value for neurons is approximately constant for all seven clusters ($R^2 = 0.114$ which means that the change interval explains only 11.4% of the variation in mean *C(l, b, t)* about a mean defined by a line with slope -0.0038), a least squares fit of the myelinated

fiber data to an exponential decrease shows the two are well-correlated ($R^2 = 0.9131$, with change interval explaining 91.31% of the variation).

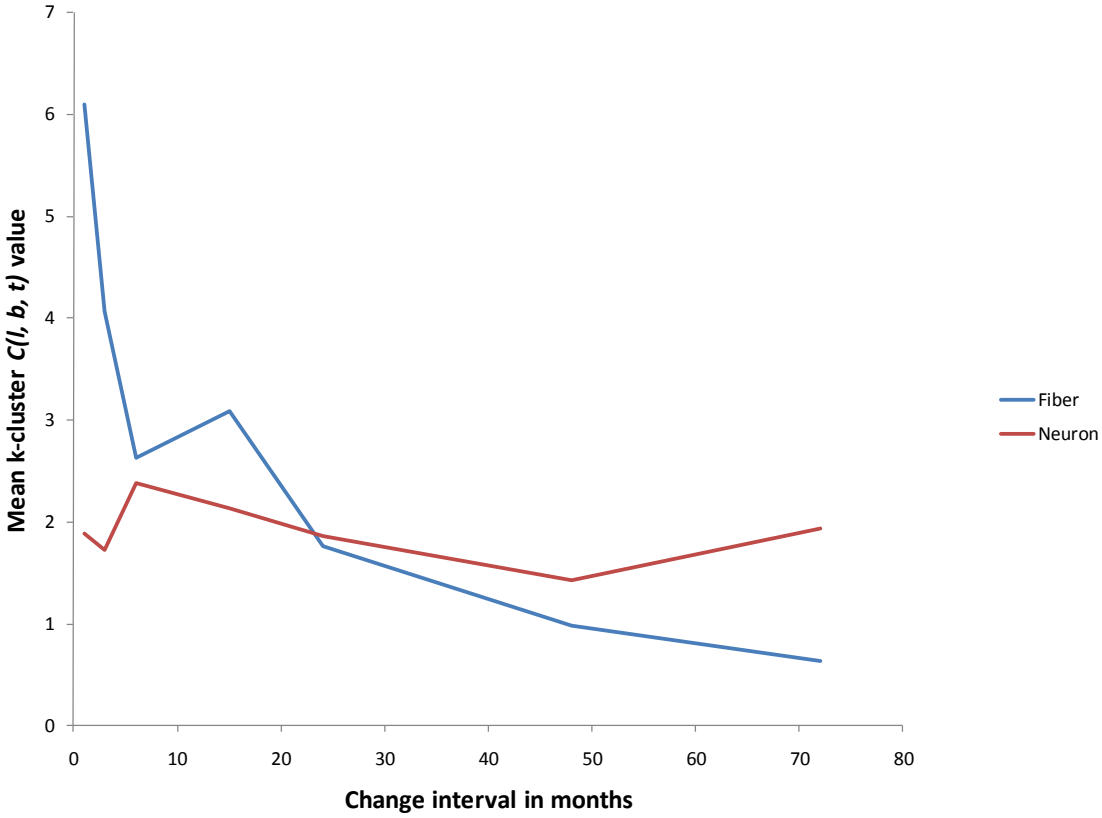


Figure 4-2. Mean k-cluster *C(l, b, t)* values do not change identically with age for neuron population and myelinated fiber density changes. Regression using least squares reveals a nearly horizontal line for change to neuron *C(l, b, t)* values, with $y = -0.0038x$. By contrast, mean k-cluster *C(l, b, t)* for myelinated fiber density is approximately $y = 4.2131\ exp(-0.029x)$, where y is the mean k-cluster *C(l, b, t)* value, and x is the change interval in months.

4.1.3 Myelinated fiber changes are subgranular from birth and shift to supragranular after the 15 month change interval

Of the 420 addresses in Conel's myelinated fiber data, Conel reported no fibers in five through 72 months: BA 45r 2 and 3a, and BA 10 2, 3a and 3b. As noted in Chapter 2 and Figure 4-2, the remaining 415 addresses fall into seven k-clusters, each with a maximum k-cluster mean *C(l, b, t)* in a unique change interval. When the k-clusters are organized according to the change interval with maximum *C(l, b, t)*, Figure 4-3 shows that maximal changes are subgranular (inferior to Layer IV) initially, and shift to supragranular after the 15 month change interval. Figure 4-3 does not depict *C(l, b, t)* data for Layer 3c-4, the closest approximation to neuron Layer 4 in Conel's data, which will be described below.

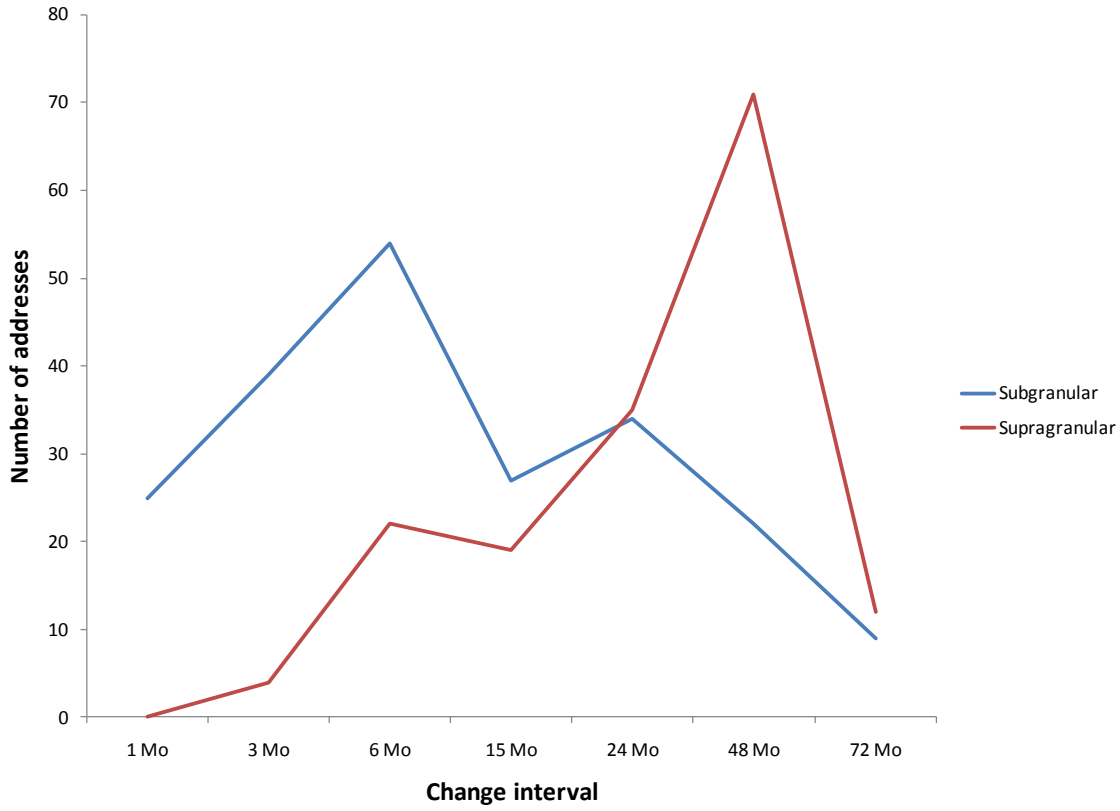


Figure 4-3. Number of sub- and supragranular addresses in myelinated fiber k-clusters. The 373 members of k-clusters that are superior to Layer 3c-4 (supragranular) or inferior to Layer 3c-4 (subgranular) are counted in the change interval for which the mean k-cluster *C(l, b, t)* value is maximal, as in Figure 4-2 for total k-cluster membership. Subgranular changes are predominant until the 15 month change interval, then supragranular changes thereafter. The five supragranular addresses with no fiber density reported to 72 months remain unallocated.

4.1.4 Fiber Layer 3c-4 demonstrates statistically significant correlations with neuron population changes

Table 4-1 compares neuron population and myelinated fiber k-cluster counts, sorted by sub-, supra- and granular address.

Table 4-1. Neuron population and myelinated fiber k-cluster membership by layer and change interval.

	NCsup	NCIV	NCsub	FCsup	FC3c-4	FCSub	Nsup	NIV	Nsub	Fsup	F3c-4	Fsub
1 Mo	5	0	6	0	0	20	11	5	8	0	0	5
3 Mo	7	2	3	4	7	32	12	5	8	0	0	7
6 Mo	5	1	2	20	7	42	6	1	4	2	1	12
15 Mo	8	7	11	18	7	11	9	5	3	1	5	16
24 Mo	8	1	7	32	8	27	5	1	5	3	1	7
48 Mo	3	3	3	69	1	22	9	3	7	2	1	0
72 Mo	5	1	3	8	1	4	18	2	4	4	3	5

Core and non-core membership is defined in Chapter 2. Abbreviations: N = neuron, F = myelinated fiber, C = core, sub = subgranular, sup = supragranular, IV = neuron Layer IV, 3c-4 = fiber Layer 3c-4. To avoid ambiguity, neuron and myelinated fiber layers are distinguished by Roman and Arabic numerals, respectively.

Table 4-2 shows the sample correlation r and p-values for the 66 pairs contained in Table 4-1. There are only four significant correlations based on a two-tailed t-test: neuron core supragranular members with fiber core Layer 3c-4 ($r = 0.7657$, $p = 0.0448$), neuron core Layer IV with myelinated fiber non-core Layer 3c-4 ($r = 0.7634$, $p = 0.0458$), neuron non-core Layer IV with myelinated fiber non-core supragranular (r = -

0.8195, p = 0.0241), and neuron subgranular with myelinated fiber non-core Layer 3c-4 (*r = -0.8121, p = 0.0265*).

Table 4-2. The sample correlation r and p-values for the 66 pairs contained in Table 4-1.

r =	NCIV	NCsub	FCsup	FC3c-4	FCSub	Nsup	NIV	Nsub	Fsup	F3c-4	Fsub
NCsup	0.3493	0.6674	-0.4151	0.7657	-0.01	-0.2483	0.1507	-0.3208	-0.1366	0.3241	0.6617
NCIV		0.6425	0.234	0.3122	-0.3278	-0.1154	0.4147	-0.4325	-0.1768	0.7634	0.5291
NCsub			-0.11	0.2773	-0.4156	-0.252	0.3893	-0.3506	-0.2079	0.6007	0.5493
FCsup				-0.0831	0.0954	-0.4174	-0.3398	-0.001	0.308	-0.0089	-0.411
FC3c-4					0.5061	-0.5864	-0.1869	-0.3788	-0.0672	0.1368	0.6842
FCSub						-0.6669	-0.2709	0.2639	-0.3051	-0.6523	0.0991
Nsup							0.2897	0.0932	0.1547	0.1915	-0.2972
NIV								0.4503	-0.8195	0.0705	0.0959
Nsub									-0.5843	-0.8121	-0.6624
Fsup										0.3161	-0.1747
F3c-4											0.5718

p =	NCIV	NCsub	FCsup	FC3c-4	FCSub	Nsup	NIV	Nsub	Fsup	F3c-4	Fsub
NCsup	0.4426	0.1014	0.3544	0.0448	0.983	0.5914	0.7471	0.483	0.7703	0.4783	0.1055
NCIV		0.1196	0.6135	0.4954	0.4729	0.8054	0.3549	0.3325	0.7045	0.0458	0.222
NCsub			0.8144	0.5471	0.3538	0.5856	0.388	0.4407	0.6546	0.1538	0.2016
FCsup				0.8595	0.8387	0.3514	0.4559	0.9983	0.5016	0.9849	0.3596
FC3c-4					0.2464	0.1664	0.6882	0.4021	0.8862	0.7699	0.09
FCSub						0.1018	0.5567	0.5674	0.5057	0.1123	0.8326
Nsup							0.5286	0.8425	0.7405	0.6808	0.5174
NIV								0.3106	0.0241	0.8807	0.8379
Nsub									0.1684	0.0265	0.105
Fsup										0.4898	0.7079
F3c-4											0.1799

Figures 4-4 and 4-5 show the positively correlated neuron and myelinated fiber counts from Table 4-1. Figure 4-4 depicts the supragranular addresses in the core neuron population change k-clusters compared to the core myelinated fiber address counts for Layer 3c-4.

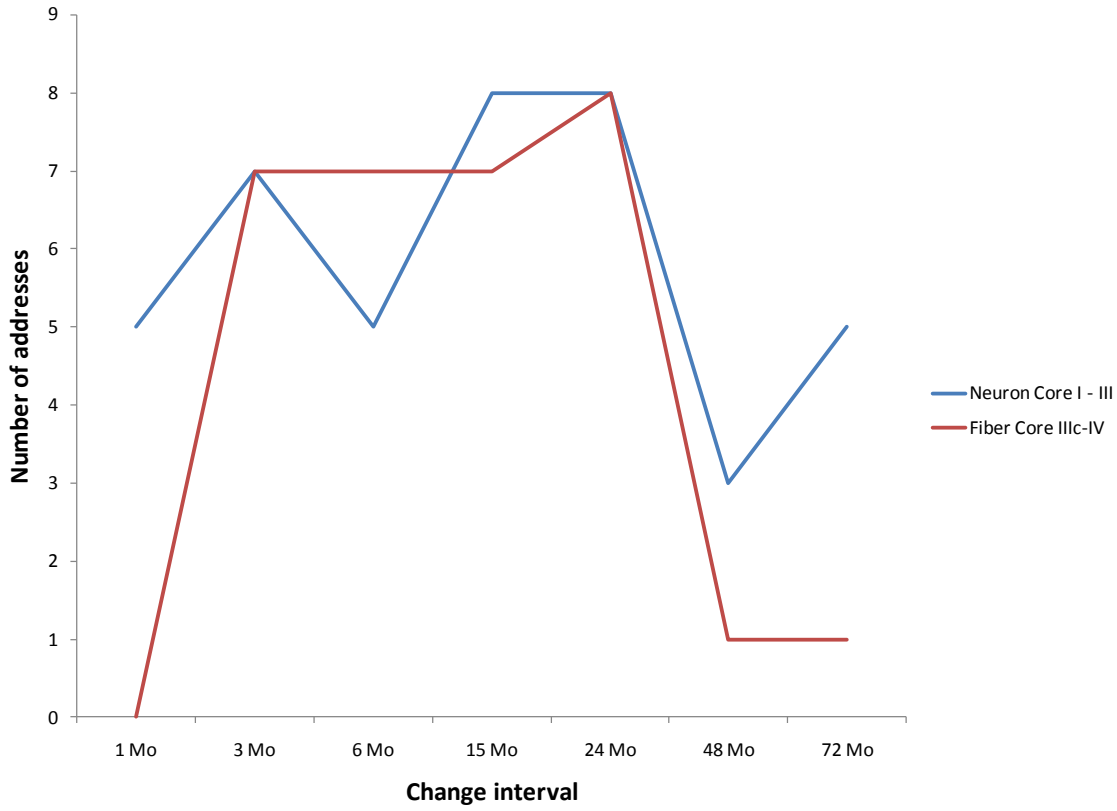


Figure 4-4. Number of addresses in core neuron supragranular populations compared to core myelinated fiber members in Layer 3c-4. Because neuron and myelinated fiber data are not directly comparable, counts are at the level of Brodmann Areas, without sub-areas. That is, core membership is based on whether an address is within a k-cluster core or whether an address in a given Brodmann Area occurs within a Brodmann Area that also has a k-cluster core address. The correlation of these two trajectories is statistically significant ($r = 0.7657$, $p = 0.0448$, two-tailed t-test). These relationships are detailed in Table 4-3.

Table 4-3 shows all addresses in Figure 4-4. None of the neuron addresses is the same as the myelinated fiber addresses at the corresponding change interval. Consequently, the right-hand column in Table 4-3 also shows the age the myelinated fiber core address demonstrated the greatest change to its myelinated fiber count.

Table 4-3. Neuron and myelinated fiber addresses included in Figure 4-4.

	Neuron Core contains LI - III	Neuron Core in other layers than LI to LIII	Fiber Core contains L3c-4	Fiber Core in another layer than L3c-4	Age for Corresponding Fiber
1 Mo	10_III				Post 72 months for 10 3a and 3b
	40_III				40 3b at 24; 40 3a at 48
	37_I				37_1 at 6
	42_III				42 3a and 3b at 24
	41_III				41 3a and 3b at 24
3 Mo	6_III	6_II	4Hn_3c-4	4L_3c-4	6 2 and 3a at 24; 6 3b (non-core) at 48
	44_II	44_III	4Hd_3c-4		44 3b at 24; 44 2 and 3a at 48
		39_II	3Hn_3c-4	3T_3c-4	39 2 at 48
	38_I	38_III		1L_3c-4	38 1 at 6 Mo; 38 3a at 72, 3b (non-core) at 48
			17_3c-4		
6 Mo	7_II		4T_3c-4		7 2 at 48
	37_II			3L_3c-4	37 2 at 48
	17_I	17_II	40_3c-4		17 1 and 2 at 48
	14_III		39_3c-4		14 3a and 3b at 48
			22_3c-4		
			41_3c-4		
			23/31_3c-4		
15 Mo	4L_III	4Hn_II		9_3c-4	***4L 3a and 3b at 15***; 4 ***Hn 2*** and 4 Hd 2 at 24
		4Hd_II		46_3c-4	
	8_III		45r_3c-4		***8 3b at 15***; 8 3a at 48
		6_I	47_3c-4		***6 1 at 3***
		19_I	1T_3c-4		19 1 at 24
	18_II		1Hn_3c-4		18 2 at 48
	21_II		7_3c-4		21 2 at 48
24 Mo	4P_I	4T_I		44_3c-4	***4T 1 at 3***
		4T_III		45_3c-4	***4T 3b at 6; 3a at 15***
		4P_III	10_3c-4		
		45_I		1Hd_3c-4	***45 1 at 24***
	14_I		36_3c-4		***14 1 at 6***
	13_I	13_II		37_3c-4	***13 1 at 6;*** 13 2 at 48
				19_3c-4	
				13_3c-4	
48 Mo	7_III			8_3c-4	***7 3a at 48***; 3b at 72 (non-core)
		40_I			***40 1 at 6***
		40_II			***40 2 at 48***
72 Mo		45_III	3Hd_3c-4		***45 3a at 48***; ***3b at 72***
	22_I	22_III			***22 1 at 48; 3a and 3b at 24***
	24_III				***24 3b at 24; 3a at 48***

The column entitled “age for corresponding fiber” refers to the change interval in which Conel’s observations of myelinated fiber density has the greatest *C(l, b, t)* value corresponding to the neuron population core k-cluster address. Red values are those for which the myelinated fiber change interval precedes the respective neuron change interval. Blue values are those for which the change intervals are in the same change interval. These begin at the 15 month change interval and then predominate.

Figure 4-5 shows the relationship of neuron core Layer IV k-cluster members with myelinated fiber non-core Layer 3c-4 k-cluster members. Whereas the correlation of these two populations is positive (*r = 0.7634, p = 0.0458*), non-core fiber Layer 3c-4 members have a statistically significant negative correlation with non-core neuron Layer V and VI as well (*r = -0.8121, p = 0.0265*).

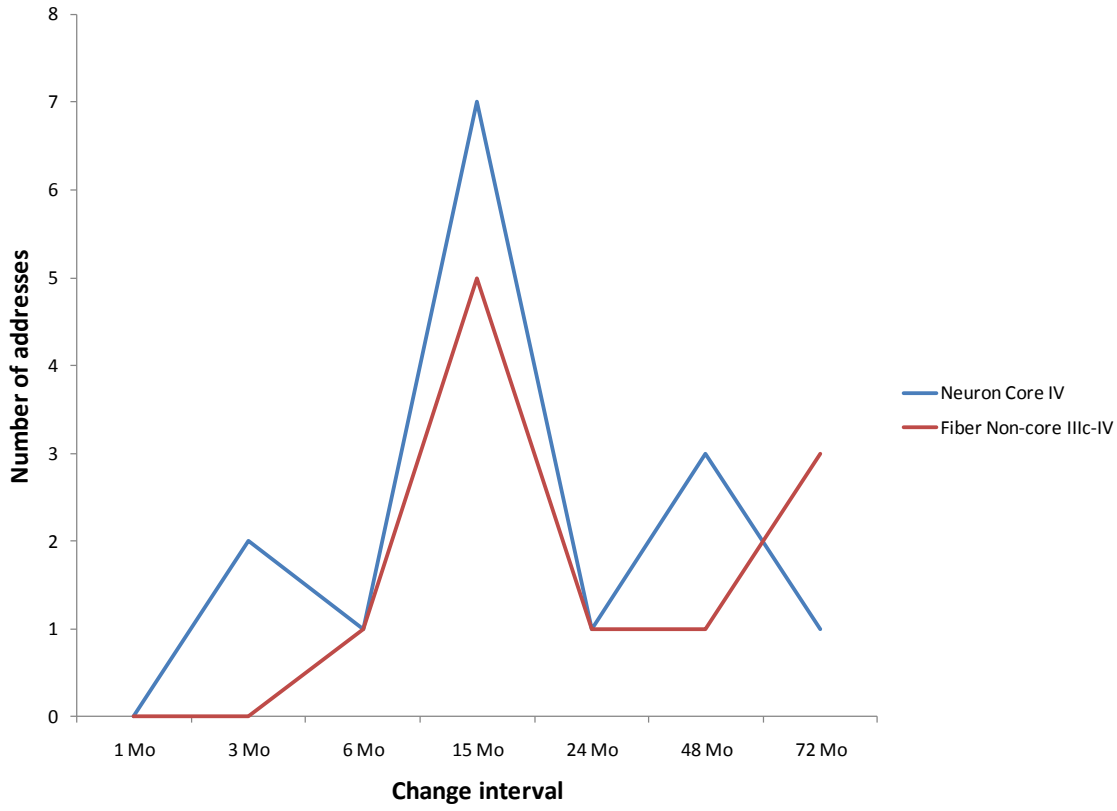


Figure 4-5. Neuron core Layer IV k-cluster members are significantly correlated with myelinated fiber non-core k-cluster members from Layer 3c-4. Of the neuron population k-clusters, membership included 15 Layer IV addresses that fell into the core as defined in Chapter 2. Similarly, 11 addresses in Layer 3c-4 fell into a fiber k-cluster core. The trajectories of the two sets distributed into their respective change interval are significantly correlated ($r = 0.7634$, $p = 0.0458$, two-tailed t-test). Details of the two sets appear at Table 4-4.

Table 4-4. Correlation of neuron core Layer IV with non-core myelinated fiber Layer 3c-4.

	Neuron Core contains LIV	LIV in Neuron Core	Fiber Non-core 3c-4
3 Mo		44_IV	
	39_IV		
6 Mo	17_IV		24_3c-4
15 Mo		4L_IV	6_3c-4
		4T_IV	6_8_9_3c-4
		4P_IV	18_3c-4
	6_IV		21_3c-4
		8_IV	14_3c-4
	19_IV		
		18_IV	
24 Mo		42_IV	38_3c-4
48 Mo	40_IV		42_3c-4
	22_IV		
		21_IV	
72 Mo	45_IV		8_9_3c-4
			30_3c-4
			29_3c-4

4.1.5 Neuron population and myelinated fiber k-cluster cores do not have the same distribution

Figure 4-6 shows that the core members of the k-clusters to not have the same relationship to the overall k-cluster membership when neuron population and myelinated fiber density data are compared.

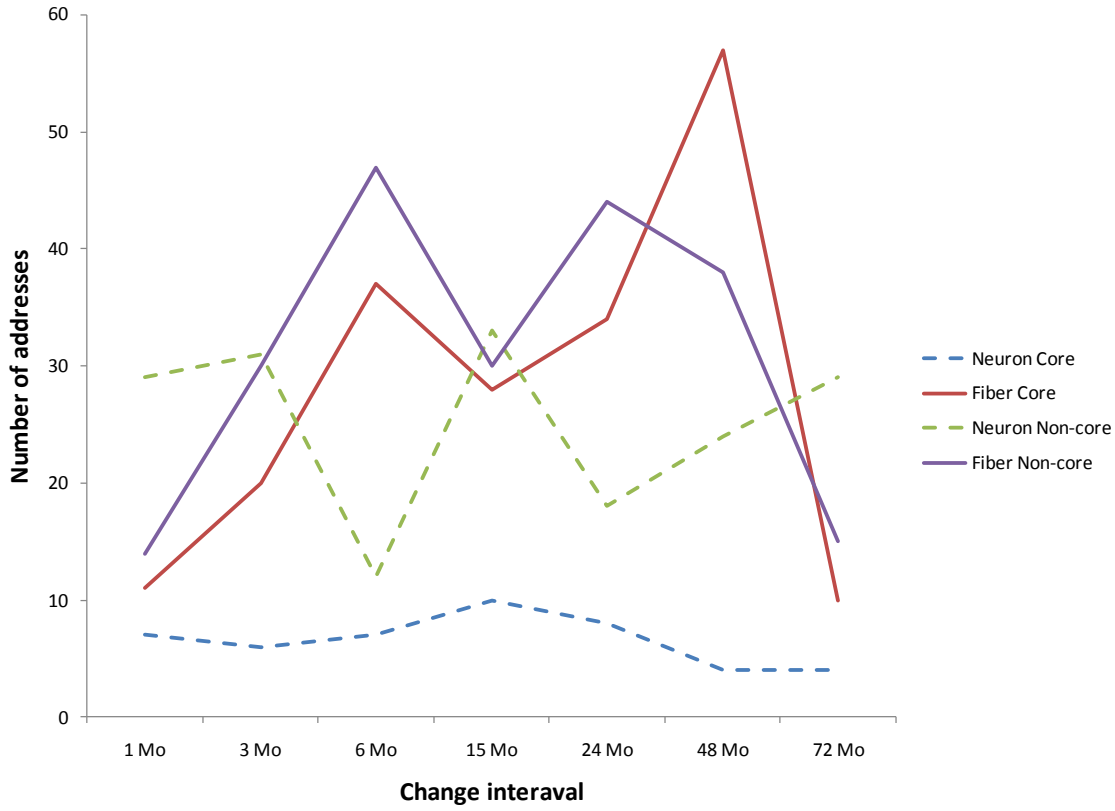


Figure 4-6. Comparison of neuron population and myelinated fiber core and non-core k-cluster membership. K-clusters are characterized by the respective change interval in which they have they maximum mean value for C(l, b, t). Neuron non-core membership in a given k-cluster is consistently greater than the core membership. This is generally the case for myelinated fibers, except during the 48 month change interval. The ratio of core/non-core membership is consistently greater for myelinated fiber density changes, exceeding 1 during the 48 month change interval.

Table 4-5 shows the correspondence of core neuron and core myelinated fiber k-cluster membership.

Table 4-5. Correspondences of core neuron and core myelinated fiber k-cluster members hip at successive change intervals.

Max C(l, b, t)	BA	Neuron			Myelinated Fiber			Comment
		Supra-granular	Granular (IV)	Sub-granular	Supra-granular	Granular (3c-4)	Sub-granular	
1 Mo	45r	0	0	1	0	0	1	N V* to F Vert*
3 Mo	6	2	0	1	1	0	2	N II/III*/V* to ***F 1****/Va/Vib
	44	2	1	0	0	0	2	N II*/III/IV F Vib*/Vert
6 Mo	45r	0	0	1	0	0	1	N ***VI**** to F ***5b-6a****
	7	1	0	0	1	0	1	N II* to ***F 1****/6b
	37	1	0	0	1	0	3	N II* to ***F 1****/5a*/5b-6a/6b
	17	2	1	0	0	0	2	N I*/II/IV to F 5b-6a*/6b*
	14	1	0	0	1	0	2	N III* to ***F 1****/5b-6a*/6b
15 Mo	4	3	2	5	3	0	0	4L N ***III**** to F ***3a*/3b**** ; 4Hd N VI*, 4P N VI*
	8	1	1	1	1	0	0	N ***III****/IV/V to F ***3b****
	29	0	0	1	0	0	1	N ***V**** to F ***5a****
24 Mo	4	4	0	1	5	0	1	4Hn N VI to F 2*/3a*/3b; 4P N I is core
	44	0	0	1	1	1	0	N V* to F 3b*/3c-4
	45	1	0	1	1	1	0	N I/VI* to F1*/3c-4/***5a****
	42	0	1	1	3	0	0	N IV/VI* to F 2*/3a*/3b*
	41	0	0	2	3	0	1	N V*/VI to F 2*/3a*/3b/Sub
	13	2	0	0	2	0	1	N I*/II to F 3b/3c-4/***Va****
48 Mo	7	1	0	1	2	0	1	N ***III****/VI to F 2*/***3a****/Vert
	40	2	1	1	2	0	0	N I/II/IV*/V to F 2*/3a*
	22	0	1	0	1	0	2	N IV* to F 1/***Sub*/Vert****
	21	1	0	1	3	0	1	N IV/VI* to F 2*/3a*/3b*/5b-6a/***Vert****

Asterisks indicate the respective neuron or fiber address is core, as defined in Chapter 2. Addresses for sub-areas in BA 4 are compared in a single row. In Chapter 3, 47 core areas appeared in the k-cluster cores, distributed across 41 different Brodmann Areas, ignoring sub-areas. Of these 41 Brodmann Areas, 21 (51.2%) appear on this table. In eight of these areas (38.1%) the myelinated fiber cores are contrary to the subgranular to supragranular pattern in Figure 4-3 (in red). In five of these eight cases (62.5%) the respective neuron and myelinated fiber layers are adjacent. In five additional cases (23.8%), the neuron and myelinated fiber cores occur in the same layer (in blue). Abbreviations: N = neuron population; F = myelinated fiber density.

4.1.6 Within k-clusters, mean C(l, b, t) for myelinated fibers is significant only for BA 1, 3 and 4

Figure 4-7 shows the mean value of *C(l, b, t)* for addresses in respective Brodmann Areas when those areas are members of a k-cluster. This incorporates *C(l, b, t)* values for all 415 addresses where those values exist, but excludes the *C(l, b, t)* values for those addresses during change intervals when they are not part of the k-cluster. There are only six cortical areas where this mean $C(l, b, t) > 3.9$, which is the threshold for $p < 0.05$ as shown in Chapter 2. Four of these are subareas within BA 4, the remaining two fall into BA 3 and BA1.

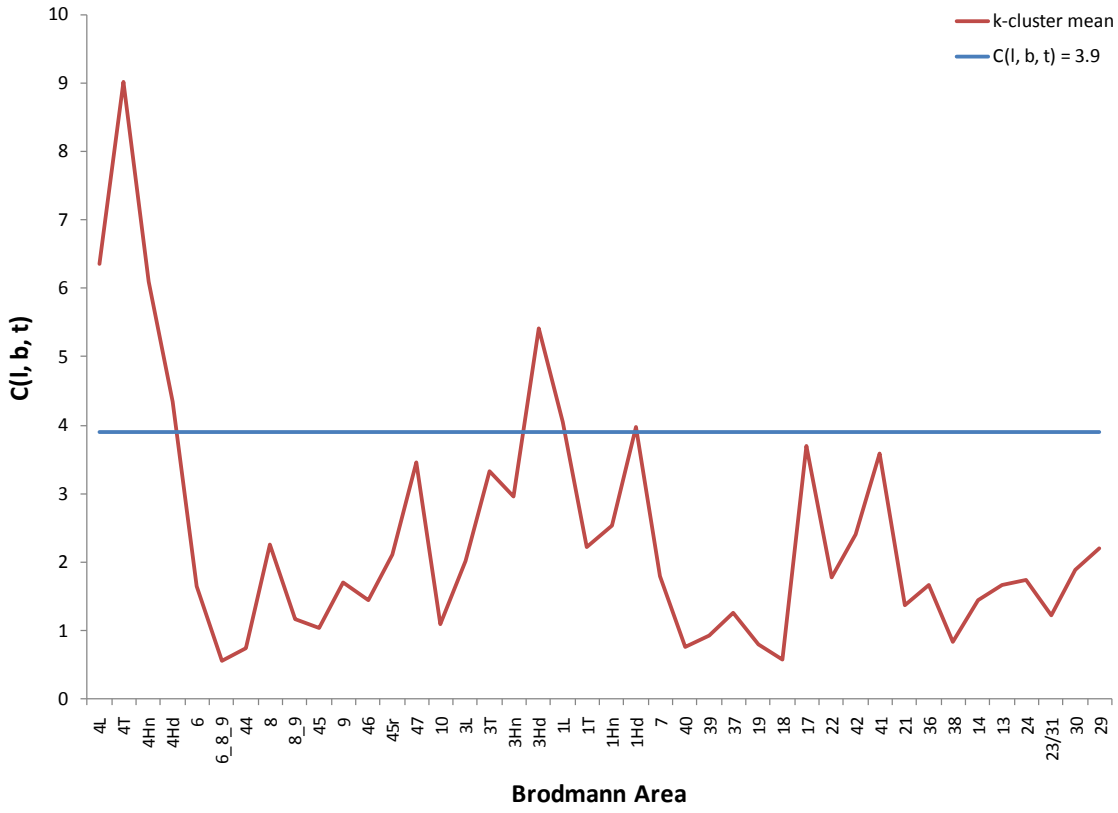


Figure 4-7. Mean values of *C(l, b, t)* within k-clusters by Brodmann Area. Mean *C(l, b, t)* for each Brodmann Area/Sub-area is calculated on the ten *C(l, b, t)* occurrences for each area when that address falls within a k-cluster. As noted earlier, there are only eight occurrences for BA 45r, and seven for BA 10, as Layer 2 and 3a densities are still zero through age six in Conel's myelinated fiber data. As noted in Chapter 2, when $C(l, b, t) > 3.9$, $p < 0.05$ based on the chi-square distribution for one degree of freedom. $C(l, b, t) = 3.9$ is shown as the blue line. BA 4 values for the leg, trunk, hand and head exceed this threshold, as do BA 3 for the head and BA 1 for the head.

4.2 Discussion

Myelinated fiber neocortical change patterns during postnatal human development are dependent on age and layer. Onset age for myelinated fiber counts and the monotone increasing pattern in Figure 4-1B recapitulate almost exactly the "inside-out" pattern for pre-natal laminar development in the cortex (Bar and Goffinet, 2000) with the exception of Layer 5a. This is also evident in the change from sub-granular to supra-granular membership of myelinated fiber k-clusters in Figure 4-3. The transition from sub- to supragranular majorities occurs during the 15 month change interval, which is precisely the time that changes in Layer 3c-4 non-core myelinated fibers are significantly correlated with changes to core neuron population changes in Layer IV, as shown in Figure 4-5. Myelinated fiber Layer 3c-4 non-core changes are also significantly correlated with core neuron supragranular layers as shown in Figure 4-4 ($r = 0.7634$, $p = 0.0458$, *two-tailed t-test*), and significantly anti-correlated ($r = -0.8121$, $p = 0.0265$, *two-tailed t-test*) with non-core neuron subgranular change patterns. The age-dependence of myelinated fiber $C(l, b, t)$ magnitudes is evident in Figure 4-2 as well, where the mean magnitude of changes to fiber counts declines exponentially with increasing age in months.

As Figure 4-2 shows, these myelinated fiber results are directly contrasted with the change patterns for neuron populations examined in Chapter 3. For neuron populations, the mean magnitude of $C(l, b, t)$ values in the k-clusters is approximately constant across all seven observed change intervals, and thus dependent on the k-cluster

membership and not on the observation ages themselves. Moreover, the obvious laminar pattern in the myelinated fiber data is not evident when k-cluster membership is compared.

Within the basic myelinated fiber age and layer-dependent pattern, statistically significant magnitudes for *C(l, b, t)* correspond to sensorimotor functions, as revealed by Figure 4-7. Addresses in all four sub-areas of BA 4 (primary motor cortex, Gazzaniga, et al., 2002) recorded in Conel's data have a mean magnitude in their respective k-cluster above the 3.9 threshold for significance. Likewise, the addresses for BA 3 for the head and BA 1 for the head (primary sensory cortex, Gazzaniga, et al., 2002), have mean *C(l, b, t)* magnitude greater than 3.9.

Table 4-5 lists all Brodmann Areas where there is a core neuron address and a core myelinated fiber address. In Chapter 3, core neuron addresses were shown to be consistent with human sensorimotor development. As noted in section 4.1.5, the majority of the core neuron addresses also have corresponding myelinated fiber core addresses. Of those corresponding addresses, 61.9% co-occur in the same layer, or the fiber layer occurs contrary to the subgranular to supragranular pattern in Figure 4-3. Since the remaining myelinated fiber addresses could be predicted on the basis age dependence alone, it is not possible to infer dependence on neuron function or age in the remaining cases.

Table 4-3 lists 22 Brodmann Areas with a core neuron address in Figure 4-4, as well as seven additional Brodmann Areas where a core neuron address occurs somewhere else in the respective column (ten addresses including three in BA 4). These are the core

supragranular neuron addresses in Figure 4-4 that are significantly correlated with core Layer 3c-4 myelinated fiber densities.

Table 4-3 shows first that the correlations between neuron-related cores and fiber-related cores do not occur in the same Brodmann Area at each change interval. Thus, if any particular function is the reason for this correlation, significant neuron population changes in supragranular layers during any particular change interval would be correlated with significant changes to myelinated fibers in core-related Layer 3c-4 in different cortical areas in that interval. For example, the first change interval in which this situation arises is for 3 months, where Table 4-3 associates neuron population changes in Layer I to III in BA 6, 38, 39 and 44 (supplemental motor cortex, the temporal pole, the angular gyrus and Broca's Area, respectively, Gazzaniga, et al., 2002) with myelinated fiber density changes in BA 1, 3, 4, and 17 (primary and secondary sensory cortex for the leg, trunk and hand; primary motor cortex for the leg, hand, and head; and primary visual cortex, respectively; Gazzaniga, et al., 2002). In Chapter 3, the neuron core-related areas were associated with verb generation in adolescent and adult subjects (Le Bihan, et al., 1993, Cao, et al., 1999), semantic processing in adults (Binder, et al., 1996), word-picture matching among children from 5 to 18 (Schmithorst, et al., 2007), as well as perception of the "end" goal of an action (Hesse, et al., 2009) and at least one component of the primate "mirror" system (Cadoret, et al., 2005). The corresponding myelinated fiber core areas in Table 4-3 provide primary vision inputs at 1 month, when children begin to focus on external objects (Piaget, 1952), as well as somatosensory inputs for the respective motor areas in the neuron core. Thus, the functional associations

continue to be consistent with the overall sensorimotor themes in neuron population changes noted in Chapter 3, including the "motor schema" construct advocated by Mandler (1999). Similarly, the association of core neuron Layer IV addresses with non-core myelinated fiber core Layer 3c-4 in Table 4-4 appears to show a similar correlation of different cortical areas with the neuron-related functions in those change intervals. Chapter 3 also noted that neuron population changes occur in these areas well before the corresponding cognitive functions manifest themselves during later childhood and adulthood. The "corresponding fiber" column in Table 4-3 shows one possible basis for this delay. Relatively large changes to myelinated fiber densities for each of the supragranular neuron layers for the 3 month change interval did not appear until 24 and 48 months for most cases, and not until 72 months for BA 38 3a. Since the completion of myelination is a signal of the onset of full functional capacity (Flechsig, 1920; Yakovlev and Lecours, 1967), the occurrence of relatively large changes at these addresses at later change intervals is a sign that these circuits were not complete until much later than the neuron population changes. This is an example of the second major pattern that appears in Table 4-3. Prior to 15 months, the corresponding age for changes to the respectively supragranular address is invariably from 24 to 72 months. Beginning at 15 months, the fiber changes are simultaneous or occur earlier than the neuron population changes. This reflects the general subgranular to supragranular shift in Figure 4-3.

This age dependence raises an important factor in interpreting core k-cluster membership for myelinated fibers. As demonstrated in Chapter 2, changes that depend

on exponential decay such as the curve in Figure 4-2 for myelinated fiber will have *C(l, b, t)* approximately zero, except for noise. Such a pattern would result in large numbers of highly correlated changes that all have small *C(l, b, t)* magnitudes. Once divided into k-clusters, this would result in maximal cliques that approach the size of the entire k-cluster. This is very probably the reason for the difference in relative size of k-cluster core membership evident in Figure 4-6. Core membership for neurons is rare and relatively interesting as a result. As noted in results, mean neuron *C(l, b, t)* in Figure 4-2 is not dependent on age. Conversely, if core membership is explained by dependence on age in myelinated fibers, the addresses that are ***not*** dependent become more interesting, as are the *C(l, b, t)* magnitudes that are statistically significant.

5. Evolution of preterit forms in English is heterogeneous and discontinuous

Chapter 1 introduced a clock hypothesis for neocortical development given the following elements: 1) a self-organizing and potentially hierarchical developmental "clock" that paces the organization of neocortical regions (Brodmann Areas) needed for specific cognitive processes; 2) differential heterogeneous response to the "clock" at the level of area/layer combinations that is subject to evolutionary pressure; and 3) the conservation of proven hierarchical developmental sequences in the neocortex. We assumed genetic, epigenetic and environmental factors acting as inputs that effect cortical addresses differentially, and predicted that correlated and statistically significant simultaneous change would have a functional basis.

Chapter 3 showed that neurotrophin- and activity-based neuron population changes induce unique k-clusters at each change interval that are consistent with sensorimotor development in young children (Piaget, 1952). K-clusters emerged at the area/layer address level, indicating the conservation of layer-based control in addition to activity-based inputs at the level of Brodmann Areas. Each k-cluster had a small core, consisting of the members of one or more maximal cliques of highly correlated addresses at each successive change interval. Core membership and statistically significant

magnitudes of *C(l, b, t)* were significantly correlated. These cores were consistent with sensorimotor development, a hierarchical process of continued adaptation and assimilation (Piaget, 1952), and coincided with subsequent linguistic function in older subjects in the case of Broca's Area, but the age at which sophisticated cognitive functions such as language or participation in the primate mirror system were well in advance of when that behavior normally emerges. The mean magnitude of *C(l, b, t)* in each k-cluster was relatively constant across change intervals, or independent of increasing age.

In Chapter 4, change to myelinated fiber density was shown to be dependent on age and layer. While time dependence is likely the result of temporal control, and layer dependence is symptomatic of the conservation of the prenatal "inside-out" development of cortical layers, the pattern of exponentially declining *C(l, b, t)* magnitudes and proportionally large k-cluster core membership contrasts with neuron population changes. This may reflect a situation where control of oligodendrocytes is correlated most closely with neuron secretions (Butt and Berry, 2000), then by expression of cell adhesion molecules (Coman, et al., 2005), and then finally by electrical activity in the neuron (Sanchez, et al., 1998; Sirevaag and Greenough, 1987; Szeligo and Leblond, 1977). As noted in Chapter 4, temporal dependence is likely to result in very small *C(l, b, t)* magnitudes and larger k-cluster cores. Nevertheless, significant correlations between neuron population and myelinated fiber density change continued to be consistent with sensorimotor development. Moreover, the layer-dependent pattern in myelinated fiber

change helps to explain completion of cognitive functional development months or years after the initial neuron population change.

For both neuron populations and myelinated fiber densities, the three elements of genetic, epigenetic, and environmental factors were present for the operation of a self-organizing clock. What happens when there is no obvious means by which genetic and epigenetic factors can operate? To answer that question, we can turn to an example from linguistic evolution, reflecting a complex cognitive behavior where the changes occur solely in the outputs of the humans that comprise the speech community.

5.1 A complex cognitive behavior: evolution of irregular verb forms in English

5.1.1 Background from linguistics

Since the days of the Neogrammarians and their critics, theories of linguistic change have alternated between two formulae: "exceptionless" laws on the one hand, and the dictum that "every word has its own history" on the other. Dialect evidence from both historic sound changes and from sound changes in progress supports both (Labov, 1994). Nevertheless, Lieberman, et al. (2007) claim to have discovered an exceptionless law, hereinafter called the L-M Hypothesis (for Lieberman-Michael): in terms of Modern English, non-'ed' preterits have been "decaying" since the Old English period, with a half-life inversely proportional to the square-root of the usage frequency. Thus, 'be', which has the highest usage frequency in the CELEX corpus (van der Wouden,

1988) used in support of the L-M Hypothesis, has a tremendously long half-life, at least 38,800 years.

The L-M Hypothesis is morphological and thus focused at a word level. It implies that a common default attractor has operated since pre-1066 days until the present. In Old English, the putative default attractor was a transform that operated on the so-called 'weak' verbs, using dental suffixes (e.g., '*-de*' or '*-te*' depending on the voicing environment) to form the preterit. For example, the transform took 'I judge' from '*iċ dēme*' to '*iċ dēmde*'. Weak verbs like '*iċ dēmde*' constituted approximately three quarters of the verbs in Old English (Quirk and Wrenn, 1994). In Modern English the default attractor, which is more or less the direct descendent of the Old English version, consists of a dental suffix ('-ed') that likewise transforms 'I deem' to 'I deemed' now. As in Old English, the suffix is pronounced as /t/ in voiceless environments, such as 'laugh/laughed'. Under the L-M Hypothesis, all other preterit transforms are decaying to this default at a regular rate.

In Old English, this transform competed with 'strong' verbs that used vowel alternations (e.g., '*beran/bær*' 'bear/bore'), descended from alternated thematic stem vowel *–e–* with *–o–* found in earlier Germanic and Indo-European forms. Other verbs, such as '*eom/wæs*' were simply irregular, then as now ('am/was'). Old English also had a few, but very frequently used pretero-presents, where preterit forms appeared in present contexts, such as 'might' does now. Old English weak verbs themselves fell into different classes. The example above, '*iċ dēme*/*dēmde*', was Type I, but others fell into Type II (with suffix '*-ode'*), such as '*iċ lufiġe/lufode'* for 'I love/loved', which we can

safely group as ‘regular’ in a loose sense. However, other weak verbs demonstrated alternations in the root consonant, such as ‘*iċ hæbbe/hæfde*’ for ‘I have/had’, which will be counted in this paper as non-‘ed’ forms (Campbell, 1959). One should remember that multiple preterit attractors existed and continue to exist, and that a contrast between a single “regular” form and all the others is really an oversimplification. Old English also distinguished between preterits and past participles, but for simplicity we will focus exclusively on the preterits.

Here we examine whether this putative default transform, embodied in the various cognates of the (unchanged) *stem + ed* form, represents the sole active attractor for preterit formation in English since Old English. If it is, under the L-M Hypothesis we would expect a monotone decreasing share of English verbs to follow alternative patterns, with the most frequently used verbs most resistant to this change. We will show that relative usage frequency bears a different relationship to accessions and decrements from the population of non-‘ed’ verbs over time. In fact, change patterns for the population of high relative frequency and low relative frequency non-‘ed’ verbs are approximate mirror images. Finally, the change patterns for these two groups of verbs appear to reflect distinct Middle English and Modern English phases, as well as highly significant cluster changes related to the completion of the Great Vowel Shift. These reveal changes tied to individual word histories, as well as the systemic operation of changes to the overall population of English verbs, but not a single monotone decreasing population of non-default verbs.

5.1.2 Modifications to data and methods in Chapter 2

The Lieberman, et al. (2007) quantification of a regularized transition process relies on elements that are subject to some criticism. First, they restrict their analysis of change steps to just the two steps from Old to Middle and from Middle to Modern English based on compilations from handbooks. This is less than precise in assigning the dates of verb transitions. Second, while their assignment of current usage frequency from the CELEX corpus was a pragmatic decision made necessary by the lack of similar corpora for earlier stages of English, their representation of verbs from the CELEX corpus had errors and omissions. For example, 'become' is not listed in their supplemental data until Modern English, but actually occurred in Old English. Other Old English strong verbs are not listed, including all the modal auxiliaries ('shall', 'will', 'can', 'may'), which belong to the exceptional pretero-present class. An additional group appears to have been weak (equivalent to the '-ed' ending), or foreign verbs borrowed with the '-ed' ending applied all along. Altogether, these amount to more than 10% of their 249 item list. Moreover, many forms ('sneaked/snuck') are now or were co-extensive. Finally, their decision to focus exclusively on the set of Old English non-'ed' verbs eliminated most of the data that would have negated their hypothesis, whether or not the list was accurate.

To correct these inconsistencies, we deleted verbs without a non-'ed' phase, and included the pretero-present class for a revised total of 228 verbs that includes additions to the non-'ed' set between the 10th and 20th centuries. To refine the transition data in the

verb lists beyond two change steps, we used the Oxford English Dictionary (OED, 1971) to create a table of transitions across the 11 centuries ending between 1000 and 2000 AD. For each verb, the OED cites forms and dates. We define a set S of verbs that do not use the *stem + ed* transform to form a preterit. For each verb *n* in S in century *c*, we assign the usage frequency *v* from the CELEX corpus, in common with Lieberman, et al. This presumes that usage is constant over all 11 centuries, which we provisionally accept as a working hypothesis. We also note that this hypothesis is technically negated by co-extensive forms and by wholesale substitution of roots (e.g., 'take' rather than *niman*). We return to these points in the discussion later.

The L-M Hypothesis is based on usage frequency. It is difficult to envision a mechanism in which a language learner (especially an infant or toddler) examines a body of data for frequency of occurrence, so we substitute the notion that the child responds to the relative frequency of forms in given contexts: greater frequency implies greater probability that a rule or construction will be adopted in each learning context (see Tomasello, 2003).

As there are no cortical layer or Brodmann Area distinctions in the data, to calculate the relationships in Chapter 2 we will return to the table of observations *T(r, c)* for r rows and c columns, redesignated in this chapter as *T(n, c)* for n verbs numbered from 1 to 228 and 11 centuries c corresponding to the 10th to the 20th centuries for which the English data are available. Equation 1 then produces *R(n, c)* for the n verbs and c centuries and Equation 3 produces *C(n, c-1)* for the *n* verbs and *c-1* change intervals

between the 11 centuries. Similarly, $c - 1 = 10$ for the indices in Equation 8 for the correlation distance used to identify the k-clusters in the verb data.

As noted in Chapter 2, we tested for the optimal value of k up to 100,000 replications, which proved to be subject to diminishing returns as shown in Figure 2-5. As noted there, optimum $k = 13$ for this set of data.

5.2 Results for the set S of non-'ed' forms

5.2.1 Log frequency results

For a given verb, if the usage frequency is the average for the set S in that century, $R(n, c) = 0$. If the usage frequency is greater than average for S in that century, $R(n, c) > 0$. A histogram of $R(n, c)$ for all verbs that belonged to the set S of non-'ed' verbs is shown in Figure 5-1. Between the 10th and 20th centuries S was not fixed in size. In the 10th century, it included 188 of the 228 corpus verbs, and increased to 214 in the 13th century, and thus fails to support the L-M half-life model since the set was not monotone decreasing. It has undergone a net loss in membership thereafter, decreasing to 157 verbs in the 20th century. Between the 10th and 20th century, S included a 'stable' membership of 125 verbs. A total of 40 'accessions' occurred of which 32 remain in the set. A total of 71 verbs 'decayed' to the equivalent of a *stem + ed* form, including 8 of the 40 accessions. We call these 'deletions' from the set S. The set of 8 verbs that were both accessions and then later deletions are called 'transitions' in Figure 5-1. Most

(64.8%) of the stable set have $R(n,c) > 0$, while nearly all (84.9%) of the deletions have $R(n,c) < 0$. Nearly half (43.8%) of the accessions have $R(n,c) > 0$. Since we do not have usage frequencies for all verbs in each century, we cannot assert that $R(n,c) = 0$ implies the average for all verbs. $R(n, c) = 0$ is simply the average usage frequency in a given century for the verbs that happened to be in set S in that century. That list was different from one century to the next.

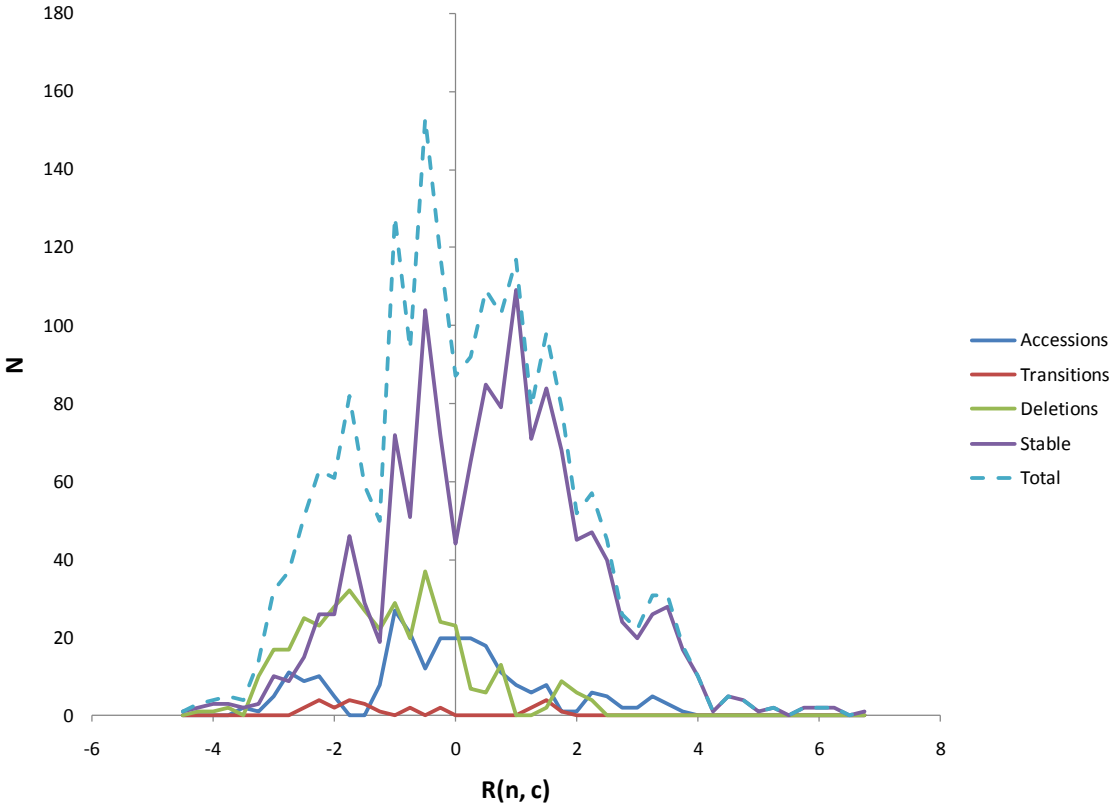


Figure 5-1. Histogram of all values of R(n, c) for the set S of non-'ed' verbs from the 10th to the 20th century. The figure is a histogram of relative usage frequencies for the 228 verbs in the corpus for S, normalized by Equation 1. At $R(n_1, c_2) = 0$, a verb n_1 in century c_2 has the average usage frequency for a member of set S in century c_2. While

the set S shows an approximately normal distribution, the constituents of S representing the stable non-'ed' verb population, accessions, deletions and transitions do not. *N* for the histogram represents numbers of occurrences of *R* values in bins of 0.25 nats without regard to the century. Thus, for 'take/took,' the highest frequency verb in the accession set with the first attestation in the 11^{th} century, the histogram contains the 9 occurrences of *R(n, c)* between 3.50 and 3.75, and another less than 3.50 since *R(n, c)* is not generally constant from century to century, even for the stable verbs. S is nearly symmetric, with 48.6% of all non-'ed' instances with R(n, c) > 0 (sample skewness = 0.236). R(n, c) > 0 for 60.9% of the stable verbs (sample skewness = 0.105), for 39.1% of the accessions (sample skewness = 0.234), for 25.9% of the transitional verbs (sample skewness = 0.713), and for 12.2% of the deletions (sample skewness = 0.477). Transitional verbs were not double counted with the accessions and deletions.

To examine for discontinuities between the accession and deletion subsets of S, Figure 5-2 shows the mean *R(n, c)* score for both categories by century. The mean accession R-values in the 11^{th}, 12^{th} and 19^{th} centuries are significant ($p = 2.39 \times 10^{-5}$, $p = 3.80 \times 10^{-4}$ and $p = 3.20 \times 10^{-7}$, respectively, two-tailed t-test with a Bonferroni correction for n = 16).

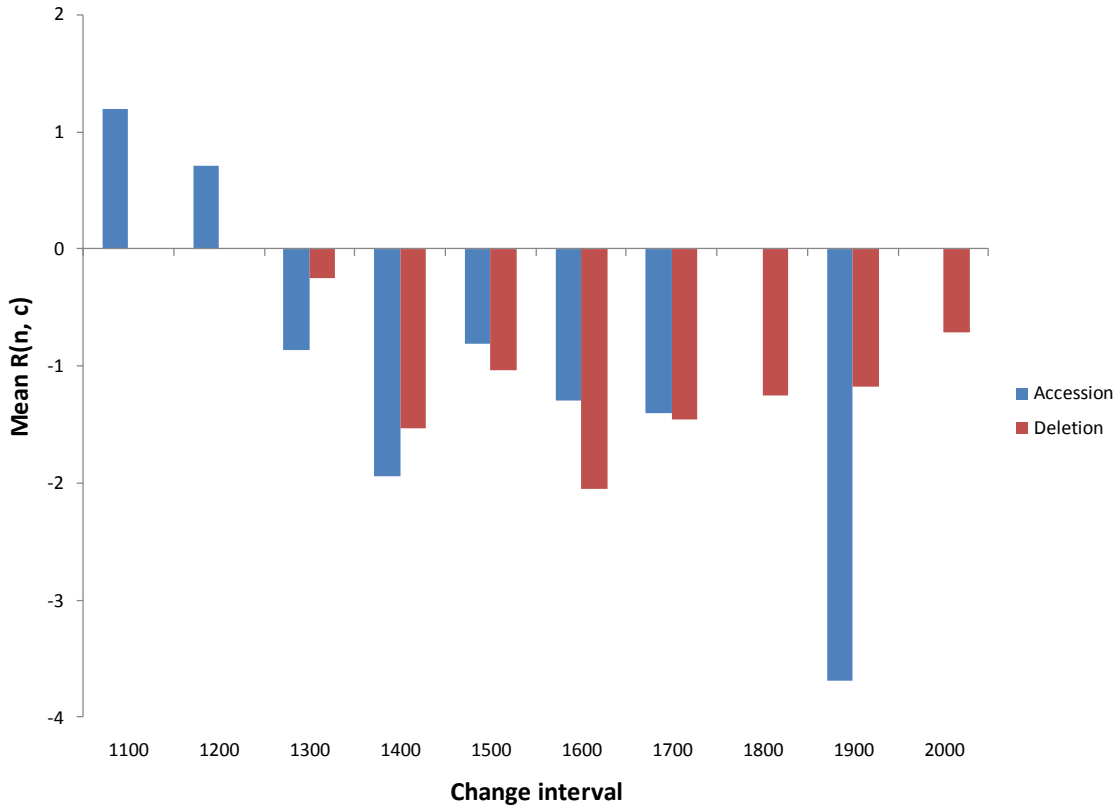


Figure 5-2. Mean values of *R(n, c)* for verbs in the accession and deletion subset of S by century. The 10th century is not depicted as these subsets appear during the change intervals from the 10th century baseline. While coextensive statistically, the two subsets are not coextensive by century. The early high frequency accessions in the 11th and 12th century include 'take/took' (R = 3.714), 'put/put' (R = 2.496), and 'spend/spent' (R= 1.487) in the 11th; and 'get/got' (R = 3.635), 'build/built' (R = 1.279), 'strike/struck' (R = 0.557), and 'spread/spread' (R = 0.407) in the 12th. The only subsequent accession verbs with *R(n, c)* > 0.4 were 'cut/cut' (R = 1.092), 'catch/caught' (R = 0.984), and 'hit/hit' = (R = 0.526), all in the 13th century. The low frequency accession to S in the 19th century was 'wed/wed' (R = -5.288), which is the accession outlier in Figure 5-1. Altogether, 13

(40.6%) of non-transitory accessions are of the ‘wed/wed’ type: monosyllabic, ending in a dental, with identical present and preterit. Including three cases like ‘build/built’ where the preterit shifts from *–d* to *–t*, the percentage increases to 53.1%.

5.2.2 K-means cluster analysis results

5.3.2.1 Changes to S depend on the usage frequency; there is only one verb with a statistically significant change step

Figure 5-3 shows that maximum *C(n, c-1)* for each of the 228 verbs is dependent on *R(n, c)* rather than on time. Moreover, only one verb, ‘wed/wed’, has a statistically significant value for *C(n, c-1)*.

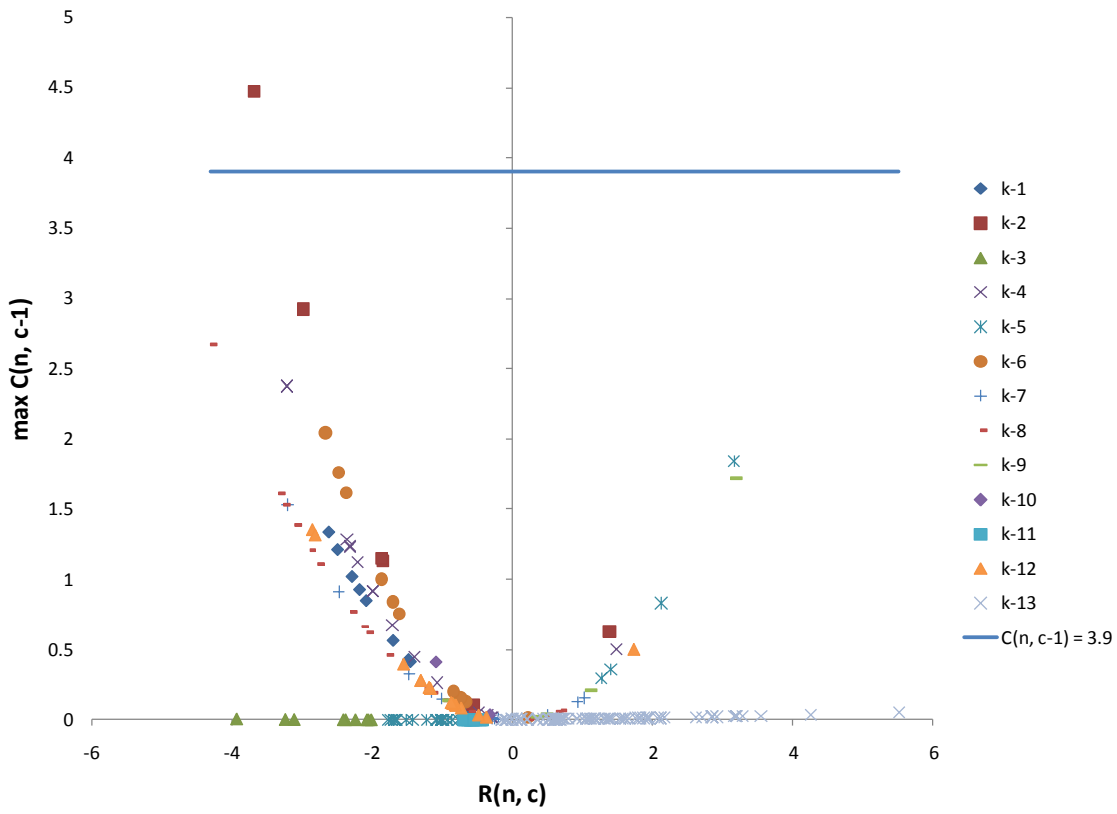


Figure 5-3. Max *C(n, c-1)* as a function of *R(n, c)*. The value of *R(n, c)* is logratio at the time the respective verb achieves its maximum value for *C(n, c-1)*. *C(n, c-1)* as defined in Chapter 2 follows a chi-square distribution with one degree of freedom, against the null hypothesis that R_0 at time zero is the same as R_1 at time one. This is approximately true for the stable verbs, which fall along the horizontal axis. By the working hypothesis in 5.2.1, usage frequency in any century is assumed constant. Since this frequency appears in *T(n, c)* in the centuries when a given verb is in set S, and the entry is otherwise zero, which applies to *R(n, c)* as well. *C(n, c-1)* will thus be approximately zero except in the century when a verb is accessed or deleted, and it will

be proportional to the square of *R(n, c)* at that time according to Equation 3. Thus, transitional verbs follow a parabola in the figure. The k-clusters follow slightly different parabolic paths. Only one verb, 'wed/wed' has a statistically significant value for *C(n, c-1)*. It is the only point above the threshold for significance, where *C(n, c-1) = 3.9* and *p = 0.5* (one-tailed chi-square with one degree of freedom). For 'wed/wed' *C(n, c-1) = 4.4745* and *p = 0.0344*. This verb was accessed into set S in the 19th century.

Figure 5-4 provides details of the 13 k-clusters. Mean *C(n, c-1)* values are similar to neuron and myelinated fiber k-clusters in showing a unique period in which the value is maximal; however, the values do not show a dependence on the century related to the k-cluster. Ten of the k-clusters contain all but four of the verbs accessed into or deleted from the set S during the preceding century, as well as 17 of the stable non-'ed' preterit verbs that were always members of S. The remaining three k-clusters—k-3, k-11, and k-13 in the random assignment of numbers to k-cluster—contain the remaining 112 of the 125 stable verbs. K-13 contains 81 of the stable verbs as well as four accessions into the set S: 'bend/bent' (11th century), 'cost/cost' (14th century), 'hurt/hurt' (12th century), and 'withdraw/withdrew' (13th century). All three of these stable k-clusters have values of *C(n, c-1)* that are not visible at the scale depicted in Figure 5-3. We will take these stable k-clusters up shortly.

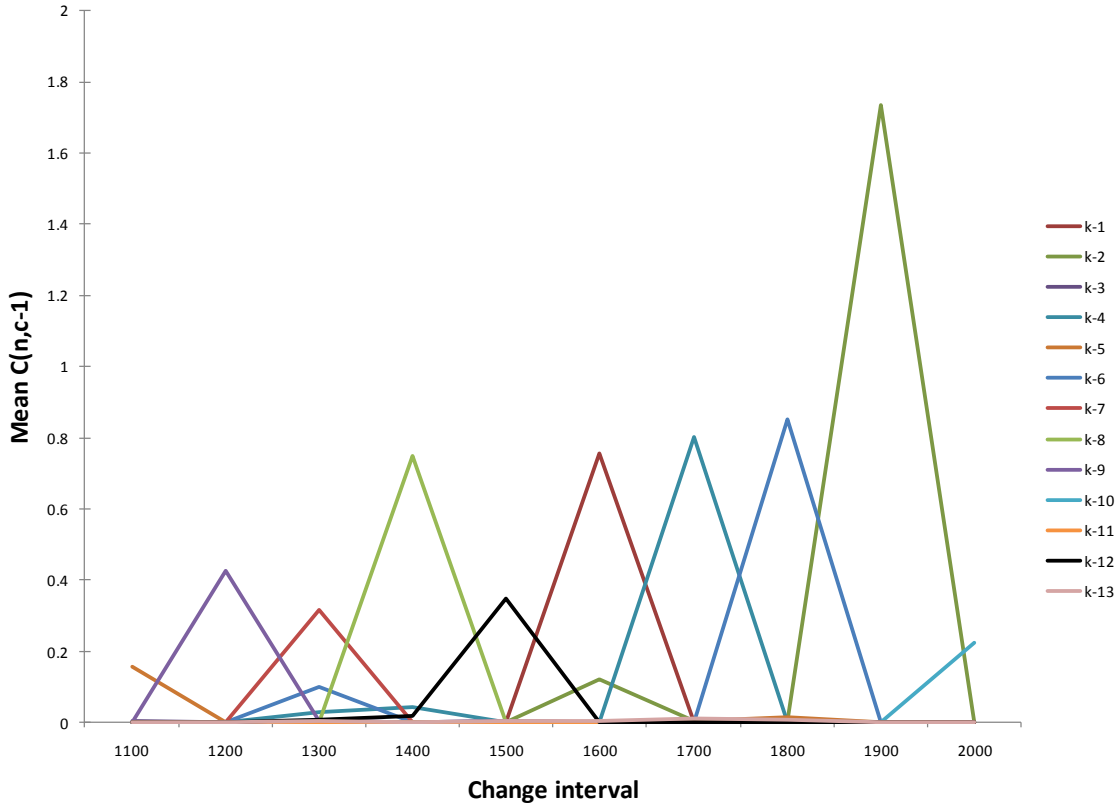


Figure 5-4. Mean change step magnitude *C(n,c-1)* by k-cluster. The mean value of *C(n, c-1)* for each 13 k-clusters is shown for each century. The value for the 11th century represents the normalized change step for verbs between 1000 AD and 1100 AD, and so on. All 13 k-clusters are depicted on the graph, but only the ten with large change steps in a single century are visible because of the scale of that change. These large magnitude values for *C(n, c-1)* result from transitions into or out of the set S of non-'ed' preterit verbs. The extreme value in the 19th century comes from *C(n, c-1) = 4.4745* for 'wed/wed'.

5.2.2.2 Nearly all k-clusters members are part of the k-cluster core

By contrast to changes to neuron populations or myelinated fiber densities, 84.6% of all 228 verbs fall into their respective k-cluster core. Figure 5-5 contrasts the three sets of data with respect to proportional core membership and proportion of data elements that are statistically significant in terms of *C(n, c-1)* (or *C(l, b, t)* for neurons and fibers).

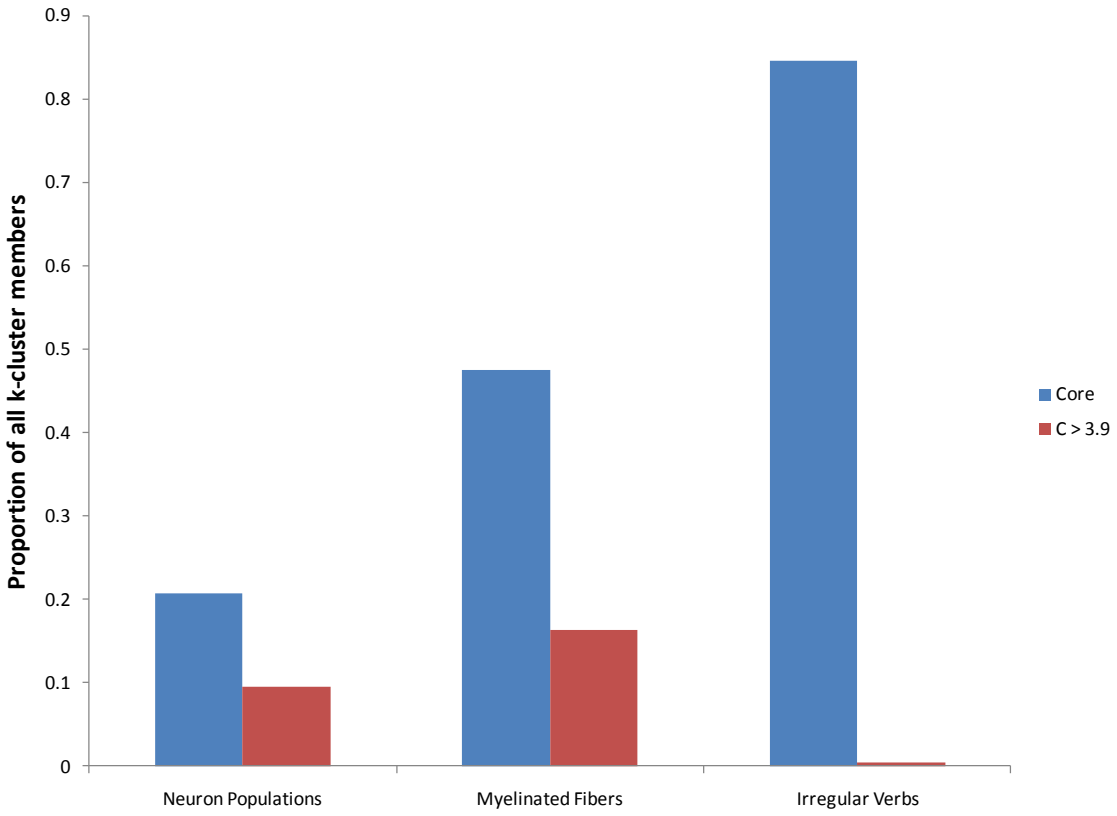


Figure 5-5. Proportion of k-cluster members that are in the k-cluster core or are statistically significant for neuron populations, myelinated fiber densities are changes to irregular verbs in English. Chapter 3 showed a functional developmental basis for synchronized changes to neuron populations in the cortex. For this data set, 20.7% of the 222 addresses belonged to one of the seven k-cluster cores. Of the 222

addresses, 9.5% had at least one value of $C(l, b, t) > 3.9$. Statistical significance of *C(l, b, t)* was significantly correlated with core k-cluster membership. Chapter 4 showed an age-related decline in *C(l, b, t)* for myelinated fiber density changes that explained 91.3% of the variance in *C(l, b, t)* for the 415 addresses where this was defined. For this data set, 47.5% of the 415 addresses were members of the seven k-cluster cores. For myelinated fiber density, statistical significance of *C(l, b, t)* was not significantly correlated with core membership; 16.4% of the addresses had significant *C(l, b, t)* values. For irregular English verbs, where membership in S depends on *R(n, c)* rather than time, 84.6% of the 228 verbs are members of one of the 13 k-cluster cores, and only one of those verbs had a statistically significant transition.

5.2.2.3 K-clusters and individual word histories

Table 5-1 summarizes the remaining 10 k-clusters, listing the century in which their greatest changes occurred, the number of verbs accessed or dropped by S during that century, their mean change step over all 10 centuries, and the mean relative usage frequency for the k-cluster.

Table 5-1. Statistics for the 10 "transition" verb clusters

Transition Century	k-Cluster	Number of accessions to S	Number of deletions from S	Mean *C(n,c-1)* over 10 centuries	Mean R(n,c) ± s.d
11th	k-5	5	1[1]	0.0180	-0.8660 ± 1.2529
12th	k-9	5	0	0.0450	0.6500 ± 1.2744
13th	k-7	10	1	0.0320	-0.8032 ± 1.1340

14th	k-8	4	15	0.0752	-1.7392 ± 1.3184
15th	k-12	1	11[2]	0.0374	-1.0079 ± 1.0547
16th	k-1	1	8	0.0762	-1.9172 ± 0.5764
17th	k-4	2	15[3]	0.0878	-1.3723 ± 1.2968
18th	k-6	0	10[4]	0.0957	-1.4018 ± 1.0086
19th	k-2	1	5[5]	0.1869	-1.1140 ± 1.9208
20th	k-10	0	2	0.0229	-0.5795 ± 0.4819

[1]'walk' subsequently deleted from S in the 18th century; k-5 has 17 stable verbs with a similar change profile

[2]'dream' and 'yell' were accessed in the 13th century, 'scrape' in the 14th century, before deletion in the 15th

[3]'blend' accessed in the 13th century, 'span' in the 14th, both deleted in the 17th

[4]'cringe' was accessed by S in the 13th century before deletion in the 18th

[5]'sneak/snuck' was an accession to S in the 16th century and a deletion in the 19th; 'snuck' is still a regional variant

These k-clusters contain the entire span for Middle English from 1100 – 1500 (Pyles, 1971; Fischer, et al., 2000) and the subsequent development of Modern English. K-5 has its largest magnitude values for *C(n,c-1)* from the interval from 1000 to 1100, and thus represents the transition from Old English to Middle English. It is exceptional among the "transitional" k-clusters in that it contains 17 of the stable verbs, but also has five accessions, of which four have such a high usage frequency that the entire cluster has a change profile resembling the other nine k-clusters containing only accessions and deletions. In order of descending relative usage frequency at accession, the five accessions to k-5 are 'take/took' (R = 3.1602) which displaced *niman* in the 11th century; 'put/put' (R = 2.1238) which appeared about 1050 according to the OED, and may also

have displaced the Old English *stellan*; ‘walke/wéolc’ (R = 1.3979) but later deleted (18th century) which represents a shift in meaning for this root from the sense of rolling in Old English to the current sense of moving about; ‘spend/spent’ (R = 1.2656); and finally ‘mislead/misled’ (R = -0.7846). As noted for Figure 2, the accessions for this century have a mean value for *R(n, c)* that is statistically significant. They have the highest relative usage frequency of the accessions in this dataset, and none of these is a member of the k-5 core. The continued active incorporation of strong preterits (‘took’, ‘wéolc’, and ‘misled’) is evident.

As evident in Figure 5-2, the 12th century accessions in k-9 continued to be unusual. K-9 consists entirely of accessions, of which four of the five had above average relative usage frequency in S in the 12th century: ‘get/got’ (R = 3.1880), possibly displacing *abiddan* (obtain); ‘build/built’ (R = 1.1219); ‘strike/struck’ (R = 0.4875); ‘spread/spread’ (R = 0.3523); and ‘thrust/thrust’ (R = -0.9120). The mean *R(n,c)* for high relative usage frequencies continued to be statistically significant, as noted previously. Unlike k-5, however, all members of k-9, including the accessions were part of the core. Active adoption of strong forms (‘got’, ‘struck’) also continues to be evident, as does the possible substitution of new roots for old.

By contrast, the remaining k-clusters have lower relative usage frequencies. For example, only three of the ten 13th century accessions in k-7 have $R(n,c) > 0$: ‘cut/cut’ (1.0248), ‘catch/caught’ (0.9234), and ‘hit/hit’ (0.4938), and mean $R(n,c) < 0$ from the 13th century for both accessions and deletions, as shown in Figure 5-2. From the 14th century on, deletions outnumber the accessions to S.

Figure 5-6 shows all of the stable non-'ed' verbs: 112 in the remaining three k-clusters, and 17 in the stable set from k-5, labeled 'k-5 (S)'. These were invisible due to scale in Figure 5-4. The vertical markers indicate time references to which we will return in the Discussion.

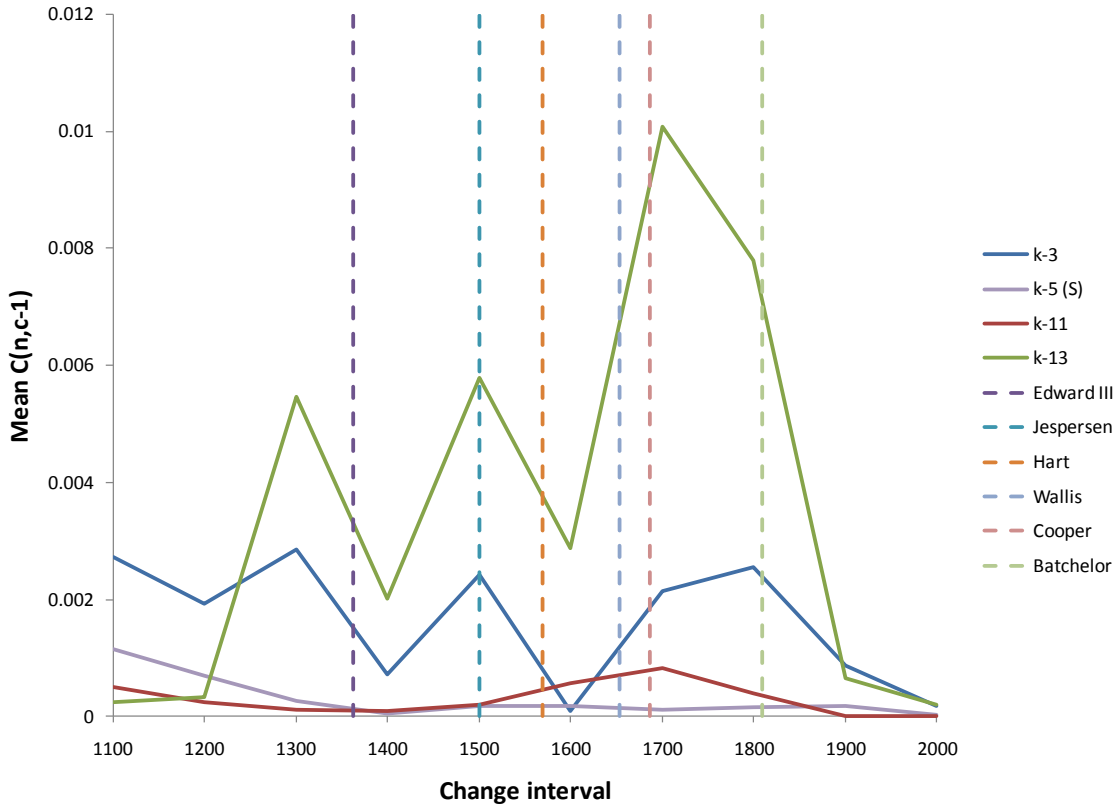


Figure 5-6. Mean *C(n,c-1)* by century for three stable k-clusters. K-3 consists of nine stable non-'ed' verbs. K-11 consists of 22 stable non-'ed' verbs. K-13 consists of 81 stable non-'ed' verbs and 4 accessions to the set S of non-'ed' verbs. The 17 stable verbs from k-5 are also shown. The vertical dashed lines indicate time references for Middle English and the transition to Modern English. 'Edward III' refers to Edward's first address to Parliament in English in 1362. 'Jespersen' refers to evidence in

Jespersen, 1909 for the initial changes in the Great Vowel Shift around 1500 that marked the end of Middle English. Hart, Wallis, Cooper and Batchelor were all orthoepists or early phoneticians who reported on exact pronunciation in their respective times: about 1570 for Hart, 1653 for Wallis, 1687 for Cooper, and 1809 for Batchelor. These were all cited in Chomsky and Halle (1968) among other works on the history of English.

Figure 5-7 shows the usage frequencies for these k-clusters by century. K-13 and k-3 had the highest and lowest mean values for $R(n, c)$, respectively, and their trajectories are nearly mirror images. Thus, k-13 had an increasing share of the set S until the 13th century that declined thereafter, while k-3 had a declining share of S until the 13th century that increased thereafter.

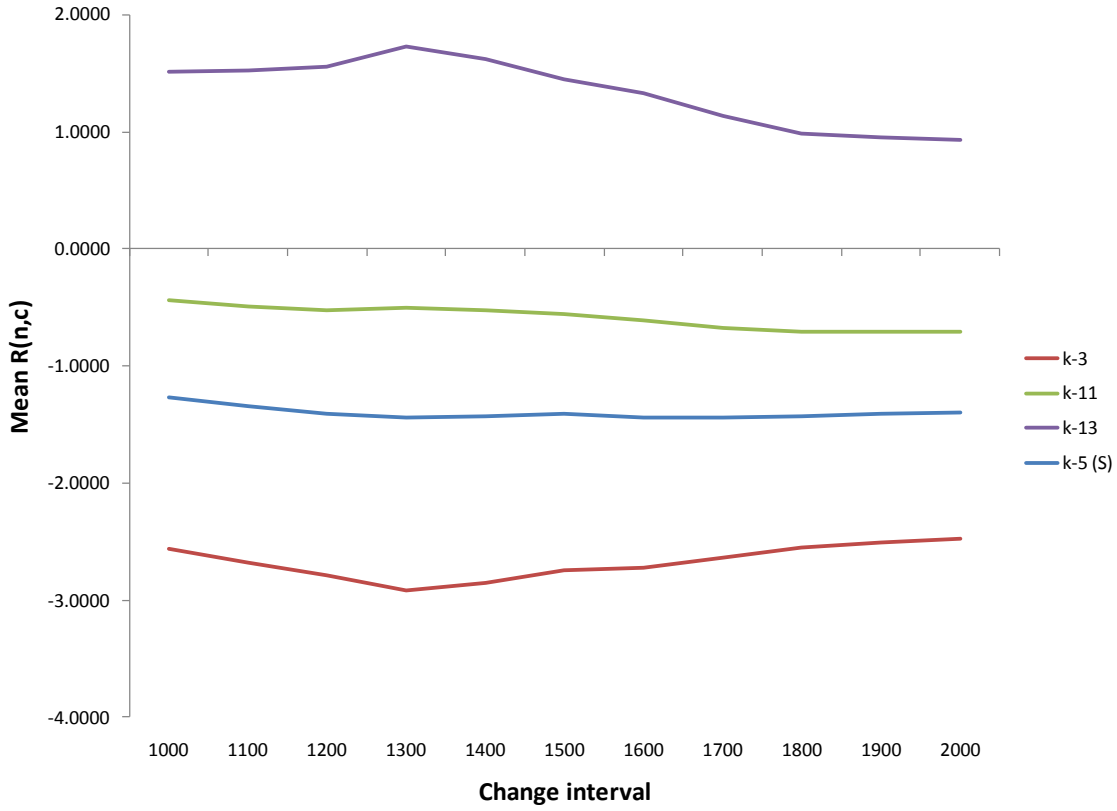


Figure 5-7. Mean relative usage frequency *R(n,c)* for the stable k-clusters. K-13 has the highest relative usage frequency (1.3542 ± 1.1598), while K-3 has the lowest (-2.6780 ± 0.6048). K-11 had a relative usage frequency near *R(n,c)= 0* (-0.5838 ± 0.1583). K-5 (S) had lower relative usage frequency, with mean *R(n,c)* = -1.3989 ± 0.3358. K-13 is the largest k-cluster with 85 members, and includes all the highest frequency stable verbs, such as 'be'.

5.3 Discussion

The L-M Hypothesis does not describe the preterit formation data in English. When examined century-by-century instead of as a two step change process, the set S of non-'ed' verbs is not monotone decreasing. Instead, S increased until the 13th century. In fact, there were no deletions from S until the 13th century, and the patterns of change from the Old English baseline in the 10th century through the 12th century evolution of Middle English demonstrate an active strong verb system and wholesale replacement of one root for another rather than the adoption of a single morpheme outside its normal range. Moreover, by focusing only on the fate of non-'ed' verbs from Old English, the L-M Hypothesis does not account for accessions, which outnumbered deletions until the 14th century, continued until the 19th in the CELEX list, and probably continue to this day, as we will see below.

When we transform sequential observations into change vectors and test for correlated changes, the temporal changes alone divide the set S of verbs that at one time or another had a documented non-'ed' preterit into ten sets (k-clusters) characterized by accessions into and deletions from S, as well as three k-clusters consisting almost entirely of stable verbs. The initial k-cluster marking the transition from Old to Middle English in the 11th century uniquely contains both stable verbs and accessions with very high relative usage frequencies.

An examination of the verbs contained in the k-clusters suffices to show that the clustering does not distinguish accession from deletion verbs when they occur in the

same century, nor does it provide a natural grouping based on verbal morphology. These k-clusters are simply correlated based on fluctuations in their relative usage frequency compared to their complement in the set S. Because set S was different from one century to the next, this means that the complements to the stable k-clusters also changed, which explains the fluctuations in Figure 5-6, particularly for the k-13, the largest k-cluster containing most stable verbs with $R(n, c) > 0$, and k-3, with the stable verbs having $R(n, c) << 0$. These are particularly useful as "yardsticks" with fixed sets of usage frequencies against which to measure the changing usage frequencies in the remainder of S. The respective magnitudes of $C(n, c\text{-}1)$ for these two k-clusters change in synchrony in Figure 5-6, and help shed some light on the actual dynamics that appeared to affect verbal morphology, particularly as Middle English became Modern English.

One interesting benchmark date, indicated in Figure 5-6, was Edward III's first address to Parliament in English in 1362. This was significant because Edward III was a Plantagenet king, descended from Norman nobility that had used Norman French at court and in the conduct of their affairs from 1066 onward. An address by the monarch to the ruling elite in Parliament in English clearly marks the emergence of English once again as the language of the ruling class in England. Chaucer was a member of his court. The low magnitude of $C(n, c\text{-}1)$ in Figure 5 indicates very little change from the 13th century, and thus a relatively stable century for the set S, regardless of the Black Death (1348) or the Hundred Years' War taking place at the time.

By contrast, the 15th century demonstrated increasing magnitudes in $C(n, c\text{-}1)$ in Figure 5-6, reflecting changes in the membership of S. The Great Vowel Shift, which

commenced the changes that produced Modern English, began during this century. The initial sound shifts noted by Jespersen (1909), who coined the name for this vowel shift, were already complete by 1500. John Hart, often cited in studies of the Great Vowel Shift by Labov (2001), Chomsky and Halle (1968) and others, was active until about 1570, and captured the system in the very early Modern English that existed during the following century, which was relatively stable according to Figure 5-6. John Hart's system is compared to Middle English in Table 5-2, which also lists the vowel systems reported by his successor English orthoepists, who all wrote on the correct pronunciation in their day.

Table 5-2. The vowel system in Middle English compared to four English orthoepists from 1570 to 1809.

Middle English	Hart	Wallis	Cooper	Batchelor
i:	ey	ey	ʌy	ʌy
e:	i:	i:	i:	iy
æ:	e:	e:		
a:	a:	æ:	e:	ey
ɔ:	o:	o:	o:	ow
o:	u:	u:	u:	uw
u:	ow	ʌw	ʌw	ɔw
e:w	yu	iw	yuw	yuw
æ:w	(y)e:w	(iw) e:w (ye:w)		
æ:y	e:y	æy	e:	ey
a:w	a:w	ɔ:	ɔ:	ɔ:
ɔ:w	o:w	əw (ʌw) (o:)	o:	ow

Middle English Hart Wallis Cooper Batchelor

ɔ:y	o:y	əy (ʌy)	ɔy	ɔy
i	i	i	i	i
e	e	e	e	e
a	a	æ	æ	æ æ:
o	o	ɔ	ɔ	ɔ ɔ:
u	u	ʌ u	ʌ u	ʌ U

Based on Chomsky and Halle (1968).

By Hart's day, the value of the vowels in various words differed between Middle English and how Hart thought it best to pronounce them, but the systems themselves continued to distinguish the same words from each other. In other words, there were no mergers, where two words such as 'meed' (a field) and 'mead' (a drink) could no longer be distinguished. This is consistent with the relative stability indicated for the 16th century in Figure 5-6.

The 17th century, however, contained the largest magnitudes for stable verb *C(n, c-1)* in the dataset. This indicates large-scale changes in the non-stable set S, and Table 1 shows that there were 2 accessions and 15 deletions in this century. One of the deletions, 'help', had a very high usage frequency (R = 1.4779). Including 'help', these were not obviously related to the vowel changes evident in John Wallis (1653-1699) who described *more* distinctions rather than mergers toward mid-century, and Christoper Cooper (1687) who described numerous mergers resulting from the changing vowel system by the end of the century. Nevertheless, both sets of observations indicate large

scale changes going on at this time: generally exceptionless changes, such as the Great Vowel Shift, and changes peculiar to particular verbs, as in k-4 and for 'help' in particular. In fact, London at the end of the 17th century provided evidence for three competing vowel systems from rhymes and puns as well as Christopher Cooper. These are sketched in Table 3. Of the three systems, only two survived into the 18th century (columns 2 and 3), of which the latter eventually predominated.

Table 5-3. Competing vowel systems in late 17th century London.

	Cooper 1685	Rhymes and puns;	Rhymes and puns; Cooper 1687
meed	i:	i:	i:
mead	e:		
made	ɛ:	e:	e:
maid			

Based on Samuels (1972).

Batchelor (1809) does not add appreciably to the possibilities for merger in Cooper (1687), but the magnitudes for *C(n, c-1)* for the 18th century are still indicative of large-scale change within the verbal systems. By Batchelor, the effects of the Great Vowel Shift were finally complete, including the diphthongs that replaced the original long vowels in Old and Middle English. The verb systems, however continued to change, including the accession of 'wed/wed', the final accession to S in the 19th century. In fact, this is the lowest frequency outlier in the data and creates the statistically significant low frequency bar in Figure 5-2. While Lieberman, et al., cite this particular verb as the next verb in danger of disappearance, it is actually the most recent to join S in

their corpus. Here is a 21st century example ('net/net') to show that the monosyllabic verbs ending in dentals still appear to be joining this class:

> She [Hillary Clinton] *net* 200,000 votes on Tuesday, but they will likely be all-but-cancelled out by an Obama win in North Carolina two weeks from now.
> Loewe D (2008)

Thus, the occurrence of 'wed/wed' as a recent accession becomes significant because it is evidence for the continued existence of an alternative attractor for preterit formation besides the *stem+ed* transform, survivals from the older strong verb vowel stem alternation, and various less regular or irregular forms. This alternative attractor may have occurred as early as the 13th century, when 'put/put' (replacing the weak Old English *stellan*) and 'cut/cut' appeared in S (Campbell, 1959; OED, 1971). From a language learner's perspective, weddings are important socially, but not nearly as frequent as situations for putting and cutting. The statistical significance of 'wed/wed' stems from this discrepancy in frequency that is also reflected by the relative infrequency of 'wed' in the CELEX corpus. The extremely low usage frequency, as well as situational frequency, argues for application of an otherwise viable paradigm rather than learning by example in this case.

While an extremely low usage form like 'wed/wed' is most likely related to adoption of an attractor that otherwise has much more frequent examplars, the highest frequency accession—'take/took', which replaced *niman* and was a statistically significant accession in the 11th century—provides a useful illustration of potential

systemic mechanisms that could affect individual word choices. Within a construction grammar context such as Tomasello's (2003), the frequency of a 'take/took' scenario within households or neighborhoods presumably did not change significantly. Yet the learned form changed nevertheless, implying that the adult role models for language acquisition had to change their behavior within previously familiar contexts that recur many times each day. Frequencies of competing forms do not necessarily apply here, but shifts in social prestige and hypercorrection may very well perform that function. Labov (1966) has shown these factors operate for sound change. Stigma or prestige factors may also account for changes in verbal syntax (Cooper, 1999). While Figure 5-2 shows that such high usage substitutions are significant only in the 11th and 12th centuries, high frequency deletions continued well beyond that time, including 'lie' (R = 1.7385) in the 15th, 'help' (R = 1.4779) in the 17th, 'walk' (R = 1.0213) in the 18th, and 'show' (R = 1.3847) in the 19th.

These considerations, as well as the decidedly sawtooth form in Figure 5-6, point to a model similar to the punctuated equilibrium hypothesis in biology. The analogy begins with the correspondence of $R(n,c)$ with homogeneous gene-flow in Ernst Myer (1954). Figure 5-1 shows that deletions from S generally occur when $R(n,c) < 0$, where 87.8% of the deletion set falls in this range. By contrast, $R(n, c) > 0$ for 39.1% of the set of accessions and 60.9% of the stable verbs. The skewness for the deletion set (0.477) is more than twice as high as for the accessions (0.234), while the stable verbs have very slight skewness (0.105). The L-M Hypothesis was an oversimplification of this basic pattern relating deletions with lower usage frequency.

The role of migration in Eldridge and Gould (1972; the E-G Hypothesis) would be played by prestige, stigma, and emulation of changing elite groups, whether or not supported by actual migration of human populations. Particularly when magnified by hypercorrection (Labov, 1966), such factors would account for high frequency substitutions like 'take' and 'put', as well as significant changes in the 17th century, whereas the L-M Hypothesis with its exceptionless and homogeneous assumption of usage frequency-dependent decay would not. The L-M Hypothesis likewise fails to predict low-frequency accessions like 'wed', as well as the alternating periods of change and stability evident in Figure 5-6.

In general, the verbs in S, like all the others that have formed a part of English, have their own history. Most words form statistical patterns that form default patterns, while others follow non-default patterns. Even fully irregular verbs like 'to be' follow patterns within particular tenses. Both words and patterns have frequencies, and relative frequencies are sufficient to describe and predict most random shifts from one category to another. Non-defaults need support, which probably requires $R(n, c) > 0$ for these classes of words, even though this is not necessary for all cases, as evidenced by 'wed/wed'. Even for the truncated set examined here, $R(n, c) > 0$ for 60.9% of the set of stable verbs. New social "in-groups" similar to the ecological isolation of populations in the E-G Hypothesis can perform the same role for language (Labov, 2001). Prestige or stigma then constitute the same conditions as migrations for evolving species, and can account for discontinuities among the general statistical regularities. Certainly, the evolution of

English preterit forms was punctuated when viewed through the lens of relative usage frequency.

With respect to a clock mechanism, there is very little support in the verbal data for a time-dependent process. Instead, the change patterns were clearly related to usage frequency, which did not change at all by hypothesis, and probably did not change much in actual fact, especially for high-frequency social transactions like ‘put’ and ‘take.’ The large proportion of core members for the verbal data in Figure 5-3 is consistent with a frequency-dependent process in the speech community, rather than a time-dependent process whether internally or externally driven. With frequencies held constant by hypothesis, both the large number of correlated changes evident in the core membership and the low number of statistically significant examples would be likely.

6. FOXP2 and Broca's Area: Complex cognitive function in light of change vector results

6.1 The clock hypothesis and a biological development model for cognitive behavior

Johnson and Morton's (1991) discussion of ocular dominance defines innate influence on emergent behavior as contingent on molecular and cellular influence. By contrast, the species-typical environment external to the organism produces primal responses, and individual but external environment of an organism produces learning. These distinctions were retained in Elman, et al., (1996) in their re-evaluation of innate behavior. They suggest that innate behavior can manifest itself at three levels as a set of constraints on what can be learned. These levels are representational (at the level of synaptic connections), architectural (at the level of cross-cortical connections), and temporal. This thesis provides results at the temporal level.

Stiles (2008) takes these observations into account in distinguishing between heritable and environmental factors in her biological development model. A key concept in Stiles' model is the inseparability of inherited factors from their environment. Inherited factors are defined as those phenotypic variances that can be accounted for by genetic variances.

> By this view, the emerging structures and functions of the brain are the product of the developmental processes created by the interaction of inherited and contextual factors. Developmental processes, and the structures and functions that derive from them, rely upon but are distinct from the inherited and contextual factors that interact to create them. This is a very different way of thinking about what it means for something to be innate. Because developmental processes rely equally on inherited and contextual information, the attempt to categorize the origins of a brain structure as the product of nature or nurture is misdirected. This model shifts the focus of inquiry to the question of development itself. Specifically, what set of developmental processes gives rise to a particular biological structure or neural mechanism, and what are the constraints on those developmental processes that lead, in most cases, to the typical trajectory of brain development?
> Stiles, 2008: p. 15

Chapter 1 introduced a clock hypothesis for the change data reviewed in this thesis, reproduced here as Figure 6-1. Chapters 3 to 5 provided three different aspects of this model, and provide a temporal component to both Elman, et al., (1996) and Stiles (2008). As noted in Chapter 4, the neuron is the locus for genetic controls affecting both neuron populations and neuron myelination. Accordingly, Conel's data for neurons applies to all three inputs levels in the clock hypothesis. Genetic and epigenetic factors are innate according to Johnson and Morton (1991) and Elman, et al., (1996). All three elements in combination are necessary to apply Stiles' (2008) perspective. Similarly, all three elements are active for Conel's myelination data as noted in Chapter 4, but with epigenetic factors, principally neuron secretions as they affect the oligodendrocytes, being most important. Stiles' (2008) model provides the greatest distinction for the verb data in Chapter 5. For Johnson and Morton (1991), the verb dataset provides a clear

example of learned behavior. For Stiles (2008) the verb dataset is a relevant environmental constraint that affected English speakers as they acquired their native language from the 10th century forward. In this case, since the normal human developmental trajectory cannot possibly depend on whether past tense forms end in '-ed', the question becomes how relevant environmental phenomena with different frequencies of occurrence can influence that developmental trajectory.

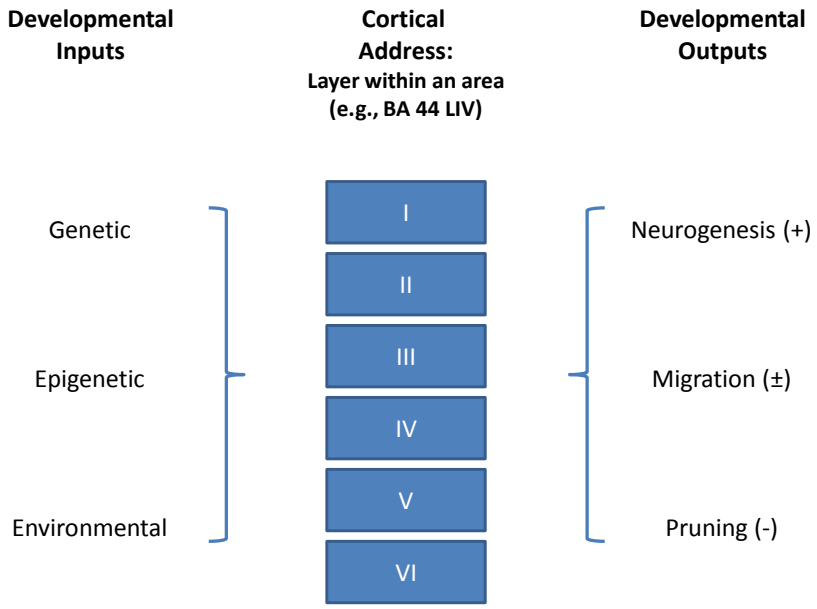


Figure 6-1. Sketch of the "clock" hypothesis for neuron populations. Genetic and epigenetic factors are innate according to the framework in Johnson and Morton (1991) and Elman, et al., (1996). In Stiles (2008) heritable and environmental factors are inseparable during development and must be taken into account at the same time. Developmental outputs for myelinated fibers would be changes to local density of those

fibers as described in Chapter 4, almost invariably increasing with age. Varying verbal frequencies, as described in Chapter 5, provide an environmental influence on genetic and epigenetic outputs, but the interaction is more subtle.

This chapter will retain Stiles' (2008) perspective, but employ the three levels of constraint discussed in Elman, et al., (1996) in reviewing the possible interaction of the speech environment on development of speech, particularly as it applies to Broca's Area as revealed in the results from Chapters 3 and 4. In the next section, the focus will be on the level of representation. The following section will explore architectural constraints. Finally, the temporal constraints that emerge from using the change vector method will be explored. Since the principal results from Chapters 3 and 4 point to a generally vision-centered sensorimotor theme to development to six years old, this focus on speech and Broca's Area is but a subset of the overall developmental patterns revealed in Conel's data. However, applying the verbal dataset will allow this discussion to explore both sides of Stiles' inseparable heritable/environmental model.

6.1.1 FOXP2 and developmental constraints at the representational level

Elman, et al., (1996: 25) define innateness at the representational level as "in the form of fine-grained patterns of synaptic connectivity at the cortical level, i.e., cortical micro-circuitry. To the best of our knowledge at the present time, this is how the brain stores its representations, whether they are innate or acquired." As noted earlier, this

level incorporates both genetic and cellular activity. Consequently, the effects of the foxp2 gene and the FOXP2 molecule, discussed in Chapter 1, would manifest themselves at this level.

For Stiles (2008: 15), “The specific causal role of a gene product in a biological process requires assessment at the molecular level, and at that level it is the interaction of inherited factors operating within specific contexts that ultimately defines a gene's function or functions and its role in development." Enard, et al., (2009) studied the principal effects of substituting the two human FOXP2 amino acids that have undergone positive selection after divergence of the human and chimpanzee lines into mice, and thus provides an example at this level. An assay of more than 300 phenotypic differences with wild type mice shows that the significant effects occur only in the brain. These include increased dendrite lengths and greater long term depression (LTD) in medium spiny neurons in the striatum. In addition, the mice, which grow into healthy adults unlike foxp2 knockout mice, have qualitatively different ultrasonic isolation calls, decreased exploratory behavior and decreased dopamine concentrations in the brain compared to wild type. This suggests that the humanized foxp2 allele affects the basal ganglia. Moreover, Enard, et al., (2009) show that humanized foxp2 affects medium spiny neurons with D1 dopamine receptors, which indicates that foxp2 functions within the direct dopaminergic pathways in the basal ganglia, and thus with disinhibition of thalamic inputs to the cortex (Parent and Hazrati, 1995). The thalamo-cortical loop is an architectural level entity and will be discussed below in section 6.1.2.

As noted in Chapter 1, foxp2 is among the five percent most conserved genes in human-mouse pairings (Enard, et al., 2002). Foxp2 was isolated because of a point mutation in a human family afflicted with inherited difficulties with pronunciation and grammar (Varga-Khadem, et al., 2005). Enard, et al., (2009) duplicated this point mutation and confirm that humanized foxp2 effects are reversed when it is present. Similarly, their phenotypic assay confirms that the selective pressure for the human form of the gene in the past 200,000 years (Enard, et al., 2002) was likely the result of brain function, despite expression of FOXP2 in multiple other tissues, such as the epithelial tissue of airway branches in lungs, in the outer mesoderm of the intestines, and in the outflow tract of the atria of the heart (Shu et al., 2001).

As pointed out in Chapter 1, there is a general upregulation of gene expression when humans are compared to chimpanzees (Preuss, et al., 2004), but foxp2 does not seem to be a clear example of this. Konopka, et al., (2009) compared the transcriptional regulation by human and chimpanzee FOXP2 *in vitro*, and showed that FOXP2 is associated with genetic cascades in both species where some genes are significantly upregulated and some are significantly downregulated in humans compared to chimpanzees. There was a function difference between up- and downregulation by FOXP2, however. Genes upregulated by human FOXP2 compared to $FOXP2^{chimp}$ were involved in transcriptional regulation of gene expression and in cell to cell signaling, whereas downregulated genes were related to protein and cell regulation. The potential creation of a phosphorylation site for protein kinase C at one of the two human sites (325) explored in Enard, et al., (2002) and noted in Chapter 1 has not yet been confirmed or

discounted. If FOXP2 function were contingent on a conformal change triggered by protein kinase C, one would not necessarily expect it to have a converse effect on protein kinase C in Konopka, et al. (2009). FOXP2 does upregulate two genes related to tyrosine kinases, however (Konopka, et al., 2009: Suppl Table 5), consistent with cell signaling.

Chapter 1 pointed out numerous parallels between human and chimpanzee anatomical features related to Broca's Area. Konopka, et al., (2009) adds another. For the perisylvian cortex, both human and chimpanzee foxp2 have an accelerated conserved non-coding sequence associated with elevated FOXP2 levels (Konopka, et al., 2009: Suppl Table 5).

6.1.2 Broca's Area and evidence for developmental constraints at the architectural level

Ullman (2004) has proposed a declarative/procedural model for human language that is written at the architectural level. In this model, word-specific knowledge is learned and used in a declarative component that is distinct from a rule-based, procedural component. The declarative component, related to declarative memory, is based on the medial temporal area of the brain, including the dentate gyrus, subiculum, the entorhinal perirhinal and parahippocampal cortex, and the hippocampus itself (Squire and Knowlton, 2000; Suzuki and Eichenbaum, 2000). These regions form a hierarchy, with the hippocampus strongly connected to the entorhinal cortex, which is strongly connected in turn to the perirhinal and parahippocampal cortices, and ultimately to temporal and parietal areas in the neocortex (Suzuki and Amaral, 1994).

These structures are implicated in general memory functions, including encoding, consolidation, and retrieval (Buckner and Wheeler, 2001; Eichenbaum and Cohen, 2001; Squire and Knowlton, 2000). In humans, declarative memory becomes less dependent on the medial temporal structures and shifts to neocortical areas, especially in the temporal lobe (Hodges and Patterson, 1997; Squire, et al., 2001), where different areas may even specialize in different types of knowledge, such as tools or animals (Damasio, et al., 1996; Martin, et al., 2000). In addition, humans appear to have acquired the ability to remember sound sequences using these areas, whereas detailed auditory memory appears to be absent in macaques (Fritz, et al., 2005). Thus, while all animals require an ability to remember, this portion of the neural architecture in humans is now adapted to support the "arbitrariness of the sign" needed to use vocabularies in the tens of thousands of words for normal human language, where each of those words has no direct connection to that which it refers (Saussure, 1916; Hurford, 2004).

Ullman (2004) also includes the ventro-lateral prefrontal cortex (VL-PFC) within this system. This area includes Broca's Area and BA 47, and appears to be involved in working memory (Smith and Jonides, 1999; Braver et al., 2001). With respect to language, the posterior portion of this region (BA 44/6) is associated with phonology, whereas the anterior region (BA 45/47) is implicated in semantic functions (Fiez, 1997; Poldrack, et al., 1999). Generally, Ullman associates the declarative system with the so-called ventral stream (Goodale & Milner, 1992; Ungerleider and Mishkin, 1982).

The procedural component of Ullman's model consists of Broca's Area, superior temporal cortex, the basal ganglia and cerebellum, as well as portions of the parietal lobe

related to the mirror neuron system (Ullman, 2004). As noted in Chapter 1, mutant foxp2 is implicated in KE family speech deficits, and FOXP2 is expressed in subgranular layers in Broca's Area. In Section 6.1.1, foxp2 was also associated with medium spiny neurons in the direct pathway in the basal ganglia, which functions to disinhibit neocortical targets of the thalamocortical loop (Young and Penney, 1993). The striatum is associated with sequences of stereotypical actions (Aldridge and Berridge, 1998), and the procedural system in Ullman's model is similarly stereotypical, slowly acquired and "encapsulated." Language depends on the two systems inter-operating, possibly competing, and compensating for each other (Ullman, 2004).

Reading, which cannot have been the outcome of a biological evolutionary process, provides an interesting contrast to this kind of solution to the evolution of spoken language. Dehaene (2009) ascribes the ability of humans to invent and teach reading to a "letterbox" in the occipito-temporal cortex (BA 37) that allowed humans to map line junction combinations that occur in natural scenes to their arbitrary but systematic phonetic counterparts in any of a wide range of existing writing systems, thus enabling a linkage to the language system. While reading is essentially an overlearned skill, it results in an automatic response to visual stimuli that is not normally subject to conscious intervention in skilled readers. While all normal humans can distinguish between phonemes in their native language, they do not recognize "phonemes" as discrete sound segments unless they can read (Morais, et al., 1986). Thus, the phonemes that are among the traditional list of "linguistic universals" (Chomsky and Halle, 1968) are not recognizable to human subjects as discrete entities unless the subject is literate.

By contrast, sequences of sounds or signs (as in American Sign Language) are "easily" acquired, as opposed to reading and writing (Arbib, 2004). This is an argument for language-related functions that are not "encapsulated" to which we will return in Section 6.3.

Like Ullman's declarative/procedural model, the motor theory of language originally developed in Liberman, et al., (1952) contained a "speech is special" architecture. However, the component of this theory whereby memory mediated between articulation and the perception of speech sounds was subsequently abandoned, as was a later modular version (Galantucci, et al., 2006). Components of the motor theory of speech that did persist and appear to be well-founded are that speech perception is tied directly to perception of oral, facial, as well as manual gestures; and that the motor system is recruited to enable the perception of speech (Galantucci, et al., 2006). These elements first emerged in trying to explain the effect of co-articulation of phonemes, which led to the generalization that when sequences of sounds that have different acoustic properties are produced by the same sequences of articulatory gestures, they are perceived to be the same (i. e., they are the same phonemes; Liberman, 1957). Liberman, et al. (1967) concluded that these features were the result of unique human phylogenetic adaption.

The McGurk Effect, in which a subject hearing /ba/ but seeing /ga/ will interpret the sound to be /da/ (McGurk and MacDonald, 1976), provides an example of the linkage of speech perception to articulatory gestures. Other supporting evidence includes better perception of speech in noisy environments when the listener can observe the speaker's

face (Sumby and Pollack, 1954), as well as superior speed in sound imitation when the face can be seen (Fowler, et al., 2003; Kozhevnikov and Chistovich, 1965; Porter and Castellanos, 1980; Porter and Lubker, 1980).

As to the involvement of the motor system with speech perception, Galantucci, et al., (2006) cite the mirror neuron system as one particularly important strand of evidence. In addition, they point out that the principle of parity in the most recent versions of the motor theory of speech perception requires the co-evolution of production and perception mechanisms. Under the principle of parity, for any pair of conspecifics, a message for one is also a message for the other, and the content of the message one sends will be the same as the content the other perceives. Whenever these parity conditions are violated in the co-evolution of the production and perception of communication, there is an opportunity for speciation (Ryan and Wilczynski, 1988; Shaw, 2000).

Sinha (2004) argues that communication behavior can provide a Baldwin effect on evolution, whereby the elaboration of symbols as in human language emerges from an initial, simpler signaling process. The potential from speciation provided by the parity principle in Ryan and Wilczynski (1988) would provide a mechanism for this effect. Arbib (2004) further argues that neuron plasticity provided the needed connectivity among brain structures for spoken language, rather than the prewiring required at the representation level, just as Dehaene (2009) argued for plasticity providing the needed connectivity in reading. This may suggest that future human evolution will also show a Baldwin effect for reading. Moreover, with respect to spoken language again, Sinha (2004) contrasts his perspective with the "nativist" construct that is more compatible with

innate representational structures. Like Stiles (2008), Sinha (2004) emphasizes the co-dependence of encoded genetic information and the external environment, which he claims results in a unidirectional developmental flow whereby the commitment of neurons to a particular path given a particular environment forecloses that possibility in the future. Thus, first languages are acquired easily, but languages acquired after that sensitive period are not.

As a counterpoint to development based on genetic information instantiated within an architecture, Sinha (2004) and Deacon (1997) have suggested that the arbitrary forms in spoken language have themselves evolved in accelerated fashion to favor structural constraints in the human brain, just as Dahaene (2009) suggests that arbitrary written forms have evolved to facilitate learning constraints in the reading circuits in the brain. In the next section, we will take up the evidence in our change vector data for innate temporal patterns related to Broca's Area and spoken language. In Section 6.2, we will then examine the contrasting pattern in change vector data that English irregular verb evolution shows.

6.1.3 Neuron and myelinated fiber change vector data and evidence for developmental timing constraints and the clock hypothesis

The corollary to innate behavior emerging from architectural adjustments based on neural plasticity in the context of an organism interacting with its external environment is some mechanism to organize the architectural changes. As noted earlier,

the Conel data on neuron populations and myelinated fiber densities are silent as to the precise nature of such a mechanism, but they do provide evidence for coordinated, simultaneous changes during development. Chapter 3 provided evidence for relatively small and complex neuron population cores within k-clusters that had a one-to-one relationship with the seven change intervals in Conel's data. Chapter 4 provided similar data for larger and simpler myelinated fiber density cores that also had a one-to-one relationship with the seven change intervals. The neuron population cores and the statistically significant changes to specific neuron addresses showed a general correlation with sensorimotor functions centered on vision. Likewise, the statistically significant fiber addresses occurred only in primary motor and primary sensory cortex, while the cores showed a systematic shift favoring subgranular layers at birth toward supragranular layers by six years. One can surmise that the neuron pattern of smaller, complex cores reflects activity dependence whereas the myelinated fiber pattern with nearly fifty percent of each k-cluster also being part of a single large k-cluster core reflects greater internal, and thus, innate coordination. The two datasets are not sufficient in themselves to evaluate this conjecture statistically.

Figure 6-2 shows the change trajectory for posterior Broca's Area (BA 44) at each of the addresses in a neuron or myelinated fiber sample column. In each case, the k-cluster membership of any address can be determined by its maximum value. Membership in a maximal clique "core" is indicated by a filled circle at the appropriate apex. Figure 6-3 shows the comparable figure for anterior Broca's Area (BA 45). Membership of the k-cluster cores for BA 44 and BA 45 is listed in Tables 6-1 and 6-2,

respectively. In the figures and tables, layers for neuron populations are listed as Roman numerals, e.g., LIV, whereas the layers for the myelinated fibers are indicated by alphanumerics using Arabic numerals, e.g., L3c-4.

Both BA 44 and BA 45 have two neuron population core addresses. BA 44 LII is the earlier core member at 3 months. BA 44 LV and BA 45 LVI co-occur at 24 months and have the largest C(r, c-1) magnitudes for core addresses in Broca's Area. BA 45 LIV is a core address at 72 months. Both BA 44 and BA 45 have very large C(r, c-1) magnitudes for myelinated fibers at 6 and 15 months that do not reflect a general cortical pattern. By contrast, both participate in a cortex-wide increase in myelinated fiber density for layers L2 and L3a at 48 months.

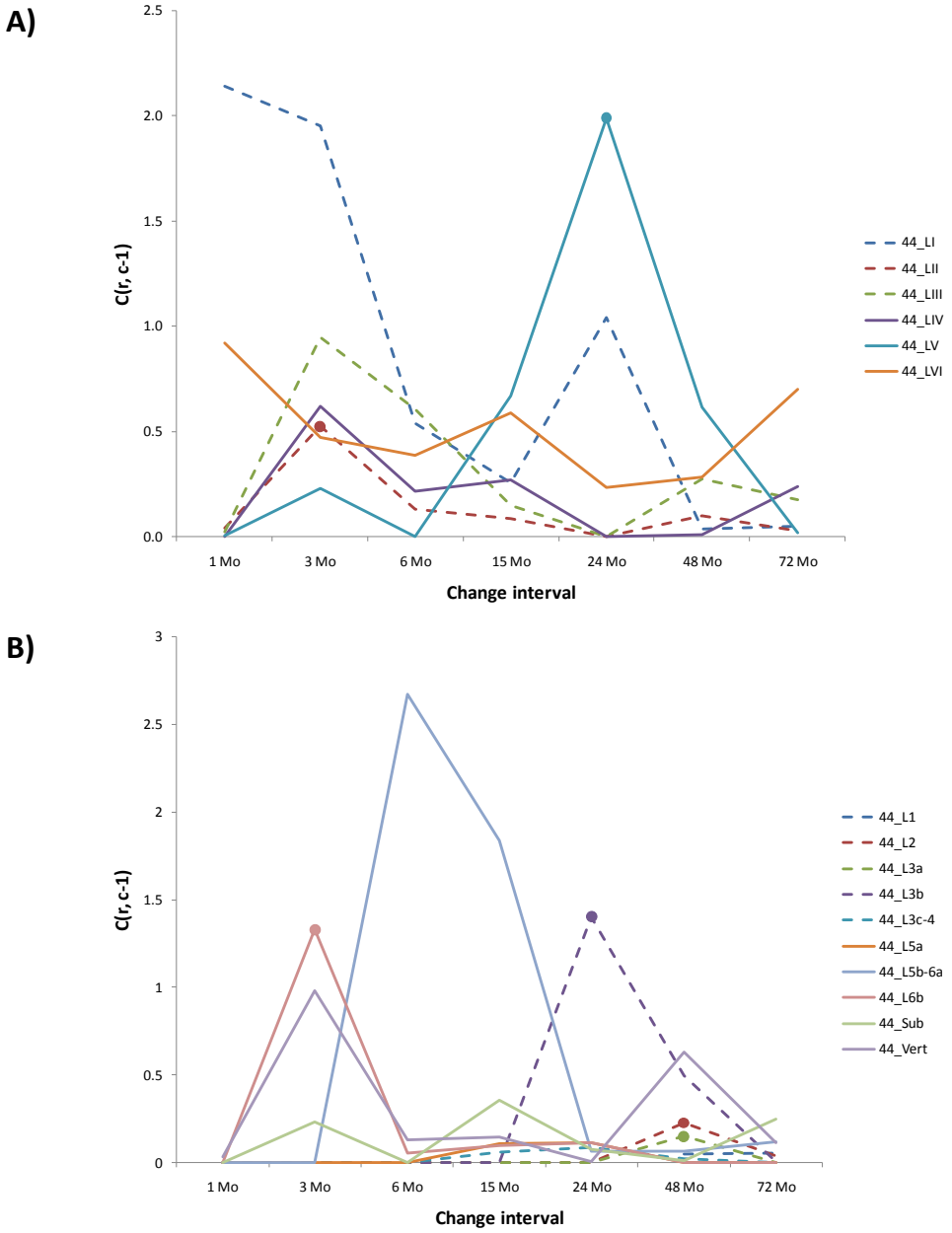


Figure 6-2. Change vector components of all sample column addresses in BA 44.

Core membership is indicated by a filled circle on the appropriate apex for a core address. A) C(r, c-1) values for the six neuron population addresses by change interval. Curves

for the supragranular layers are labeled with dashed lines and colors as indicated in the legend. Layer IV and the subgranular layers have solid curves. B) C(r, c-1) values for the ten myelinated fiber addresses by change interval. Curves for the supragranular layers, including Layer 3c-4, are labeled with dashed lines; subgranular layers have solid lines. BA 44 has both a core neuron and core fiber address in the 3 month interval as well as the 24 month interval. These are listed in Table 6-1, as well as the core membership of the core membership at 48 months that contains two BA 44 addresses. The neuron population core member at 3 months at LII is contemporary with maximal C(r, c-1) values for LIII and LIV as well. The change pattern for LI is local to BA 44 and not common across the cortex. The two core fiber addresses at 48 months are from two layers that comprise 73.7 percent of the k-cluster core during that change interval. The double spike pattern for vertical fibers with peaks at 3 and 48 months is not a core pattern, and thus more peculiar to BA 44. Similarly, the large-scale changes for L5b-6a at 6 and 15 months are echoed by BA 45, but not across the cortex.

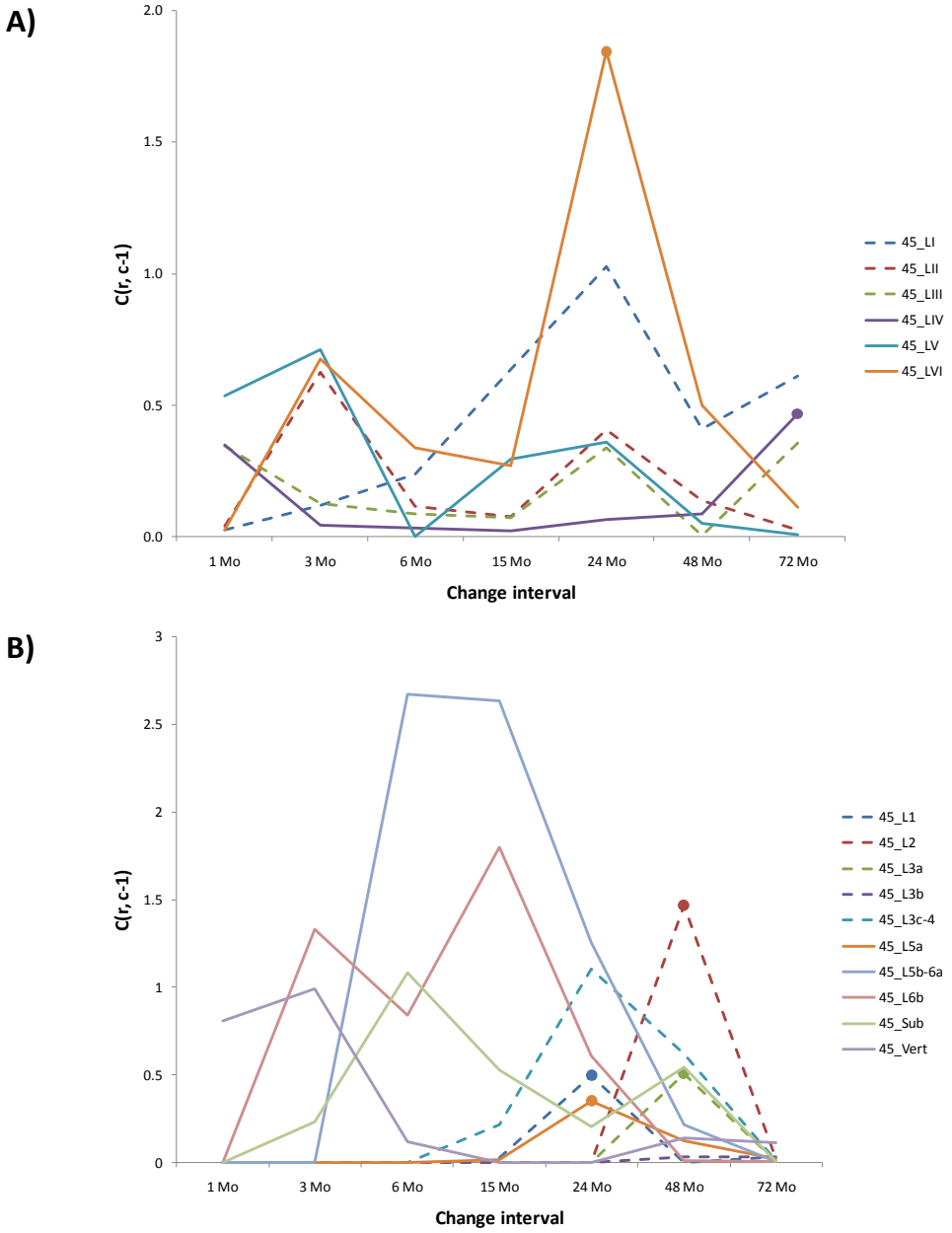


Figure 6-3. Change vector components of all sample column addresses in BA 45.

Core membership is indicated by a filled circle on the appropriate apex for a core address.

A) C(r, c-1) values for the six neuron population addresses by change interval. Curves

for the supragranular layers are labeled with dashed lines and colors as indicated in the legend. Layer IV and the subgranular layers have solid curves. B) C(r, c-1) values for the ten myelinated fiber addresses by change interval. Curves for the supragranular layers, including Layer 3c-4, are labeled with dashed lines; subgranular layers have solid lines. BA 45 LVI is a neuron population core address that co-occurs at 24 months with BA 44 LV, as well as myelinated fiber core membership for L1 and L5a. The two core myelinated fiber addresses at L2 and 3a for the 48 month change interval do not correspond to a neuron core and are part of a cortex-wide increase in C(r, c-1) magnitude for those fiber layers during that period.

Table 6-1 lists the core membership for BA 44. Both neuron cores correspond to a myelinated fiber core at 3 months and 24 months. The two core addresses at 48 months for L2 and 3a are part of a general cortical pattern at period in development.

Table 6-1. Core membership for addresses in BA 44.

BA 44 change interval	Type	Core Member	Other core members
3 Mo	Neuron	LII	6 LIII & V; 39 LIV; 38 LI; 13 LIV
	Fiber	L6b	4 Hn L1, 3c-4 & 5a; 4 Hd L1, 3c-4 & 5a; 4 T L6b; 6 L1; 9 Sub; 3 T L5a & sub; 3 Hn L3c-4; 1 T L5a; 1 L L5a; 1 Hd L6b; 17 L3c-4 & 5a; 42 L5b-6a; 41 L5b-6a
24 Mo	Neuron	LV	45 LVI; 42 LVI; 13 LI
	Fiber	L3b	4 Hn L2 & 3a; 4 Hd L3b; 6 L2 & 3a; 8_9 sub; 45 L1 & 5a; 10 L3c-4; 1 Hd L3a; 7 L5b-6a; 40 L3b; 39 L3b; 37 L3b; 19 L1 & 5b-6a; 22 L2, 3a, 3b & 5b-6a; 42 L2, 3a & 3b; 41 L2 & 3a; 36 L3c-4 & 6b; 13 L5a; 24 L3b & 6b; 23/31 L5b-6a; 30 L3b; 29 L3b

48 Mo	Fiber	**L2 & 3a**	4 Hd 3a; 6_8_9 **L2 & 3a**; 8 **L2, 3a** & 6b; 8_9 L2 & 3b; 45 **L2 & 3a**; 9 L5a; 46 L3b; 45r L3b; 47 **L2, 3a** & 3b; 3 Hd L3b; 1 L L3b;1 T L2 & 3b; 1 Hn L2; 7 **L2 & 3a**; 40 **L2 & 3a**; 39 **L2 & 3a**; 37 **L2 & 3a**; 19 L3a; 18 **L2 & 3a**; 17 **L2, 3a** & 3b;22 sub & vert; 21 **L2, 3a**, 3b & vert; 36 **L2 & 3a**; 14 **L2, 3a** & 3b; 13 L2; 24 **L2 & 3a**; 23/31 **L2 & 3a**; 30 **L2 & 3a**; 29 **L2 & 3a**

"BA" has been omitted from Brodmann Areas to avoid repetition. Layers are listed after an "L" which is also used for BA 1 for the leg in "1 L L3b." Bolded layers for fiber addresses at 48 months highlight the 18 occurrences of that pair during that change interval. There are 57 addresses in the k-cluster core during that interval altogether. Other abbreviations: Hd = head, Hn = hand, r = rostral, T = trunk, sub = subcortical, vert = vertical.

At 3 months, Broca's Area activates when the child is listening to sentences (Dehaene-Labertz and Dehaene, 1994; Cheour, et al., 1998; Dehaene-Lambertz, et al., 2002; Pena, et al., 2003). By this time the child is attending to the contours of the speech signal (Mehler, et al., 1988), and the capacity to distinguish sounds, such as /ba/ from /ga/ was established soon after birth, and thus in an earlier change interval (Kuhl, 2004). During the 3 month change interval, children are generally transitioning from non-speech sound production to production of vowel-like sounds. This enables the child to produce language specific sounds during the next change interval to 6 months, based on statistical learning of the distributional frequencies in the child's environment according to Kuhl (2004).

The Conel data show that this coincides with BA 44 LII as a core address, associated with BA 6 (supplementary motor), BA 39 (angular gyrus), BA 13 (insula), and BA 38 (temporal pole). Like Broca's Area, in forward speech in the mother language, BA 39 is also activated in 3-month-old subjects listening to sentences (Dehaene-Lambertz, et al., 2002). While listening to delayed repetition or modification of sentences after intervals up to 14 seconds, infants at 3 months use Broca's Area to reinforce the repetition or to detect the change. Moreover, there is a temporal gradient during this process that begins at Heschl's gyrus with significant time lags as activity moves posteriorly to the angular gyrus and anteriorly down the temporal lobe to Broca's Area (Dehaene-Lambertz, et al., 2006). As is clear from Figure 6-2, there are no significant changes in myelination in the 1 month change interval, but the 3 month interval also has BA 44 L6b as a core address, corresponding to seven other addresses in primary motor cortex (BA 4), six in primary sensory cortex (BA 3 and 1), as well as two primary vision addresses (BA 17) and primary and secondary auditory addresses (BA 41 and 42, respectively). All of this is fully consistent with the sensorimotor pattern in Chapter 3 and 4, as well as the motor theory of speech production discussed earlier in this chapter. In fact, learning patterns observed in children at this age suggest a "magnet" or "prototype" mechanism for derivation of local language sounds (Kuhl, et al., 1992; Kuhl, 1991; Rosch, 1975). Moreover, as BA 44 L6b is subgranular, this address is also consistent with the articulatory issues related to the KE family foxp2 mutation described in Chapter 1, as well as the participation of Broca's Area in a language-related thalamocortical loop. By the same token, BA 44 LII is supragranular, which potentially

implicates mirror neuron connections and processing in a different set of circuits as discussed in Chapters 1 and 3.

The two portions of Broca's Area have core addresses at 24 months, as shown above for BA 44 and in Table 6-2 for BA 45.

Table 6-2. Core membership for addresses in BA 45.

BA 45 change interval	Type	Core Member	Other core members
24 Mo	Neuron	LVI	44 LV; 42 LVI; 13 LI
	Fiber	L1 & 5a	4 Hn L2 & 3a; 4 Hd L3b; 6 L2 & 3a; 8_9 sub; 44 L3b; 10 L3c-4; 1 Hd L3a; 7 L5b-6a; 40 L3b; 39 L3b; 37 L3b; 19 L1 & 5b-6a; 22 L2, 3a, 3b & 5b-6a; 42 L2, 3a & 3b; 41 L2 & 3a; 36 L3c-4 & 6b; 13 L5a; 24 L3b & 6b; 23/31 L5b-6a; 30 L3b; 29 L3b
48 Mo	Fiber	**L2 & 3a**	4 Hd 3a; 6_8_9 **L2 & 3a**; 8 **L2, 3a** & 6b; 8_9 L2 & 3b; 44 **L2 & 3a**; 9 L5a; 46 L3b; 45r L3b; 47 **L2, 3a** & 3b; 3 Hd L3b; 1 L L3b;1 T L2 & 3b; 1 Hn L2; 7 **L2 & 3a**; 40 **L2 & 3a**; 39 **L2 & 3a**; 37 **L2 & 3a**; 19 L3a; 18 **L2 & 3a**; 17 **L2, 3a** & 3b;22 sub & vert; 21 **L2, 3a**, 3b & vert; 36 **L2 & 3a**; 14 **L2, 3a** & 3b; 13 L2; 24 **L2 & 3a**; 23/31 **L2 & 3a**; 30 **L2 & 3a**; 29 **L2 & 3a**
72 Mo	Neuron	LIV	6 LVI; 22 LI; 24 LIII

At 24 months, BA 44 LV and BA 45 LVI are parts of a discrete 4-address component of the k-cluster core. The other members of this group are BA 42 (secondary auditory cortex) and BA 13 (insular cortex). At this change interval, BA 44 L3b and BA 45 L1 and 5a form part of a myelinated fiber k-cluster core with a total of 32 members. The sensorimotor areas form a much smaller portion than at 3 months, including three addresses in primary motor cortex (BA 4), two in BA 6, and one in primary sensorimotor

cortex (BA 1 for the head). There are four addresses in the parietal lobe (BA 7, 37, 39 and 40), and nine from the auditory areas in the temporal lobe (BA 22, 41, and 42). As noted previously in Chapter 1, the parietal areas are possibly linked to mirror neuron circuits, whereas most of the remaining neuron and myelinated fiber core areas are clearly relevant to audition and language. By this point during development, both BA 44 and 45 have undergone significant increases in myelinated fiber density, as shown on Figures 6-2 B and 6-3 B. In addition, since both Broca's Area addresses are subgranular, the possibility that these core addresses would be affected by mutations in foxp2 exists for this change interval as well as the earlier one at 3 months. By this stage in language development, children have learned the phonemes in their mother language, have ceased to distinguish phonemes in other languages, and have begun to produce words and sentences in their native language (Kuhl, 2004).

As mentioned previously, the two BA 44 and 45 addresses that occur in the 48 month change interval appear to be part of a cortex-wide trend toward myelination of layers L2 and 3a. These may thus not provide a functional clue for Broca's Area in particular.

The neuron core address for BA 45 is linked to BA 6 (supplementary motor cortex), BA 22 (tertiary auditory cortex) and BA 24 (anterior cingulate). These continue to be consistent with the linkage of audition, motor function, and language.

6.2 Inseparability of external and inherited factors: possible effects on the verbal data

According to Stiles (2008) neuron population changes are largely neurotrophin and activity-based, whereas changes to myelinated fiber densities is largely controlled by neuron secretions and driven by neuron firing activity only to a limited extent. As we have seen, the cores for neuron populations are smaller compared to the respective k-cluster than they are for myelinated fibers, and this relationship was maintained for Broca's Area in particular. Moreover, as noted in Chapters 3 and 4, neuron cores were more complex than fiber cores, with one case of dual maximal cliques (where BA 44 and 45 were in one of the two groups), as well as cores with overlapping and "unique" members. Finally, individual addresses for both neurons and myelinated fibers were not members of k-cluster cores, but did have significantly large values for C(r, c-1).

For Broca's Area, there were no outlier addresses, but its addresses with core membership followed these patterns. The individual subjects who contributed to Conel's datasets were likewise immersed in a speech environment that constituted the inseparable second half of an innate behavior according the Stiles' (2008) construct. The verbal material described in Chapter 5 provides the other half.

That dataset has three characteristics that distinguish it from the neuron and myelinated fiber sets. First, only one of the 228 verbs had a statistically significant value for C(n, c-1). That verb was 'wed/wed' which appeared to be an instance of a new preterit attractor. There was certainly no evidence that it was an internally driven shift of

pattern. Second, k-clusters showed no evidence for coordinated shifts by members of particular morphological groups, except the wholesale replacement of 'niman' by 'take' in the late Old English/early Middle English period. There is therefore almost no evidence to rule out a purely random process for accessions and deletions into the set S of non-'ed' verbs. It is interesting that k-cluster cores for the verb data comprise 84.6% of the k-clusters, by contrast to neuron populations, where the core is 20.7%, and myelinated fiber density, where it is 47.5%. Finally, membership in the set S depended on usage frequency that was assumed to be reasonably stable, and thus independent of time. This essentially rules out the operation of a clock mechanism, and is consistent with an independent stochastic process.

For spoken language, Deacon (1997) has suggested that this kind of environment provided the opportunity to the speech communities to evolve their respective languages to match human learning constraints. Cooper (1999) suggests an attractor-based mechanism for this process that conforms to "ambiguity landscapes" that arise as random changes occur. Dehaene (2009) suggests a similar mechanism for the evolution of writing systems in historical times that has resulted in relatively efficient methods for writing language despite the impossibility of evolutionary support for learning to read in that short period of time.

6.3 Conclusions

Broadly speaking, Chapter 1 described the puzzle of Broca's Area: the first human

brain area to which a particular function could be ascribed, but where that linguistic function was impossible in non-human primates in which the equivalent to that area also appears. Applying the change vector method introduced in Chapter 2 to Conel's neuron and myelinated fiber data in Chapters 3 and 4, it appeared that human cortical development in general appeared to be driven by a clock-like sensorimotor process that particularly favored vision. This was consistent with primate development, perhaps including the crucial mirror neuron system, and also consistent with the details of the development of Broca's Area as reviewed in this chapter.

As noted in Galantucci, et al., (2006), the discovery of the mirror neuron system has resulted in wider acceptance of the motor theory of language perception outside of linguistics than within it, but that the theory has received a considerable amount of theoretical support, particularly in terms of architectural innateness in the framework described at the beginning of this chapter. While many linguists have ignored all of these issues in favor of Chomsky's (1986) nativist argument (including Pinker, 1994), the branch of linguistics known as "cognitive linguistics" (e.g., Taylor, 2002) provides a line of argument entirely consistent with a sensorimotor basis for language that also incorporate mirror neurons and a heavy reliance on thalamocortical pathways that appears to emerge for BA 44 at 3 months in the Conel data.

Figure 6-4 provides a final picture to show these linkages. It is based on Croft (1991) who used this breakdown of cause and effect to explain thematic roles that produce such things as case systems (e.g., 'he' for subjects, 'him' for objects, 'his' for possession, etc.) across hundreds of languages. Figure 6-4 also provides a depiction of

how the sensorimotor perception system underlying the mirror system can also explain the fact that "languages encode either experience or stimulus as the causally antecedent participant" (Croft, 2001: 164). That is, languages either mark the object and leave the subject unmarked, such as English (contrast subject = 'he', unmarked, and object = 'him,' marked and therefore longer); or they have the opposite pattern. The pattern in English is called "accusative" where the object is marked ('him'), and the contrasting ergative pattern is called "absolutive." These terms appear in the figure below.

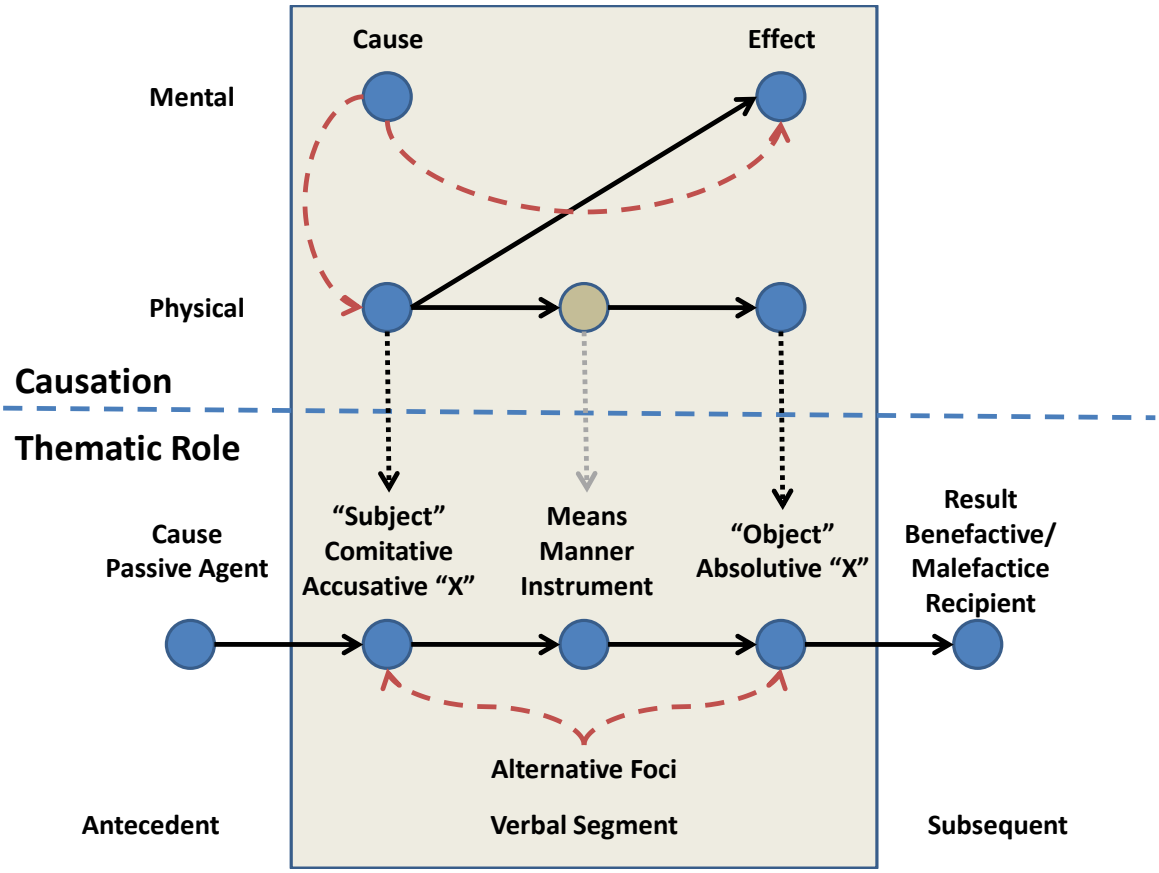


Figure 6-4. Causal chains and linguistic phenomena. Perception of causation requires an initiator and endpoint, with no presumption of telekinesis. For mental events, which

requires a "theory of mind," it is possible for a mental stimulus to affect another mental state. A mental initiator can also affect a physical endpoint, which may initiate other physical effects. Similarly, physical causes can have an influence on mental states. They sometimes operate by an intermediate physical mechanism, which may or may not be explicitly stated. They also affect a physical endpoint. These causal relationships are reflected directly in how human language describes the equivalent events. Thematic roles are explicitly encoded in various human languages. The causal role immediately precedes the event described by the verbal segment. Passive agents are also expressed external to the verb segment when there is a passive construction. "Subjects" are logical initiators of verbal actions. Their expression may vary by language. For example, in accusative languages such as English, or other European languages, "X" indicates the unmarked form (lowest number of morphemes, or carriers of meaning). The comitative role is performed by the entity that participates in this causal chain as the initiator. Means, manner, and instrumental roles are informative about how the verbal action is implemented. The "object" is the logical recipient of the verbal action. In ergative languages, this role receives the lowest number of morphemes, indicated by absolutive "X." Thus, across languages, either end of the verbal segment can receive the focus of the coding system. Result, benefactive (good result), malefactive (bad result), and recipient roles follow the span of the verbal segment. The observation of physical causation, which also corresponds to basic verbal descriptions, is common to the primate goal-oriented perception in mirror neurons.

REFERENCES

REFERENCES

Abercrombie M. (1946) Estimation of nuclear population from microtome sections. The Anatomical Record 94:239-247.

Albus JS (1971) A theory of cerebellar function. Mathematical Bioscience 10: 25-61.

Aldridge JW, Berridge KC (1998) Coding of serial order by neostriatal neurons: a "natural action" approach to movement sequence. Journal of Neuroscience, 18(7), 2777–2787.

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HBM, Zilles K (1999) Broca's region revisited: Cytoarchitecture and intersubject variability. The Journal of Comparative Neurology 421: 319-341.

Arbib MA (2004) How far is language beyond our grasp? A response to Hurford. In Kimbrough D, Griebel U (eds.) Evolution of Communication Systems: A Comparative Approach. Cambridge, MA: MIT Press, pp. 315-321.

Azevedo FAC, Carvalho LRB, Grinberg LT, Farfel JM, Ferretti REL, Renata E.P. Leite REP, Filho WJ, Lent R, Herculano-Houzel (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J Comp Neurol 513:532-41.

Bai J, Ramos RL, Ackman JB, Thomas AM, Lee RV, LoTurco JL (2003) RNAi reveals doublecortin is required for radial migration in rat neocortex. Nature Neuroscience 6: 1277-1283.

Bar I, Goffinet, AM (2000) Evolution of cortical lamination: the reelin/Dab1 pathway. In: Bock GR, and Cardew G (eds) Evolutionary developmental biology of the cerebral cortex, Novartis Foundation Symposium 228. Chichester, John Wiley & Sons, pp 114-125.

Barbas H (1997) In Sakata H, Mikami A, Fuster J, eds. The Association Cortex: Structure and Function. Amsterdam: Harwood Academic, pp. 99-116.

Barbas H, De Olmos J (1990) Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. Journal of Comparative Neurology 301: 1-23.

Barbas H, Ghashghaei H, Dombrowski SM, Rempel-Clower NL (1999) Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. Journal of Comparative Neurology 410: 343-367.

Barbas H, Henion TH, Dermon CR (1991) Diverse thalamic projections to the prefrontal cortex in the rhesus monkey. Journal of Comparative Neurology 313: 65-94.

Batchelor T (1809) Orthoëpical Analysis of the English Language.

Bates E, Wilson SM, Saygin AP, Dick F, Sereno M, Knight RT, Dronkers NF (2003) Voxel-based lesion-symptom mapping. Nature Neuroscience 6: 448-450.

Becksei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405: 590-593.

Becksei A, Séraphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks; cell differentiation by graded to binary response conversion. European Molecular Biology Organization 20: 2528-2535.

Belton E, Gadian DG, Vargha-Khadem F (2003) Evidence for specific motor programming deficit in developmental verbal dyspraxia but not in SLI. Soc Neurosci Abstr 196: 20.

Belton E, Salmond CH, Watkins KE, Vargha-Khadem F, Gadian DG (2003) Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. Human Brain Mapping 18: 194-200.

Berman RA, Colby CL, Genovese CR, Voyvodic JT, Luna B, Thulborn KR, Sweeney JA (1999) Cortical networks subserving pursuit and saccadic eye movements in humans: an FMRI study. Human Brain Mapping 8:209–25.

Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW (1996) Function of the left planum temporale in auditory and linguistic processing. Brain 119:1239-1247.

Binkofski F, Buccino G, Stephan K, Rizzolatti G, Seitz R and Freund H (1998) Human anterior intraparietal area subserves prehension: a combined lesion and functional MRI activation study. Neurology 50:1253-1259.

Blacker D, Byrnes ML, Mastaglia FL, Thickbroom GW (2006) Differential activation of frontal lobe areas by lexical and semantic language tasks: A functional magnetic resonance imaging study. Journal of Clinical Neuroscience 13:91–95.

Blake WJ, Kærn M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. Nature 422: 633-637.

Braver TS, Barch DM, Kelley WM, Buckner RL, Cohen NJ, Miezin FM, Snyder AZ, Ollinger JM, Akbudak E, Conturo TE, Petersen SE (2001) Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. Neuroimage, 14(1 Pt 1), 48–59.

Broad KD, Curley JP, Keverne EB (2009) Increased apoptosis during neonatal brain development underlies the adult behavioral deficits seen in mice lacking a functional paternally expressed gene 3 (Peg3). Developmental Neurobiology 69:314-25.

Broca P (1861a) Sur le principe des localisations cérébrale. Bulletin de la Société d'Anthropologie 2: 190-204.

Broca P (1861b) Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche. [Sur le siège de al faculté du langage]. Bulletin de la Société d'Anthropologie 2: 235-238.

Broca P (1861c) Remarques sur le siege de la faculté du langage articulé, suivies d'une observation d'aphémie perte de la parole. Bulletin de la Société Anatomique 36: 330-357.

Broca P (1861d) Nouvelle observation d'aphémie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolution frontales gauches. Bulletin de la Société Anatomique 36: 398-407.

Brown JW, Braver TS (2005) Learned predictions of error likelihood in the anterior cingulated cortex. Science 307: 1118-1121.

Buckner RL, Wheeler ME (2001) The cognitive neuroscience of remembering. Nature Review Neuroscience, 2(9), 624–634.

Burgess N (2008) Spatial cognition and the brain. Ann N Y Acad Sci 1124:77-97.

Butcher A (1981) Aspects of the speech pause: phonetic correlates and communicative functions. Kiel: University of Kiel.

Butler AB Hodos W (2005) Comparative Vertebrate Neuroanatomy: Evolution and Adaptation, 2d ed., Hoboken, NJ, John Wiley & Sons.

Butt AM, Berry M (2000) Oligodendrocytes and the control of myelination in vivo: New insights from the rat anterior medullary velum. J Neurosci Research 59: 477-488.

Buxhoeveden D, Switala A, Litaker M, Roy E, Casanova M (2001) Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex. Brain Behavioral Evolution 57: 349-358.

Buzsáki G (2006) Rhythms of the brain. Oxford, UK: Oxford University Press.

Cadoret G, Mackey S, Petrides M (2005) Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature 435:1235-1238.

Campbell A (1959) Old English Grammar. Oxford: Clarendon Press.

Cantalupo C, and Hopkins WD (2001) Asymmetric Broca's area in great apes. Nature 414: 505.

Cao Y, Vikingstad E, George K, Johnson A, Welch K (1999) Cortical language activation in stroke patients recovering from aphasia with functional MRI. Stroke 30:2331–40.

Caplan D, Alpert N, Waters G, Olivieri A (2000) Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Human Brain Mapping 9: 65-71.

Carlsson P, and Mahlapouu M (2002) Forkhead transcription factors: Key players in development and metabolism. Developmental Biology 250: 1-23.

Carroll SB (2003) Genetics and the making of Homo sapiens. Nature 422: 849-857.

Chen C, Kano M, Abeliovich A, Chen L, Bao S, Kim JJ, Hashimoto K, Thompson RF, Tonegawa S (1995) Impaired motor coordination correlates with persistent multiple climbing fiber innervation in PKCγ mutant mice. Cell 83: 1233-1242.

Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Naatanen R (1998) Development of language-specific phoneme representation in the infant brain. Nature Neuroscience 1: 351-353.

Chomsky N (1986) Knowledge of Language: Its Nature, Origin, and Use. New York: Praeger.

Chomsky N, Halle M (1968) The Sound Pattern of English. Cambridge, MA: MIT Press.

Coman I, Barbin G, Charles P, Zalc B, Lubetski C (2005) Axonal signals in central nervous system myelination, demyelination and remyelination. Journal of the Neurological Sciences 233: 67-71.

Compact Edition of the Oxford English Dictionary (1971) Oxford: Oxford UPress.

Conel JL (1939-1967) Postnatal development of the human cerebral cortex, vols 1-8. Cambridge, MA: Harvard University Press.

Cooper C (1685) Grammatica Linguae Anglicanae.

Cooper C (1687) The English Teacher.

Cooper DL (1999) Linguistic attractors: The cognitive dynamics of language acquisition and change. Amsterdam: John Benjamins.

Cooper DL, Gentle JE, Barreto E, Olds JL (2010) Synchronized changes to relative neuron populations in postnatal human neocortical development. J Cogn Neurodyn DOI 10.1007/s11571-010-9103-3.

Creuzfeldt O (1995) Cortex cerebri. Springer-Verlag

Croft W (1991) Syntactic categories and grammatical relations: The cognitive organization of information. Chicago: University of Chicago Press.

Croft W (2001) Radical construction grammar: Syntactic theory in typological perspective. Oxford: Oxford University Press.

Cuneod C, Bookheimer S, Hertz–Pannier L, Zeffiro T, Theodore W, Le Bihan D (1995) Functional MRI during word generation using conventional equipment: A potential tool for language localization in the clinical environment. Neurology 45:1821–7.

Damasio H, Grabowski T, Tranel D, Hichwa R,Damasio A (1996) A neural basis for lexical retrieval. Nature, 380(6574), 499–505.

Dean C, Leakey MG, Reid D, Friedemann S, Schwartz GT, Stringer C, Walker A (2001) Growth processes in teeth distinguish modern humans from Homo erectus and earlier hominins. Nature 414: 628-631.

Decety J, Grèzes J, Costes N, Perani D, Jeannerod M, Procyk E, Grassi F, Fazio F (1997) Brain activity during observation of actions. Influence of action content and subject's strategy. Brain 120: 1763-1777.

Dehaene S (2009) Reading in the brain: The science and evolution of a human invention. New York: Viking.

Dehaene-Lambertz G, Dehaene S (1994) Speed and cerebral correlates of syllable discrimination in infants. Nature 370: 292-295.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. Science 298: 2013-2015.

Dehaene-Lambertz G, Hertz-Pannier L, Dubois J, Mériaux S, Roche A, Sigman M, Dehaene S (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. Proceedings of the National Academy of Sciences 103: 14240-14245.

Dehay C, Kennedy H (2007) Cell-cycle control and cortical development. Nature Reviews Neuroscience 8:438-450.

Dhavan R, Tsai L-H, Tsai L-H (2001) A decade of CDK5. Nature Reviews Molecular Cell Biology 2:749-759.

DiCristo G, Chattopadhyaya B, Kuhlman SJ, Fu Y, Bélanger MC, Wu CZ, Rutishauser U, Maffei L, Huang ZJ (2007) Activity-dependent PSA expression regulates inhibitory maturation and onset of critical period plasticity. Nature Neuroscience 10:1569-77.

Donoghue MJ, Rakic P (1999) Molecular gradients and compartments in the embryonic primate cerebral cortex. Cerebral Cortex 9:586-600.

Dronkers N (1996) A new brain region for coordinating speech articulation. Nature 384: 159-161.

Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. Cognition 92:145-177.

Dum RP, Strick PL (1991) The origin of corticospinal projections from the premotor areas in the frontal lobe. The Journal of Neuroscience 11: 667-689.

Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. Nature Neuroscience 3: 1184-1191.

Economo Cv, Koskinas GN (1925) Die Cytoarchitectonik der Hirnrinde des erwachsenen Menschen. Berlin: Springer.

Eichenbaum H, Cohen NJ (2001) From conditioning to conscious recollection: memory systems of the brain. New York: Oxford University Press.

Ellison PAT, Semrud-Clikeman M (2007) Child neuropsychology: assessment and interventions for neurodevelopmental disorders. New York, Springer.

Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K (1996) Rethinking Innateness: A Connectionist Perspective on Development. Cambridge, MA: MIT Press.

Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297: 1183-1186.

Eldredge N, Gould SJ (1972) Punctuated equilibria: an alternative to phyletic gradualism. In Schopf, T. J. M., ed., Models in Paleobiology. San Francisco: Freeman Cooper. pp. 82-115.

Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Favor J, Fuchs H, Gailus-Durner V, Hans W, Hölzlwimmer G, Javaheri A, Kalaydjiev S, Kallnik M, Kling E,7, Kunder S, Moßbrugger I, Naton B, Racz I, Rathkolb B, Rozman J, Schrewe A, Busch DH, Graw J, Ivandic B, Klingenspor M, Klopstock T, Ollert M, Quintanilla-Martinez L, Schulz H, Wolf E, Wurst W, Zimmer A, Fisher SE, Morgenstern R, Arendt T, Hrabé de Angelis M, Fischer J, Schwarz J, Pääbo S (2009) A humanized version of foxp2 affects cortico-basal ganglia circuits in mice. Cell 137: 961–971.

Enard W, Przeworski M, Fisher SE, Lai CSL, Wiebe V, Kitano T, Monaco AP, Pääbo S (2002) Molecular evolution of FOXP2, a gene involved in speech and language. Nature 418: 869-872.

Ewens WJ, Grant GR (2001) Statistical methods in bioinformatics. New York: Springer.

Feldman DE, Nicoll RA, Malenka RC (1999) Synaptic plasticity at thalamocortical synapses in developing rat somatosensory cortex: LTP, LTD, and silent synapses. J Neurobiol 41:92-101.

Ferland RJ, Cherry TJ, Preware PO, Morrisey EE, Walsh CA (2003) Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. Journal of comparative Neurology 460: 266-279.

Ferrell JE, Jr., Machleder EM (1998) The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. Science 280: 895-898.

Fiez JA (1997) Phonology, semantics, and the role of the left inferior prefrontal cortex. Human Brain Mapping, 5, 79–83.

Finger S (2004) Paul Broca (1824-1880). J Neurol (2004) 251 : 769–770.

Finlay BL, Darlington RB, Nicastro N (2001) Developmental structure in brain evolution. Behavioral and Brain Sciences 24:263–308.

Fischer O, van Kemenade A, Koopman W, van der Wurff W (2000) The Syntax of Early English. Cambridge, UK: Cambridge University Press.

Flechsig P (1920) Anatomie des menschlichen Gehirns und Rückenmarks auf myelogenetischer Grundlage. Leipzig: G. Thieme.

Fowler CA, Brown JM, Sabadini L, Weihing J (2003) Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. Journal of Memory & Language, 49: 396-413.

Frankland PW, Bomtempi B (2005) The organization of recent and remote memories. Nature Reviews Neuroscience 6: 119-130.

Friedman M (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association 32: 675-701.

Fritz J, Mishkin M, Saunders RC (2005) In search of an auditory engram. Proc Natl Acad Sci U S A. 2005 Jun 28;102(26):9359-64.

Galantucci B, Fowler CA, Turvey MT (2006) The motor theory of speech perception reviewed. Psychonomic Bulletin & Review 13: 361-377.

Gallego-Diaz V, Schoenwolf GC, Alvarez IS (2002) The effects of BMPs on early chick embryos suggest a conserved signaling mechanism for epithelial and neural induction among vertebrates. Brain Research Bulletin 57:289-291.

Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. Brain 119: 593-609.

Gannon PJ, Holloway RL, Broadfield DC, Braun AR (1998) Asymmetry of chimpanzee planum temporale: Humanlike patterns of Wernicke's brain language area homolog. Science 279; 220-222.

Gazzaniga MS, Ivry RB, Mangun GR (2002) Cognitive neuroscience: The biology of the mind. New York: W. W. Norton.

Gdalyahu A, Ghosh I, Levy T, Sapir T, Sapoznik S, Fishler Y, Azoulai D, Reiner O (2004) DCS, a new mediator of the JNK pathway. European Molecular Biology Organization 23: 823-832.

Gentle JE (2002) Elements of computational statistics. New York, Springer.

Geschwind N (1967) The varieties of naming errors. Cortex 3: 97-112.

Godsil C, Royle G (2004) Algebraic Graph Theory. Berlin: Springer.

Gohlke JM, Griffith WC, Faustman EM (2007) Computational models of neocortical neuronogenesis and programmed cell death in the developing mouse, monkey, and human. Cerebral Cortex 17:2433-2442.

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends in Neuroscience, 15(1), 20–25.

Gould SJ (1977) Ontogeny and phylogeny. Cambridge, MA: Belknap.

Grafton ST, Arbib MA, Fadiga L, Rizzolatti G (1996) Localization of grasp representations in humans by positron emission tomography. 2. Observations compared with imagination. Experimental Brain Research 1: 103-111.

Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. Physica D 9:189-208.

Grèzes J, Costes N, Decety J (1998) Top-down effect of strategy on the perception of human biological motion: a PET investigation. Cognitive Neuropsychology 15: 553-582.

Habeck C, Rakitin BC, Moeller J, Scarmeas N, Zarahn E, Brown T, Stern Y (2005) An event-related fMRI study of the neural networks underlying the encoding, maintenance, and retrieval phase in a delayed-match-to-sample task. Cognitive Brain Research 23:207-220.

Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C (2004) FoxP2 expression in avian vocal learners and non-learners. The Journal of Neuroscience 24: 3164-3175.

Hanashima C, Li SC, Shen L, Lai E, Fishell G (2004) Foxg1 suppresses early cortical cell fate. Science 303:56-59.

Hawkins DL (1989) Using U statistics to derive the asymptotic distribution of Fisher's Z statistic. The American Statistician 43, 235-237.

Hesse MD, Sparing R, Fink GR (2009) End or means—the ‘‘what’’ and ‘‘how’’ of observed intentional actions. Journal of Cognitive Neuroscience 21:776–790.

Hodges JR, Patterson K (1997) Semantic memory disorders. Trends in Cognitive Sciences, 1: 68–72.

Holland PWH, Takahashi T (2005) The evolution of homeobox genes: implications for the study of brain development. Brain Research Bulletin 66:484-490.

Holloway RL (1980) Indonesian “Solo” Ngandong endocranial reconstructions: Some preliminary observations and comparisons with Neandertal and Homo erectus groups. American Journal of Physical Anthropology 53: 285-295.

Hopkins WD, Leavens DA (1998) The whole-hand point: The structure and function of pointing from a comparative perspective. Journal of Comparative Psychology 112: 95-99.

Hopkins WD, Marino L, Riling JK, MacGregor L (1998) Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging MRI. NeuroReport 9: 2913-2918.

Horwitz B, Rumsey JM, Donohue BC (1998) Functional connectivity of the angular gyrus in normal reading and dyslexia. Proc Natl Acad Sci U S A) 15:8939-44.

Huang EJ, Reichardt LF (2001) Neurotrophins: roles in neuronal development and function. Annual Review of Neuroscience 24:677-736.

Hurford, JR (2004) Language beyond our grasp: What mirror neurons can, and cannot, do for the evolution of language. In Kimbrough D, Griebel U (eds.) Evolution of Communication Systems: A Comparative Approach. Cambridge, MA: MIT Press, pp. 297-313.

Iacoboni M (2005) Neural mechanisms of imitation. Curr Opin Neurobiol 15:632-7.

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. Science 286:2526-2528.

Ito M (1989) Long-term depression. Annual Review of Neuroscience 12: 85-102.

Jarvis ED, Güntürkün O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, Striedter G, Wild JM, Ball GF, Dugas-Ford J, Durand SE, Hough GE, Husband S, Kuikova L, Lee DW, Mello CV, Powers A, Siang C, Smulders TV, Wada K, White SA, Yamamoto K, Yu J, Reiner A, Butler AB (2005) Avian brains and a new understanding of vertebrate brain evolution. Nature Reviews Neuroscience 6: 151-159.

Jespersen O (1909) A Modern English Grammar on Historical Principles, Part 1: Sounds and Spellings. Carl Winter: Heidelberg.

Johnson MH, Morton J (1991) Biology and Cognitive development: The Case of Face Recognition. Oxford: Blackwell.

Just M, Carpenter P, Keller T, Eddy W, Thulborn K (1996) Brain activation modulated by sentence comprehension. Science 274: 114-116.

Kano M, Hashimoto K, Chen C, Abeliovich A, Aiba A, Kurihara H, Watanabe M, Inoue Y, Tonegawa S (1995) Impaired synapse elimination during cerebellar development in PKCγ mutant mice. Cell 83: 1223-1231.

Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature 433: 638-643.

Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc Natl Acad Sci U S A. 87: 2264-8.

Katz LC, Shatz CJ (1996) Synaptic activity and the construction of cortical circuits. Science 274:1133-1138.

Khazipov R, Sirota A, Leinekugel X, Halmes GL, Ben-Ari Y, Buzsáki G (2004) Early motor activity drives spindle bursts in the developing somatosensory cortex. Nature 432:758-761.

Kim KHS, Relkin NR, Lee KM, Hirsch J (1997) Distinct cortical areas associated with native and second languages. Nature 388: 171-174.

Koch C (2004) The quest for consciousness: A neurobiological approach. Englewood, CO: Roberts and Company.

Koelsch S, Gunter TC, Cramon DYv, Zysset S, Lohmann G, Friederici AD (2002) Back speaks: A cortical "language-network" serves in the processing of language. Neuroimage 17: 956-966.

Koelsch S, Gunter T, Friederici AD, Schröger E (2000) Brain indices of music processing: "nonmusicians" are musical. Journal of Cognitive Neuroscience 12: 520-541.

Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: Action representations in mirror neurons. Science 297: 846-848.

Kolb B, Whishaw IW (1990) Fundamentals of human neuropsychology, 3d. ed., New York: W. H. Freeman.

Kolpak A, Zhang J, Bao ZZ (2005) Sonic hedgehog has a dual effect on the growth of retinal ganglion axons depending on its concentration. J Neurosci 25:3432–41.

Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F, Peng S, Preuss TM, Wohlschlegel JA, Geschwind DH (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature 462: 213-218.

Koski L, Paus T (2000) Functional connectivity of the anterior cingulated cortex within the human frontal lobe: a brain-mapping meta-analysis. Experimental Brain Research 133: 55-65.

Kozhevnikov VA, Chistovich LA, (1965) Speech, articulation and perception. Washington, DC: U.S. Department of Commerce.

Kuhl PK (1991) Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. Percept. Psychophys. 50: 93–107.

Kuhl PK (2004) Early language acquisition: Cracking the speech code. Nature Reviews Neuroscience 5: 831-843.

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255: 606–608.

Kullback S, Leibler RA (1951) On information and sufficiency, Annals of Mathematical Statistics 22: 79-86.

Kunishio K, Haber S (1994) Primate cingulostriatal projection: Limbic striatal versus sensorimotor striatal input. Journal of Comparative Neurology 350: 337-356.

Labov W (1966) Social Stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics.

Labov W (1994) Principles of Linguistic Change: Internal Factors. Oxford: Blackwell.

Labov W (2001) Principles of Linguistic Change: Social Factors. Oxford: Blackwell.

Lai CSL, Gerrelli D, Monaco AP, Fisher SE, Copp AJ (2003) FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. Brain 126: 2455-2462.

Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413: 519-523.

Langacker RW (1998) Conceptualization, symbolization, and grammar. In Tomasello M, editor. The new psychology of language: Cognitive and functional approaches to language structure. Mahwah, NJ: Lawrence Erlbaum: 1-39.

Le Bihan D, Jezzard P, Turner R, Cuneod C, Zefferio T (1993) Analysis of functional brain MR images with Z maps. J Magn Reson Imaging 3:141.

Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Pitcher K, Crinion J, Mechelli A, Frackowiak RSJ, Green DW, Price CJ (2007) Anatomical traces of vocabulary acquisition in the adolescent brain. J Neurosci 27:1184-1189.

Letinic K, Zoncu R, Rakic P (2002) Origin of GABAergic neurons in the human neocortex. Nature 417:645-649.

Li S, Weidenfeld J, Morrisey EE (2004) Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Molecular and Cellular Biology 24: 809-822.

Liberman AM (1957) Some results of research on speech perception. Journal of the Acoustical Society of America, 29, 117-123.

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of speech code. Psychological Review, 74: 431-461.

Liberman AM, Delattre P, Cooper FS (1952) The role of selected stimulus-variables in the perception of the unvoiced stop consonants. American Journal of Psychology, 65, 497-516.

Lichtheim L (1885) Űber Aphasie. Deutsches Archiv für klinische Medicin 36: 204-268.

Lieberman E, Michel JB, Jackson J, Tang T, Nowak MA (2007) Quantifying the evolutionary dynamics of language. Nature 449: 713-716.

Liégeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F (2003) Language fMRI abnormalities associated with FOXP2 gene mutation. Nature Neuroscience 6: 1230-1237.

Linsker R (1993) Deriving receptive fields using an optimal encoding criterion. In Hanson JS, Cowan JD, Giles CL, editors. Neural Information Processing Systems 5. San Mateo, CA: Morgan Kaufmann: 953-960.

Lisman JE, Fellous J-M, Wang X-J (1998) A role for NMDA-receptor channesl in working memory. Nature Neuroscience 1: 273-275.

Lisman JE, Zhabotinsky AM (2001) A model of synaptic memory: A CaMKII/PP1 switch that potentiates transmission by organizing an AMPA receptor anchoring assembly. Neuron 31: 191-201.

Lisman J, Schulman H, Cline H (2002) The molecular basis of CaMKII function in synaptic and behavioural memory. Nature Reviews Neuroscience 3: 176-190.

Llinás R (2001) I of the vortex: From neurons to self. Cambridge, MA; MIT Press.

Loewe D (2008) Hillary's new inevitability. http://commentisfree.guardian.co.uk Manchester Guardian on-line edition, 23 April 2008.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. Nature 412: 150-157.

Lowenstein Y, Mahon S, Chadderton P, Kitamura K, Sompolinsky H, Yarom Y, Häusser M (2005) Bistability of cerebellar Purkinje cells modulated by sensory stimulation. Nature Neuroscience 8: 202-211.

Lu M, Li S, Yang H, Morrisey EE (2002) Foxp4: a novel member of the Foxp subfamily of winged-helix genes co-experssed with Foxp1 and Foxp2 in pulmonary and gut tissues. Gene Expression Patterns 2: 223-228.

Lukaszewicz A, Savatier P, Cortay V, Giroud P, Huissoud C, Berland M, Kennedy H, Dehay C (2005) G1 phase regulation, area-specific cell cycle control, and cytoarchitectonics in the primate cortex. Neuron 47:353-64.

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432-38.

Maess B, Koelsch S, Gunter TC, Friederici AD (2001) Musical syntax is processed in Broca's area: an MEG study. Nature Neuroscience 4: 540-545.

Mandler JM (1999) Preverbal representation and language. In: Bloom P, Peterson MA, Nadel L, Garrett MF (eds). Language and space. Cambridge, MA, MIT Press, pp. 365-384.

Marcotte AC, Morere DA (1990) Speech lateralization in deaf populations: evidence for a developmental critical period. Brain and Language 39:134-52.

Marin-Padilla M (1999) The development of the human cerebral cortex. A cytoarchitectonic theory. Reviews in Neurology 29: 208-216.

Marr D (1969) A theory of cerebellar cortex. Journal of Physiology 202: 437-470.

Marr D (1980) Vision. San Francisco: W. H. Freeman.

Martin A, Ungerleider LG,Haxby JV (2000) Category specificity and the brain: the sensory/motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 1023–1036). Cambridge, MA: MIT Press.

Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, Lueders HO (2004) Functional connectivity in the human language system: a cortico-cortical evoked potential study. Brain 127:2316-2330.

Maviel T, Durkin TP, Menzaghi F, Bontempi B (2004) Sites of neocortical reorganization critical for remote spatial memory. Science 305: 96-99.

Mayr E (1954) Change of genetic environment and evolution. In Huxley J, Hardy AC, Ford EB. Evolution as a Process. London: Allen and Unwin: 157-180.

McArthy G, Blamire A, Rothman D, Gruetter R, Schulman R (1993) Echoplanar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. Proc Nat Acad Sci USA 90:4592–6.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264: 746-748.

Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertoncini J, Amiel-Tison C (1988) A precursor of language acquisition in young infants. Cognition 29: 143-178.

Michel GF, Tyler AN (2005) Critical period: A history of the transition from questions of when, to what, to how. Developmental Psychobiology 46:156-162.

Middleton FA, Strick PL (2000) Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Research Reviews 31: 236-250.

Moggi-Cecchi J (2001) Human evolution: Questions of growth. Nature 414: 595-597.

Molnár Z, Metin C, Stoykova A, Tarabykin V, Price DJ, Francis F, Meyer G, Dehay C, Kennedy H (2006) Comparative aspects of cerebral cortical development. Eur J Neurosci 23:921-34.

Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD (2007) Neuronal subtype specification in the cerebral cortex. Nature Reviews Neuroscience 8:427-437.

Montaron M-F, Buser P (1988) Relationships between nucleus medialis dorsalis, pericruciate cortex, ventral tegmental area and nucleus accumbens in cat: an electrophysiological study. Experimental Brain Research 69: 559-566.

Monuki ES, Porter FD, Walsh CA (2001) Patterning of the dorsal telencephalon and cerebral cortex by a roof plate–Lhx2 pathway. Neuron 32:591-604.

Morais J, Bertelson P, Cary L, Alegria J (1986) Literacy training and speech segmentation. Cognition 24: 45-64.

Morecraft RJ, Van Hoesen GW (1992) Cingulate input to the primary and supplementary motor cortices in the rhesus monkey: Evidence for somatotopy in areas 24c and 23c. Journal of Comparative Neurology 322: 471-489.

Mower GD, Chen L (2003) Laminar distribution of NMDA receptor subunit (NR1, NR2A, NR2B) expression during the critical period in cat visual cortex. Molecular Brain Research 119:19-27.

Murakami S, Okada Y (2006) Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. J Physiol 575: 925-936.

Murata A., Fadiga L, Fogassi L, Gallese V, Raos V, Rizzolatti G (1997) Object representation in the ventral premotor cortex area F5 of the monkey. Journal of Neurophysiology 4: 2226-30.

Muzio L, DiBenedetto B, Stoykova A, Boncinelli E, Gruss P, Mallamaci A (2002) Conversion of cerebral cortex into basal ganglia in Emx2-/-Pax6Sey/Sey double-mutant mice. Nature Neuroscience 5:737-745.

Muzio L, Mallamaci A (2005) Foxg1 confines Cajal–Retzius neuronogenesis and hippocampal morphogenesis to the dorsomedial pallium. J. Neurosci. 25:4435-4441.

Nadarajah B, Parnavelas JG (2002) Modes of neuronal migration in the developing cerebral cortex. Nat Rev Neurosci. 3:423-32.

Nelissen K, Luppino G, Vanduffel W, Rizzolatti G, Orban GA (2005) Observing others: Multiple action representation in the frontal lobe. Science 310: 332-336.

Nishitani N, Uutela K, Shibasaki H, and Hari R (1999) Cortical visuomotor integration during eye pursuit and eye-finger pursuit. Journal of Neuroscience 19:2647-2657.

Noden DM (1991) Vertebrate craniofacial development: the relation between ontogenetic process and morphological outcome. Brain, Behavior and Evolution 38:192-225.

Nomura T, Takahashi M, Hara Y, Osumi N (2008) Patterns of neurogenesis and amplitude of reelin expression are essential for making a mammalian-type cortex. Plos One 1:e1454.

Ozbudak EM, Thattai M, Kurser I, Grossman AD, Oudenaarden Av (2002) Regulation of noise in the expression of a single gene. Nature Genetics 31: 69-73.

Parent A, Hazrati L-N (1995) Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res Rev 20: 91-127.

Patel AD (2003) Language, music, syntax and the brain. Nature Neuroscience 6: 674-681.

Patel RS, Bowman FD, Rilling JK (2006) Determining hierarchical functional networks from auditory stimuli fMRI. Human Brain Mapping 27:462-470.

Patel A, Gibson E, Ratner J, Besson M, Holcomb PJ (1998) Processing syntactic relations in language and music: An event-related potential study. Journal of Cognitive Neuroscience 10: 717-733.

Paus T, Koski L, Zografos C, Westbury C (1998) Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulated cortex: A review of 107 PET activation studies. Neuroreport 9: 37-47.

Paus T (2001) Primate anterior cingulated cortex: where motor control, drive and cognition interface. Nature Reviews Neuroscience 2: 417-424.

Pedraza JM, Oudenaarden Av (2005) Noise propagation in gene networks. Science 307: 1965-1969.

Peirce CS (1868) On a new list of categories. Proceedings of the American Academy of Arts and Sciences 7: 287-298.

Pena M, Maki A, Kovacic D, Dehaene-Lambertz g, Koizumi H, Bouquet F, Mehler J (2003) Sounds and silence: An optical topography study of language recognition at birth. Proceedings of the National Academy of Sciences 100: 11702-11705.

Pentonnen M, Buzsáki G (2003) Natural logarithmic relationship between brain oscillators. Thalamus Relat Syst 48:1-8.

Petrides M, Cadoret G, Mackey S (2005) Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature 435 : 1235-1238.

Piaget J (1952) The origins of intelligence in children. New York: International Universities Press.

Pilati N, Barker M, Panteleimonitis S, Donga R, Hamann M (2008) A rapid method combining Golgi and Nissl staining to study neuronal morphology and cytoarchitecture. J Histochem Cytochem 56:539-550.

Pinker S (1994) The Language Instinct: How the Mind Creates Language. New York: William Morrow.

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD (1999) Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage, 10(1), 15–35.

Polleux F, Ince-Dunn G, Ghosh A (2007) Transcriptional regulation of vertebrate axon guidance and synapse formation. Nature Reviews Neuroscience 8:331-340.

Ponce de León MS, Zollikofer CPE (2001) Neandertahl cranial ontogeny and its implications for late hominid diversity. Nature 412: 534-538.

Pouget A, Latham P (2002) Digitized neural networks: Long-term stability from forgetful neurons. Nature Neuroscience 5: 775-782.

Porter RJ Jr, Castellanos FX (1980) Speech-production measures of speech perception: Rapid shadowing of VCV syllables. Journal of the Acoustical Society of America 67: 1349-1356.

Porter RJ Jr, Lubker JF (1980) Rapid reproduction of vowel–vowel sequences: Evidence for a fast and direct acoustic–motoric linkage in speech. Journal of Speech & Hearing Research 23: 593-602.

Preuss TM, Goldman-Rakic PS (1991) Myelo- and cytoarchitecture of the granular frontal cortex and surrounding regions in the streshirhine primate Galago and the anthropoid primate Macaca. Journal of Computational Neurology 310: 429-474.

Preuss TM, Cáceres M, Oldham MC, Geschwind DH (2004) Human brain evolution: insights from mircroarrays. Nature Reviews Genetics 5: 850-860.

Pulvermüller F (2002) The neuroscience of language: On brain circuits of words and serial order. Cambridge: Cambridge University Press.

Pyles T (1971) The Origins and Development of the English Language, 2d ed. New York: Harcourt Brace Jovanovich.

Quirk R, Wrenn CL (1994) An Old English Grammar. DeKalb, IL: NIU Press.

Rakic P (1988) Specification of cerebral cortical areas. Science 241:170-176.

Rakic P (2001) Neurobiology: neurocreationism—making new cortical maps. Science 294:1011-1012.

Raser JM, O'Shea EK (2004) Control of stochasticity in eukaryotic gene expression. Science 304: 1811-1814.

Renner J, Stafford D, Lawson A, McKinnon J, Friot E, Kellogg D (1976) Research, teaching, .nd learning with the Piaget model. Norman, OK, U Oklahoma Press.

Rice SH (2001) In Mingh-Purvis, N, and McNamara, KJ. eds., Human evolution through developmental change. Baltimore: Johns Hopkins: 154-170.

Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, Paulesu E, Perani D, Fazio F (1996) Localization of grasp representations in humans by PET: 1. Observation versus execution. Experimental Brain Research 2: 246-252.

Rizzolatti G, Fogassi L, Gallese V (2000) In Gazzaniga, MS. Ed., The Cognitive Neurosciences, 2d ed. Cambridge, MA: MIT Press: 539-552.

Rizzolatti G, Camarda R, Fogassi L, Gentilluci M, Luppino G, Matelli M (1988) Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. Experimental Brain Research 71: 491-507.

Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience 2: 661-670.

Rosch E, (1975) Cognitive reference points. Cognit. Psychol. 7: 532–547.

Rose GJ, Goller F, Gritton HJ, Plamondon SL, Baugh AT, Cooper BG (2004) Species-typical songs in white-crowned sparrows tutored with only phrase pairs. Nature 432: 753-758.

Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. Science 307: 1962-1965.

Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. Journal of Molecular Biology 323: 785-793.

Ross DS, Bever TG (2004) The time course for language acquisition in biologically distinct populations: evidence from deaf individuals. Brain and Language 89:115-21.

Rost B (1996) Phd: Predicting one-dimensional protein structure by profile-based neural networks. Methods in Enzymology 266: 525-539.

Ryan MJ, Wilczynski W (1988) Coevolution of sender and receiver: Effect on local mate preference in cricket frogs. Science 240: 1786-1788.

Sahin NT, Pinker S, Halgren E (2006) Abstract grammatical processing of nouns and verbs in Broca's Area: evidence from fMRI. Cortex 42:540-562.

Sakata H, Taira M, Mine S and Murata A (1992) Hand-movement related neurons of the posterior parietal cortex of the monkey: their role in visual guidance of hand movements. In Caminiti R, Johnson PB, Burnod Y (eds.), Control of arm movement in space: 185-198. Berlin: Springer-Verlag.

Saleem RA, Banerjee-Basu S, Berry FB, Baxevanis AD, Walter MA (2003) Structural and functional analyses of disease-causing missense mutations in the forkhead domain of FOXC1. Human Molecular Genetics 12: 2993-3005.

Samengo I and Treves A (2001) Representational capacity of a set of independent neurons. Phys Rev E Stat Nonlin Soft Matter Phys. 63:011910.

Sampaio RC and Truwit CL (2001) Myelination in the developing brain. In Nelson CA and Luciana M, eds. Handbook of developmental cognitive neuroscience. Cambridge, MA: MIT Press, pp. 35-44.

Samuels ML (1972) Linguistic Evolution: with special reference to English. Cambridge, UK: Cambridge University Press.

Sanchez MM, Hearn EF, Do D, Rilling JK, Herndon JG (1998) Differential rearing affects corpus callosum size and cognitive function of rhesus monkeys. Brain Research 812: 38-49.

Saussure, F de (1916) Cours de lingistique générale. Paris: Payot.

Schmithorst VJ, Holland SK, Plante E (2007) Object identification and lexical/semantic access in children: a functional magnetic resonance imaging study of word-picture matching. Human Brain Mapping 28:1060-1074.

Shamir M, Sompolinsky H (2006) Implications of neuronal diversity on population coding. Neural Comput 18:1951-86.

Shankle WR, Romney AK, Landing BH, Hara J (1998) Developmental patterns in the cytoarchitecture of the human cerebral cortex from birth to 6 years examined by correspondence analysis. Proc Nat Acad Sci USA 95:4023-4028.

Shankle RS, Landing BH, Rafii MS, Hara J, Fallon JH, Romney AK, Boyd JP (2000) CYBERCHILD: A database of the microscopic development of the postnatal human cerebral cortex from birth to 72 months. Neurocomputing 32-33:1109-1114.

Shankle RS, Hara J, Fallon JH, Landing BH (2002) How the brain develops and how it functions: application of neuroanatomical data of the developing human cerebral

cortex to computational models. In: Ascoli AA (ed) Computational neuroanatomy: principals and methodsTotowa, NJ, Humana Press, pp. 402-436.

Shaw KL (2000) Interspecific genetics of mate recognition: Inheritance of female acoustic preference in Hawaiian crickets. Evolution 54: 1303-1312.

Shu W, Ho JY, Jiang Y, Zhang M, Weisz D, Elder GA, Schmeidler J, De Gasperi R, Sosa MAG, Rabidou D, Santucci AC, Perl D, Morrisey E, Buxbaum JD (2005) Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. Proc Nat Acad Sci USA 102:9643-9648.

Shu W, Yan H, Zhang L, Lu MM, Morrisey EE (2001) Characterization of a new subfamily of winged-helix/forkhead Fox genes that are expressed in the lung and act as transcriptional repressors. Journal of Biological Chemistry 276: 27488-27497.

Sidman RL, Rakic P (1973) Neuronal migration, with special reference to developing human brain: a review. Brain Research 62: 1-35.

Sinha C (2004) The evolution of language: From signals to symbols to system. . In Kimbrough D, Griebel U (eds.) Evolution of Communication Systems: A Comparative Approach. Cambridge, MA: MIT Press, pp. 218-235.

Sirevaag ER, Greenough WT (1987) Differential rearing effects on rat visual cortex synapses, III. Neuronal and glial nuclei, boutons, dendrites, and capillaries. Brain Research 424: 320-332.

Smith EE, Jonides J (1999) Storage and executive processes in the frontal lobes. Science, 283: 1657–1661.

Solecki DJ, Model L, Gaetz J, Kapoor TM, Hatten ME (2004) Par6α signaling controls glial-guided neuronal migration. Nature Neuroscience 7: 1195-1203.

Sompolinsky H, Yoon H, Kang K, Shamir M (2001) Population coding in neuronal systems with correlated noise. Phys Rev E Stat Nonlin Soft Matter Phys 64:051904.

Squire LR, Clark RE, Knowlton BJ (2001) Retrograde amnesia. Hippocampus, 11(1), 50–55.

Squire LR, Knowlton BJ (2000) The medial temporal lobe, the hippocampus, and the memory systems of the brain. In Gazzaniga MS (Ed.), The new cognitive neurosciences (pp. 765–780). Cambridge, MA: MIT Press

Stiles J (2008) The fundamentals of brain development: integrating nature and nurture. Cambridge, MA, Harvard U Press.

Strand AD, Aragaki AK, Baquet ZC, Hodges A, Cunningham P, Holmans P, Jones KR, Jones L, Kooperberg C, Olson JM (2007) Conservation of Regional Gene Expression in Mouse and Human Brain. PLoS Genet 3(4): e59 doi:10.1371/journal.pgen.0030059.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America, 26: 212-215.

Sun Z, Rao X, Peng L, Xu D (1997) Prediction of protein secondary structures based on the artificial neural network method. Protein Engineering 10: 763-769.

Suzuki WA, Eichenbaum H (2000) The neurophysiology of memory. Annals of the New York Academy of Sciences, 911, 175–191.

Suzuki WA, Amaral DG (1994) Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferants. Journal of Comparative Neurology, 350(4), 497–533.

Szeligo F, Leblond CP (1977) Response of the three main types of glial cells of cortex and corpus callosum in rats handled during suckling or exposed to enriched, control and impoverished environments following weaning. J Comp Neurol 172: 247-263.

Takahashi K, Liu FC, Hirokawa K, Takahashi H (2003) Expression of Foxp2, a gene involved in speech and language, in the developing and adult striatum. Journal of Neuroscience Research 73: 61-72.

Takahashi T, Goto T, Miyame S, Nowakowski RS, Caviness VS, Jr. (1999) Sequence of neuron origin and neocortical laminar fate: relation to cell cycle of origin in the developing murine cerebral wall. J. Neuroscience 19:10357-10371.

Taylor JR (2002) Cognitive Grammar. Oxford: Oxford U Press.

Teramitsu I, Kudo LC, London SE, Geschwind DH, White SA (2004) Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. The Journal of Neuroscience 24: 3152-3163.

Thompson PM, Cannon TD, Narr KL, Erp Tv, Poutanen V-P, Huttunen M, Lönnqvist J, Standertskjöld-Nordenstam C-G, Karpio J, Khaledy M, Dail R, Zoumalan CI, Toga TW (2001) Genetic influences on brain structure. Nature Neuroscience 4: 1253-1258.

Tomasello M (1999) The cultural origins of human cognition. Cambridge, MA: Harvard University Press.

Tomasello M (2003) Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard.

Turing AM (1936) On computable numbers, with an application to the Entscheidungsproblem. Proc. London Maths. Soc. Ser 2 42: 230-265.

Turkeltaub Pe, Gareau L, Flowers DL, Zeffiro TA, Eden GF (2003) Development of neural mechanisms for reading. Nature Neuroscience 6: 767-773.

Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In Ingle DJ, Goodale MA, Mansfield RJW(Eds.), Analysis of visual behavior (pp. 549–587). Cambridge, MA: MIT Press.

Van der Wouden T (1988) in Magay T, Zigány J, eds. Papers from the 3rd International EURALEX Congress 363–373. Akadémiai Kiadó: Budapest.

Vargha-Khadem F, Gadian DG, Copp A, Mishkin M (2005) FOXP2 and the neuroanatomy of speech and language. Nature Reviews Neuroscience 6: 131-138.

Vasconcelos DB, Viana RL, Lopes SR, Batista AM, de S. Pinto SE (2004) Spatial correlations and synchronization in coupled map lattices with long-range interactions. Physica A 343:201-218.

Verney C, Takahashi T, Bhide PG, Nowakowski RS, Caviness VS, Jr. (2000) Independent controls for neocortical neuron production and histogenetic cell death. Developmental Neuroscience 22:125-38.

Vogt BA, Barbas H (1988) In Newman, JD, ed. The physiological control of mammalian vocalization. New York: Plenum: 203-225.

von Economo, Baron C, Koskinas GN (1925) The Cytoarchitectonics of the adult human cortex. Vienna, Springer.

Wallis J (1653 – 1699) Grammatica Linguae Anglicanae.

Wang B, Lin D, Li C, Tucker P (2003) Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. Journal of Biological Chemistry 278: 24259-24268.

Wernicke K (1874) Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis. Breslau: M. Crohn und Weigert.

Wilks SS (1962) Mathematical Statistics. New York: Wiley.

Woodward TS, Cairo TA, Ruff CC, Takane Y, Hunter MA, Ngan ETC (2006) Functional connectivity reveals load dependent neural systems underlying encoding and maintenance in verbal working memory. Neuroscience 139:317-325.

Yakovlev PI, Lecours AR (1967) The myelogenetic cycles of regional maturation of the brain. In Minkowski A, ed. Regional development of the brain in early life. Philadelphia: Davis, pp. 3-69.

Yamamoto N, Higashi S, Toyama K (1997) Stop and branch behaviors of geniculocortical axons: a time-lapse study in organotypic cocultures. J Neurosci 17:3653-63.

Yamamoto N, Inui K, Matsuyama Y, Harada A, Hanamura K, Murakami F, Ruthazer ES, Rutishauser U, Seki T (2000) Inhibitory mechanism by polysialic acid for lamina-specific branch formation of thalamocortical axons. J Neurosci 20:9145-51.

Young AB, Penney JB (1993) Biochemical and functional organization of the basal ganglia. In Jankovic J, Tolosa E (Eds.), Parkinson's disease and movement disorders (2nd ed.) (pp. 1–11). Baltimore, MD: Williams and Wilkins.

Zhang J, Webb DM, Podlaha O (2002) Accelerated protein evolution and origins of human-specific features: Foxp2 as an example. Genetics 162: 1825-1835.

CURRICULUM VITAE

David L. Cooper retired from the US Army after 22 years of service as a Lieutenant Colonel who specialized in Military Intelligence and served as an Arabic and French-speaking Foreign Area Officer in the Middle East and North Africa. He is currently employed as a project scientist in the National Geospatial-Intelligence Agency.

Previous degrees
A. B. in German and Linguistics from Princeton University, 1975
M. A. in International Relations from the University of Southern California, 1978

Publications:

Cooper DL (1999) Linguistic attractors: The cognitive dynamics of language acquisition and change. Amsterdam: John Benjamins.

Cooper DL (1996) Broca's arrow: evolution, prediction, and language in the brain. Anat Rec B New Anat 289: 9-24.

Cooper DL, Gentle JE, Barreto E, Olds JL (2010) Synchronized changes to relative neuron populations in postnatal human neocortical development. J Cogn Neurodyn DOI 10.1007/s11571-010-9103-3.