

**THE PHYSICS OF LANGUAGE:
TOWARD A PHASE-TRANSITION OF LANGUAGE
CHANGE**

by

Nathalie Prévost

B.A., Université de Montréal, 1986

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the Program

of

Special Arrangements: Philosophy, Computing Science

© Nathalie Prévost, 2003

SIMON FRASER UNIVERSITY

July 2003

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Nathalie Prévost
Degree: Doctor of Philosophy
Title of Thesis: The Physics of Language: Toward a Phase-Transition of Language Change

Examining Committee:

Dr. Jonathan C. ~~Driver~~, Chair

Dr. Raymond E. Jennings, Senior Supervisor
Professor, Department of Philosophy

Dr. F. David Fracchia, Supervisor
Assistant Professor, School of Computing Science

Dr. Pierre Zakarauskas, Examiner
Assistant Professor, Department of Ophthalmology
University of British Columbia

Dr. Bernard H. Bichakjian, External Examiner
Professor Emeritus, French Linguistics
University of Nijmegen, The Netherlands

Date Approved:

July 28, 2003

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

The Physics of Language: Toward a Phase-Transition of Language

Change

Author:

(signature)

(name)

July 28, 2003
(date)

Abstract

Interdisciplinary work, such as this, often requires a reform of more traditional notions in order to include novel views. *Functionalization* is such a notion and much of the thrust of this work is that for purposes of explanatory theory, the notion must be redefined. In this thesis, *functionalization* involves a much broader class of linguistic behavior than what is usually understood in linguistics. We understand *functionalization*, as linguists understand it, to involve a loss of lexicality, but by this we mean a loss of a particular sort of lexicality and that particularization has the consequence that the term as we use it applies to a much broader class of linguistic items than is usually comprehended by the term.

Our specific sample involves instances of *perceptually bound lexical* vocabulary that acquire a function of *perceptually divorced uses* - such as connectives, but also mental, religious and ethical vocabulary. This definition of *functionalization* includes but is not bound by the notion of *grammaticalization* and as such departs from its traditional linguistic use. We have chosen this vocabulary because it lends itself more readily to reform than does the more entrenched notion of *grammaticalization*.

Given this definition we notice that a large portion of our linguistic uses are functionalized - that is, have perhaps discoverable causal roles, but causal roles, the explanation of which *must be* an evolutionary explanation that draws upon the causal roles of ancestral forms of speech. If such vocabulary is to be given definition, then it must be given definition in terms of language that has undergone a similar transformation

The research reported in this thesis treats language as a physical system subject to the same forces as other systems describable in physical terms. Moreover, we suggest that the language of mechanical transformations such as first-order phase transitions is directly applicable to such fundamental structural changes in language. When it is borne in mind that the process of grammaticalization is itself evolved, it becomes evident further that

such physical phenomena as first-order phase transitions could explain the emergence of grammaticalization itself, and therefore the original advent of syntax.

With much love, to my parents Claire and Alexandre. For the obvious reasons and their unconditional support.

“Anything that, in happening, causes itself to happen again, happens again”

— *Douglas Adams, MOSTLY HARMLESS, 1992*

Acknowledgments

There are three people I want to acknowledge most especially:

First, R.E. Jennings, who saw something in my initial incoherent babble and helped me transform my insight into something I can be proud of. Because of him, I think I can call myself a scientist and a researcher, roles that I have always aspired to.

I consider both you and Mary my family. Thank you.

Second, Loki Jorgenson, who stepped in generously after Dave Fracchia left SFU. He was willing to entertain my crazy ideas about phase transitions and language and helped me understand the underlying physics that could sustain my claims. His insight (with writing as well), friendship and support has been indispensable.

Loki, a long time ago you promised that you would not let me fail and you have fulfilled your commitment. You are always there for me and I truly appreciate it. Thank you.

Third, F. David Fracchia who told me early on that he was giving me his support because he thought I would go far. He didn't know it at the time but that kind of unconditional support gave me wings. I am only part of the way now but I hope I will fulfill his expectations.

I was sorry to see you leave SFU because I knew that I would miss your boundless optimism. You really went to bat for me both in the difficulties I faced early on and in the form of indispensable financial support, later on. Thank you.

My two best girl friends also deserve thanks and acknowledgment: Maria Lantin who provided the Java version of my Ising model and Gaia Marsden who helped make sense of my writing.

Thank you so much.

Also, thanks to Elli Epp for many insightful conversations and references.

You are a friend and in many ways I consider you a mentor. Thank you.

I would like to thank Phylis Wrenn and the Dean of Graduate Studies staff who were so helpful in ironing out many of the administrative difficulties I faced early on. Thank you.

I also thank Michel Fleury whose brilliance and scientific curiosity stirred me towards higher learning again and this particular path. Thank you.

Finally, I thank the Social Science and Humanities Research Counsel(SSHRC), who provided three years of funding, the Center for Experimental and Constructive Mathematics (CECM) and the Laboratory for Logic and Experimental Philosophy (LLEP), for providing funds, equipment and a place to work. Thank you.

Contents

Approval	ii
Abstract	iii
Dedication	v
Quotation	vi
Acknowledgments	vii
Table of Contents	ix
List of Figures	xiii
1 Introduction	1
1.1 Opening Remarks	2
1.1.1 Philosophy of language	5
1.2 Time scales to linguisticity	9
1.2.1 Grammaticalization and evolution of language	10
1.2.2 Language; incidental or not	11
1.2.3 Criticality and language	12
1.3 The Underlying Facts of Language: Evasion and Vocables	16
1.3.1 Versions of a Language	19
1.4 Propagation and Efficiency	22
1.4.1 Attenuation in vocables	25
1.4.2 The dynamics of shareability	26
1.4.3 Shareability of language in a population	29

1.4.4	Transcategorical changes	31
1.5	Phase Transition in the Evolution of Language	33
1.6	The Ising model	35
1.6.1	Calvin and competition	38
1.6.2	Integration of Calvin and Ising	39
1.7	Description of Chapters	41
2	Loss of Meaning	43
2.1	Evolution of language	43
2.1.1	Logic and understanding	43
2.1.2	The process of delexicalization and functionalization	46
2.1.3	Scope misapprehension	47
2.2	How Physics and Language are Related	50
3	Efficiency and Language	51
3.1	The Swiss Army Knife	51
3.2	Language and Efficiency	53
3.2.1	Language as a physical intervention	54
3.2.2	Exploitation of incidental effects	54
3.3	Efficiency in Other Language Theories	56
3.3.1	Neurobiology's account of language emergence	56
3.3.2	Psychological parallels	57
3.3.3	An argument against structural predisposition	58
3.3.4	The plasticity of language	58
3.3.5	Head-first and head-last languages	59
4	The Physics of Language	62
4.1	Dynamics of Change in Language	62
4.1.1	The Coordination role of linguistic activity	65
4.1.2	Time scale	67
4.2	The Language of Physics	68
4.2.1	Dispersion	68
4.3	Attenuation	71
4.3.1	Eccentricity and elasticity	72

4.3.2	Elasticity in physics	74
4.4	The Effects of Attenuation	75
4.4.1	Lexical to functional	75
4.4.2	Attenuated non-lexical vocabulary	76
4.4.3	Attenuated functional vocabulary	76
4.5	Attenuated Language and Degenerate Physical States	78
4.5.1	Degeneracy in physics	78
4.5.2	Degenerate states in language	79
4.5.3	Structural changes and degeneracy	81
4.6	Phase Transition	82
4.6.1	Phase Transition and Language	84
4.6.2	Phase transition and the evolution of language	85
4.7	Self-Perpetuating	86
5	Phase Transition Models in Biology and Economics	88
5.1	Biology and Phase Transition	88
5.1.1	Kaufman and percolation	89
5.2	Economics and Phase Transition	91
5.2.1	Economics and the Ising model	92
5.3	Neuroscience and Phase Transition	92
5.3.1	Neural networks and the Ising model	93
6	The Ising Model	95
6.1	The Ising Simulation	95
6.1.1	First-order phase transition and degenerate states	98
6.1.2	Computational implementations	101
6.2	Ising and Language	102
6.2.1	Ising for language	106
6.3	The Experiment	109
6.3.1	The simulation	111
6.3.2	Results and interpretation	111
6.3.3	What we have demonstrated	113

7	Neural Darwinism and Physics	115
7.1	Superficial Layers of the Neocortex	115
7.2	Triangular Array and Hexagonal Firing Patterns	116
7.2.1	Spatio-temporal patterns or active memory vs spatial patterns of lin- gering memory	119
7.3	Schemas and metaphors	120
7.3.1	Algorithmic shortcuts	120
7.3.2	Calvin and phase transition	121
7.3.3	Phase transition and functional vocabulary	122
7.3.4	From pre-grammatical to grammatical	123
7.3.5	From infant to adult	124
7.3.6	The difference between metaphorical and functional language	124
7.3.7	Institutional vs spontaneously occurring metaphor	125
7.3.8	Population Dynamics and Calvin	128
8	Population Dynamics, Prosody and Neural Effects	129
8.1	Neural Competition Amongst Many Brains	130
8.1.1	Early functionalization	131
8.1.2	Exploitation of auditory processes	132
8.2	The Process of Generalization in Perceptual Structures	133
8.3	Prosodic Forces	135
8.3.1	Plausibility of context	137
8.3.2	Prosody and dispersion	138
8.3.3	Exploitation of Incidental Effects in Population Dynamics	139
8.4	Attenuation in Neural Terms	141
9	Conclusion	143
A	The Physics of Language: The Fabric Metaphors	146
A.1	Quine's Fabric	146
A.2	Our Fabric	148
	Bibliography	152

List of Figures

6.1	Meanfield solutions for a first-order phase transition. The red (thin) lines show all possible theoretical solutions for the order parameter, while the green (thick) line shows the stable solutions. Notice the abrupt change in spin orientation as the heat exchange increases.	99
6.2	Meanfield solutions for a second-order phase transition. The red (thin) lines show all possible theoretical solutions for the order parameter, while the green (thick) line shows the stable solutions. The dash lines illustrate the second-order phase transition. Notice the slow change in spin orientation; as the heat exchange increases, spin orientation becomes distributed between up and down.	100
6.3	The Cartesian grid is wrap-around to reduce boundary conditions. Because the grid is square it is difficult to illustrate that corners are also wrapped so there is continuity between all edges.	102
6.4	Vocable (spin) grid. Vocables can be in 1 of 2 states; light is lexical while dark is functional. The white line indicate the ratio of light to dark for every row.	103
6.5	Attenuation grid. The different shades represent degrees of attenuation; in this case constituents can be attenuated up to a 7 to 1 ratio.	104
6.6	This graph illustrates how the critical point (green or thick lines) at which a first-order phase transition occurs moves inwards for constituents ranging from 1 state of degeneracy to 3 states of degeneracy. The x's indicate the state of constituents. Those with 3 states of attenuation have become functionalized because the activity level is past their critical point. The other constituents remain lexical because their respective critical points have not been reached.	105

6.7	Activity grid. This is a "Maxwell" demon grid. A random level activity is assigned to every vocable constituent in the system. The overall average is the activity in the system. It is equivalent to the level of linguistic activity that can occur in a population of language users.	107
6.8	State space of a two dimensional Ising model with a first-order phase transition. The light represents the solid portion while the dark represents the liquid portion The dashed lines illustrates the critical point and the critical temperature at which a first-order phase transition takes place.The region outside of the interface is a space of incoherence where spin activity is so high that it is in disordered state.	110
6.9	In our simulation, constituents can be attenuated up to a 3 to 1 ratio. The darkest area indicates that constituents in that region do not have havean attenuation bias.	111
6.10	The spin grid illustrates the effects of activity in the system in the context of degrees of attenuation in vocables. Notice that the lower portion of the graph shows a greater distribution of dark, indicating that functional instances are in a majority. The middle portion illustrates some clustering which may indicate occasional functional instances however, unstable.	112
7.1	Calvin's model: The triangular grid models the particular way in which pairs of neuron that are in mutual re-excitation mode tend to entrain sub-threshold neurons that sit equidistant from the members of the pair. Furthermore, adjacent edges of excited triangles may co-opt a forth and a fifth cell to complete a large hexagonal synchronous patterns of triangular arrays. This creates a kind of "hot spot" that can potentially excite entire structures through its influence over nearest-neighbour interactions - interactions with surrounding constituents. Here, the green and red hexagon represent competing patterns for neural territory.	117

Chapter 1

Introduction

Language is an integral part of what it is to be human. It is a physical reality. Ironically, although it is such a large part of our daily lives, we do not intuitively perceive it as something that we can observe, something about which we can conduct experiments and make scientific observations. This thesis shows that language *can* be observed and that meaningful and interesting claims can be made about it.

In this thesis, the claims are formed in the framework of physics and using its scientific tools. Speaking about language may strike some as a poetic pursuit - or even an ineffable one - but we believe it actually invites the rigor and elegance possible within physics.

Beginning with the assumption that language is a physical phenomenon, we conclude with a theory that explains the way it has emerged and developed across time and place. In the process, we uncover some startling facts that support our approach. Specifically, the language of physics that we engage here examines the *phase transition* from crude wielders of a *proto-language* to the subtle users of a complex *grammatical system*. The kinds of transitions that occur in the dynamics of language are familiar to us because they occur in many other physical systems.

All of us who have watched ice cubes melting in our scotch have observed a phase transition, in this case, a structural change in the arrangement of water molecules from crystal to fluid. More broadly, a phase transition is associated with a change in a state variable. Consider the spatial relationship between molecules of a substance such as H_2O . Emergent properties can be the consequence of a “transitional process”. Referring back to our drink example, the emergent property is the amorphous fluid quality of water. We will invoke this vocabulary in our theory of language behavior.

Should we leave our scotch sitting too long on our desk, we will see the ice demonstrate the features of phase transition. In other words, the ice will change completely from crystal to liquid form. In the process some features are lost such as the presence of a solid and a clear distinction between water and alcohol. We will experience the results of *attenuation* - much to our dismay - should we now take a sip of our diluted drink.

Attenuation, we will show, is a key concept in our language theory. It is the erosion of *perceptual structures* - synaptic features - that distinguish usages between instances of vocabulary. To help define attenuation we rely on a diachronic account of language that begins with lexical vocabulary. Attenuation is a process that provides for the shareability of language and promotes structural changes in several aspects of *linguisticity* such as syntax, morphology and phonetic features. To demonstrate our claims, we employ a well understood and formally tractable computational model known as the *Ising* model. It is often associated with the modeling of structural changes in a system. The usefulness of the Ising model is in its simplicity and its mathematical predictability. Our thesis endeavours to unpack this terminology and make meaningful our claim that fundamental structural changes can occur in language without changing the basic nature of the system, that even the most primitive utterances can give rise to complex grammatical structures.

1.1 Opening Remarks

Language is a physical phenomenon. Organisms¹ may emit complexes of sounds, and commit complexes of marks to surfaces, and each of these modes of action seems to have physical effects on certain classes of other organisms of the species. In other words, people speaking and writing affect other people or cause them to react. Indeed we can categorize organisms according to the kinds of aural and inscriptional complexes they are capable of having as systematic effects.

The foregoing can be expressed in more familiar terms. By means of spoken and written language we communicate, for example, our desires to other humans. Success depends upon their understanding of what we say or write. And although our choice of interlocutor determines which fellow human is actually persuaded to do our bidding, in relevant respects the bidding that gets done is the same, and that commonality accounts for what we refer

¹for present purposes *organism* means *human organism*

to as the meaning of the spoken or written language. For example, the effects of asking for a cup of coffee will be similar whether Bill or Jeff executes the deed.

In this thesis we will eschew the more familiar language - the language with which we are more familiar, our everyday tongue - in favour of the more neutral account. We seek an explanation, and the comfortable familiarity of our language is apt to fool us into thinking that there is no aspect of the phenomenon that we do not already sufficiently understand.

In the language of linguistic description, we say that the causal significance of spoken and written language (in the example, the connection between the complex of sounds or inscriptions on the one hand, and the abstracted commonality on the other) is *conventional*. The sound complex works because the addressee is *au fait* with the conventions; the interlocutor understands the convention. But what, in more neutral terms, can be meant in saying that the connection is *conventional*? At least that there is nothing in the physical character of the complex itself that accounts for the connection. In consequence, many distinct complex-types will have the same effect within a single linguistic community, and no single complex-type will have the same effect across all linguistic communities. Independently of a certain attunement or adjustment of the interlocutor to the complex, the complex has no power to bring about the effect. It must mean at least that the causal connections that do exist between complex-types and these abstracted commonalities must themselves have causal histories. It has in some way *come about* that a particular sound complex has that connection with that physical state of things.

Perhaps even the requirements of *understanding* on the part of the addressee and of the presence of a *convention*, are an exaggeration.

In broad terms, David K. Lewis[40] analyses social conventions as regularities in their solution of recurrent coordination problems, that is, situations of interdependent decision in which common interest predominates. For Lewis, systems of communications, such as language, are proven to be *conventional* by the assignment of particular truth conditions to sentences or other items in the vocabulary.

It is unclear how much we can verbalize the understanding we have (or require to have) of the individual elements of speech. And in the ordinary way in which we speak of conventions, we have *learned* no such conventions, much less attended any. We simply find ourselves, at a certain stage of linguistic development, attuned to the requirements of speech. Perhaps the neutral language of dispassionate observation provides a more accurate account of the available data.

This thesis concerns itself with providing an observationally supportable physical description of what is usually referred to as the *conventionality* of language. In fact, the very notion of conventionality provides an excellent illustration of the features of language that interest us, since it is vexingly difficult to say what it is, yet language users typically have no difficulty in knowing when and how to use the term. *Meanings* are said to be conventional. That is, for a particular meaning, there is nothing about the physical character of a word that better suits it to bearing that meaning than some other. In this way, words are *arbitrarily* matched with meanings. This is not to say, however, either that words have, in general, been arbitrarily *chosen* to go with meanings, or that it is locally accidental that words have acquired the meanings they have. The connections between words and their meanings do have histories, and their histories do afford some explanation of their association. To be sure there may have been some random or unpredictable events in that history but nothing that would warrant that understanding of arbitrariness as it is applied to language and meaning. We stress that, *meaning* itself, as a word and a concept, also presents a semantic challenge. Its use is subject to the forces which influence the use of all vocabulary.

There is, however, an apparent element of commonality in our uses of language, that may warrant our speaking of the conventionality of use. Thus, though we have difficulty in saying what conventionality is, we generally find nothing puzzling about the uses of the word that we encounter in the speech of other people. Somehow we come to participate in these conventions. But the conventions are of a character that we can participate in them without being aware in any detail of what the conventions are. It cannot be excluded that the conventions of language use are in part rooted in our physiology and, in part, in the nature of the features of our social organization. To put the point more dramatically, it cannot be excluded that the conventionality of human linguistic behavior is like the conventionality of human mating behavior. Some aspects come under conscious control, but much is triggered in ways over which we have little control at all, and neither category of behavior need be precisely or deeply or even consciously understood by individual human organisms for them to function in ways that occasion suitable responses. And again, though some such conventions have a cultural history, and would have been different had cultural development been different, we do not suppose that they are arbitrary in any other sense.

Not all of language presents this kind of puzzle and, in respect of such vocabulary the question of conventionality settles itself differently. For example, indexical vocabulary (such as pronouns of person, place, and time, (eg. I, here, now)), object vocabulary (such as

concrete nouns (eg. giraffe, chair, house)), property vocabulary (such as adjectives of color and size), are comparatively unproblematic. We can readily give an account of the occasions in which they are used, and of their meaning by reference to things of the right sort. Within this category fall those meanings for which Kripke's theory [39] most readily applies. Kripke proposed a causal theory of meaning in which proper names and natural kinds are not solely definite descriptions. More specifically, they are rigid designators whose reference point to objects identically across all possible worlds.

Nevertheless, though we can readily accede to Kripke's account of meaning for some such range of vocabulary, doing so depends upon not questioning too closely what we understand by the word *meaning* in his claims. And the words *meaning* and *causal* are ones to which his theory less adequately applies. Indeed, much of the language with which philosophers find themselves preoccupied - the language that occasions attempts at the construction of philosophical theories, theories of meaning, theories of causation, theories of mind, and so on - falls within the category that concerns us here. As mentioned, this account does not itself offer anything that could be called a philosophical theory, though it does have implications for the likelihood of success of such theories.

1.1.1 Philosophy of language

Philosophy of language addresses how the uses of some vocabulary have particular effects by saying that it is because of their meaning. Part of the problem with this answer is that *meaning* is borrowed from conversational use and is not given much explanation. Philosophy of language is more preoccupied with the effects of meanings than a detailed description of what a meaning is. The tenets of philosophy of language are well outlined in the standard reference works.

Consider the *Encyclopedia of Philosophy* [1]. William P. Alston enumerates six fundamental questions asked by philosophy of language:

1. What is it for a linguistic expression to have a certain meaning?
2. Under what conditions do two linguistic expression have the meaning?
3. Under what conditions are we justified in saying that a word has two same senses in two different contexts?
4. What are the ways in which the meaning of an expression may be more or less vague?

5. What is the difference between a literal and a figurative use of a word?
6. What kinds of meaning are there? For example, is there any distinction between cognitive and emotive meaning?

These questions are all dependent on a definition of *meaning*. As of yet, no theory of meaning has satisfactorily met the criteria of general inquiry. The yardstick of a good theory should include the following criteria:

- that it provides adequate detail,
- that it is consistent with other theories,
- that it offers adequate predictability,
- and that the subject matter does not resist the idiom in a way that important features may be ignored in order to comply with the idiom².

The Platonic notion of universals has lingered even into the twenty-first century in which the idea of truth is intimately linked to meaning. For Plato, every object in the world is an imperfect instantiation of its form, or of the idea of itself, that is, every object in the world holds some essential quality that is universal, ideal. These essential qualities are a priori, that is, somehow extant before things came to be in the world and they are assumed to be timeless. There are layers of universals and the purest form is *Truth*. It is assumed that there is *Truth* in everything in the world but that our perceptual tools are misleading. Since our senses are themselves imperfect, much like the objects of this world, we must rely on rhetoric

²Scientific curiosity used to be of philosophical interest but somehow, the requirements of physical theories alienated philosophical concerns that seemed to belong to a different category, that is, all that can be classified as *human experience*. Human experience is that elusive sense of being in the world with all that it entails; notions of beliefs, hope, intentions, mind. This vocabulary seems to capture something that philosophers would say is *intuitive* about our state of mind. For many philosophers it seems that these words capture something essential about being human, something that would be lost if we were to abandon this language and adopt a purely physical idiom. It may be that the sense of privacy that seems to be attached to mentalist language holds a kind of magic that inspires awe, the kind of awe that we feel witnessing a magic trick. Once the trick is explained some may say that the magic is lost and that we have lost something essential. But could it be that the cleverness of the stunt inspires awe as well?

It could be a matter of taste for it seems that in addition to the idea that there is a category of phenomena that can only be described using private language, philosophers of language have presumed that language is one of these concerns that belong to the type of human experience that cannot be fully described in terms of a physical theory. This may be why philosophers of language have clung to so much of a language that has been inherited from the ancient Greeks.

to access universals and specifically *Truth*. Rhetoric is the practice of rational discussion about objects in the world. A consensus amongst philosophers about the resemblance that is shared as some abstract quality between objects, is a bridge but not a definite proof of universality and a window to the *true* nature of things [15]. The importance of apprehension of *Truth* in Plato is that it leads to moral behavior. Though Plato has not focused on a study of language, many after him have tried to give a detailed account of Universals that could be applied to the objects of language.

Like Plato, Frege presupposes a priori truths that are somehow different from a sensory apprehension of the world, which can sometimes be partial. Frege's identity claim postulates a separation in linguistic propositions, between the essential truth of an outside reference and a sensory truth about the presentation.

Frege has discriminated between sense and reference to establish the nature of meaning. Frege's inquiry is an attempt to disambiguate the notion of identity. This study started in mathematics with a formal investigation of the constraints of the identity (=) symbol. Later on, Frege extended his theory to propositions of natural language. His example: *the morning star = the morning star* is a *true* statement of a priori identity while the statement *the morning star = the evening star* is also *true* because of they both refer to the planet Venus. However, different information is put forth accordingly as the morning star or the evening star is used to refer to Venus or not. Frege goes on to claim that both expressions have the same reference, that is, a relationship of identity to itself, the planet, but different senses as the presentation of the statements give us different relevant information about the object of reference. For Frege, the truth of each statement of identity lies in its reference. However, we cannot ignore the manner in which the information is presented and how the sense of a statement adds to the truth of an identity relation.

Russell [54] argues that logical vocabulary such as *or* occurs in natural language as a choice of the form *p or q*. This kind of statement can only be true by the facts that *p* and by the facts that *q* represents. The use of *or*, Russell says, does not represent any facts. It belongs to the world of propositions. In the world of propositions *or* refers to a person's disposition of hesitation.

Some, like Grice [26], suppose that the inability or difficulty in defining meanings is a problem linked to lexical vocabulary which is thought to be semantically difficult. He thinks connective vocabulary is semantically easier, that it lends itself more readily to rational inquiry because a formal account of it can be given from truth functions. We do not think

this is so.

In fact we think the opposite is true. Later on we will show that a semantic theory is more readily available to lexical vocabulary and that connective vocabulary cannot be reduced to a truth-functional account.

However, Grice is important to us because he connects hearer and speaker in a relationship. In this research we preserve this relationship and we will show that it is an important aspect of our theory.

This research diverges from the philosophy of language approach dramatically as we have found the account for meaning unsatisfactory for our goals. In pursuit of theoretical understanding of language, we direct our focus to the fundamentals of scientific process and substance. Going back to the root of scientific inquiry allows us to assess the usefulness of scientific methods and theories in our attempt to describe language.

Fundamentally, we argue that Nature is economical in its use of mechanisms and dynamics. In that light, we suppose that a physical phenomenon such as language may exhibit behaviors that belong to a class of behaviors common to a more basic phenomenon. For example, we might find a similarity between the behaviors of fundamental systems such as simple gases and seemingly abstracted systems such as language. So we borrow methods of inquiry, mostly from physics, to explore the phenomenon of language. We expect a physical theory of language to be consistent with findings in physics since it shares in its method and idiom. We also assume that such a theory will provide an explanation of why a theory of meaning seems to resist a clear description and why it is discontinuous with the body of scientific knowledge.

We offer a causal history of language and the complex dynamics involved in the inheritance of it, across a population and its generations, as an explanation for the relationship between linguistic utterances and their effects. In doing so, we will use several models offering a range in detail, situated somewhere between the levels of descriptions of synaptic functions and population dynamics. That is, terms from these various disciplines are borrowed and repurposed, just as physicists may talk about *surfaces*, *interactions* or the way mathematicians refer to a *field*. Theories, such as physics and mathematics, are bound to use theoretical language. A theory defines object types that are of interesting generality so that important features are represented. With our theory of language, there is no need for a reform of language, merely a redefinition of certain terms borrowed from conversational language.

Dispersion is such a term. It is commonly used but also precisely defined in physics. We will apply it, and others of that sort, in the context of language analysis.

1.2 Time scales to linguisticity

We may assume that the greater the temporal distance, the greater the differences in *linguisticity*. In the short term, differences may be hard to distinguish, but in the medium- to long-term, differences monotonically increase over time. In the extreme, we may not even recognize the differences as being part of linguisticity at all. Nonetheless we assume that they are related.

We also assume that the character of short-term changes has itself changed over time. If we divide time into ten thousand year epochs, then during each previous epoch there may have been changes in the fundamental character of the linguisticity of the human or proto-human population. To describe the changes between any pairs of more recent epochs, we might require a different sort of observational language. Whether there could be a general theory that covered the dynamical character of both examples may also be doubtful. One may have grammatical categories that the other lacks, for example. Certainly, over a sufficiently long period there would be no reason to suppose that the grammatical categories were the same.

Derek Bickerton's [10] research into language evolution has led him to believe that some evolutionary traits have emerged by radical steps towards the use of language from *proto-language* - loosely put, a language without a syntax. In his view proto-language lacks the grammatical components that a natural language does.

In this case, the grammatical categories that describe proto-languages are absent and there is a marked difference between the description of Bickerton's proto-language and natural language. This would suggest that not all evolutionary increments are achieved at a regular slow pace and also the increments are quick, fairly large steps that do not involve intermediate ones.

A theory of sudden evolutionary changes for a biological phenomenon, such as the developments towards our linguisticity, does not sit well with traditional evolutionary theories and idiom. However, Bickerton is not alone in his assumptions, Stuart A. Kaufman has suggested similar *jumps* in genetic evolution[37]. Both have had to borrow from several idioms to explain their views.

This thesis uses a similar strategy. We will borrow from physics to describe behaviors in the evolution of language that could be understood as sudden changes in linguisticity. Moreover, we will suggest physical models that account for the behaviors.

The language of a theory of language may have to change with major changes in the language itself. A theoretical language of sufficiently high order, such that it might be applicable to language at all stages of its evolution may be practically inaccessible. And we might find nothing to say.

1.2.1 Grammaticalization and evolution of language

If we are to talk about the boundary between linguisticity and non-linguisticity then we must include some remarks about grammaticalization because the study of it can provide a more fine-grained boundary.

Grammaticalization is part of a physical theory of language - and so it is part of our theory as well - as it describes the evolution and emergences of several features in language. Grammaticalization is the study of emergent features - changes in construction to be understood as structural or organizational - that we categorize as grammatical relative to non-grammatical forms[44]. Opinions vary as to what should be included in a definition of grammaticalization. For the purpose of our research, grammaticalization will refer to the emergence of syntax in language.

Traditional inquiry in the field tends to describe grammaticalization as the evolution of morphemes by phonemic transformation [44]; a somewhat unidirectional process of the erosion of lexical uses through the loss of phonemic parts. Early lexical vocabulary undergoes reduction - shortening of expressions - and relies increasingly on affixes - at the end of vocables - to produce effects in linguistic interactions. Others, like Hurford [30], argue against the unidirectional trend towards increasingly abstracted forms for lexical vocabulary. There are examples that show re-lexicalization, that is, vocabulary that migrate into a different category, hence regaining an independent status, through *affixes mutation* [22].

The processes by which grammaticalization occurs are, as of yet, not transparent. However, most linguists will agree that all dynamics leading to grammaticalization should be included in this area of research[23].

The description of grammaticalization phenomena may be extended to include the dynamics that describe the emergence of linguistic forms from pre-linguistic forms. This is not an obvious extension because linguists do not all agree on the nature of the evolutionary

process by which our linguisticity was produced. Changes in vocabulary may indicate a kind of evolution. However it may be that the source of these changes is the consequence of the fine tuning of a *Language Acquisition Device* (LAD). This device would be genetically programmed and would not be subject to evolutionary forces[30].

So far, grammaticalization theory has concentrated mostly on the changes in lexical linguistic forms and meanings, and the subsequent formal and semantic changes that such material undergoes over time. These kinds of changes are observable in all human languages.

1.2.2 Language; incidental or not

Grammaticalization is most often studied in the larger context of language evolution. Though most who are interested in language agree that Darwinian evolution has a hand in the emergence of the phenomenon, there is no real consensus as to how this happened. There seems to be an intuitive notion that in whatever manner language happened, its development was a choice. It is not clear that this intuition can be well founded.

As we have seen, Hurford, postulates that the genetic evolution of humans has led to a LAD which is now heritable, imbedded in our genes. The concept of LAD implies that its function is to deal with grammaticalized language and is *for* communicating.

Michael C. Corballis [16] is also one that believes in language as a human/hominid pursuit though not from a genetic standpoint. Corballis believes that before humans used spoken language they used a form of gestural language. Corballis emphasizes that spoken language is fairly recent in the history of humanity - only around 50 000 years old. He thinks that grammatical languages might have emerged before the event of spoken language - a form of American sign language - and that the capability for spoken language may have been possible long before it happened. Early humans may have adopted speech simply because it seemed like a *good idea*. This would make the event of speech a cultural invention much like writing, and not the result of structural changes in the brain.

Corballis has based some of his ideas on the fact that 50 000 years ago there was an explosion of new technologies: new tools, drawings, textiles, musical instruments etc. In contrast, there was little evolution in artifact 2 000 000 years before that. Corballis hypothesizes that with the advent of spoken language, hands were now free to do other things.

However, there is another point of view: Language, as we know it, is the result of a lengthy process subject to evolutionary pressures of which language is a consequence and not the cause. That is, language is an incidental consequence of a series of events.

John Haiman [28] suggests that two fundamental principles, habituation and emancipation, are fundamental dynamics that can usefully describe the process of grammaticalization. Habituation and emancipation are concepts borrowed from psychology and ethology that describe ritualization of behavior through repetition. Habituation may have its roots in communicative behavior that become eroded in form and consequently in function; however, emancipation refers to non-communicative behaviors that become freed from their instrumental action to develop a semiotic or symbolic role.

William H. Calvin postulates that language is the incidental effect of the required fine tuning of motor functions for ballistic movements and not a consequence of an evolutionary process that led to a (LAD). Calvin proposes that motor functions and the processes of language are synaptically related. The requirements of ballistic movement are such that once a missile is launched the trajectory cannot be modified. Calvin believes that in order to target with any accuracy, early humans had to develop, in conjunction, the ability to fine-tune hand and arm motion and the ability to pre-process synaptically the consequences of their throwing action. This, Calvin believes, also promoted the specialization of particular motion controls in specific areas of the brain, that is, the superficial layers of the neocortex.

1.2.3 Criticality and language

Many language theorists assume that there is a clear boundary between linguisticity and non-linguisticity. They believe, then, that spoken language is what distinguishes human from non-human. Unfortunately, this categorial distinction is challenged by Darwinian evolution, advances in evolutionary biology are more finely grained than a single definitive jump to linguisticity.

On second viewing, however, their assumption may find backing in the language of *criticality*. The notion of *criticality*, as understood in statistical mechanics, covers a range of behavior in Nature that has one thing in common - the onset of a sudden change in observable features that constitute a system. The concept of system is one that is adequately supported by the framework of *statistical ensemble* - the descriptors of groups of elements that share common features. Examples of criticality involve the cyclical explosion in gypsy moth populations, structural changes in water as it turns to ice, catastrophic environmental changes such as earthquakes, volcanic explosions, tsunamis and so on. Here examples of systems are the populations of gypsy moths and water molecules.

Physics, biology and economics are fields that have used the concept of *criticality* to explain and formalize many phenomena. *Criticality* is often associated with structural changes in systems. These changes can be striking enough to confuse us in our beliefs about the nature of a particular system. For example: without experiencing the melting of ice it may be difficult to realize that both solid and liquid are related states of H_2O . The molecular constituents of H_2O are unaffected whether it be in a liquid state - water - or in a solid state - ice. The states are dramatically different. In the case of H_2O , the state transformation from one to the other is brought about suddenly at a critical point at which the presence of heat, or lack thereof, will either melt the ice or freeze the water. We believe that in the study of more complex phenomena, similar behavior occurs can go unnoticed because the structural changes are so dramatic that related states of a system are considered separate systems. Language falls into this category.

Chomsky [13], for example, was one of the first people to advocate a *Big Bang* theory of language with the suggestion that change occurred at the “genetic” level. He was implying that fundamental structures - genes - had been affected with the emergent features - linguistic capacities - being the consequence.

We would like to paraphrase that view in terms of elemental physical processes. For our purposes, these processes can be described as mechanical, chemical and nuclear, each operating at increasingly fundamental levels.

- *Mechanical* processes do not transform the constituents of materials; instead, they typically are rearranged. The change from water to ice is an example of a mechanical transformation from liquid to crystalline solid.
- *Chemical* processes can transform the constituents of materials at the molecular level. Fire is a good example, consuming certain substances and producing others that are composed of the same atoms but different molecules.
- *Nuclear* processes like fusion or fission cause changes an atomic level. Stars start with hydrogen and produce first helium, and then heavier elements.

The simplest processes, mechanical, are the most abundant and least fundamental. For Chomsky, the changes that gave rise to the new linguistic abilities occurred at a more fundamental level altering the underlying structure, not just rearranging the constituents. His use of “genetic” is comparable to the description of chemical.

For ourselves, we advocate a physical change akin to a mechanical one and not a chemical one.

This being said, we are more inclined to side with Calvin and Bickerton on the idea that language is incidental to other modes of interaction, and with Bickerton and Corbalis on the idea that criticality in structural changes plays a role in the emergence of language (but without necessarily involving mutations at the genetic level or fundamental structural changes with the size and general morphology of the brain).

Derek Bickerton [10] thinks that humans stumbled incidentally onto language. He also holds that language is not a cultural event that derived from the need to communicate. Rather *off-line* thinking - the capacity to imagine - possibly brought about from proto-language, opened the possibility for the emergence of language. Bickerton's theory stresses the impact of evolution on humans and how language has been used to preserve homeostasis, that is, the preservation of conditions that sustain the the viability of an organism. Simply put, language is a by-product of survival and more specifically, an incidental effect of the capacity to access situations pre-emptively.

As we have said, some of Bickerton's ideas imply notions of criticality in structural changes in the emergence of language from proto-language. Bickerton defines proto-language as a primitive form of language that would be as much part of early human biological endowment as present human language is, though without any of the grammatical or formal features that are usually associated with language. He argues that proto-language is more robust than language because it has been part of hominid biological endowment for much longer than language. Moreover, proto-language does not have a critical period of acquisition though it depends on the presence of lexical input. Bickerton thinks that the faculty of acquisition for proto-language and language are disjoint and the acquisition of proto-language does not necessarily lead to the acquisition of a grammaticalized form of language.

Bickerton based this supposition on several studies of pidgin languages (such as pidgin Hawaiian, pidgin English, etc.), linguistic features in early childhood, ape language, and human adults deprived of a linguistic environment. He calls forth these examples to describe the criteria that define a proto-language. According to his theory, proto-languages lack the formalisms of natural language such as a grammatical structure, the use of tenses, and non-lexical vocabulary.

From such examples as pidgin languages, he also attempts to demonstrate how proto-language can lead to language without intermediate steps. Infants make the transition from

infant forms of language to adult language in a few months while the transition from pidgin Hawaiian to Hawaiian creole was achieved only after one generation (Bickerton considers creole languages as meeting all the criteria of natural languages) . Bickerton suggests that hominids may have gone through a similar critical period from proto-language to language, though a *sudden transition*. It may have taken a few thousand years, a short time in hominid/human history [9].

Interestingly enough, in *The Language Instinct* Steven Pinker [46] uses similar linguistic samples to make an argument for the genetic innateness of a linguistic faculty and a shared universal grammar amongst humans. Unlike Chomsky, he is an advocate of evolutionary selection, but believes that the precursor of language is somewhat mysterious. He suggests that linguisticity has probably evolved from genetics (such as primate calls) rather than learned behavior³.

Despite present physical evidence, it is very difficult to establish the origins of language. This is due in part to the difficulty of establishing the scale of evidence, the granularity at which language should be considered. There is a perceived boundary between a language and pre-language state and it is unclear how to capture its quality. Is the boundary gradual

³Primate calls originate mostly from the *forebrain limbic area*, specifically the *hippocampus*, *amygdala* and *cingulate cortex* and hypothalamus. The linguistic ability in humans is usually associated with *Broca's area* clustered around the *Sylvian fissure*, that separates the temporal lobe from the parietal and frontal lobe, roughly behind the prefrontal cortex, rostral to the premotor area, and *Wernicke's area* roughly situated with the *superior* and *middle temporal gyrus*. The midline areas are usually associated with the *initiation of speech*, mostly controlled from the *supplementary motor cortex* and the *arousal/attentional control in comprehension and production of speech* through *Wernicke's area* [18].

It also has been shown that the ability to talk is dependent on cortical motor functions while primate calls are not [35]. But the relevance of primate calls in the production of speech cannot be so easily dismissed. Uwe Jurgen of the *German Primate Centre*, raises this point with the suggestion that primate calls may be involved in the production of speech, or at least that both functions have common cortical structures. Also, Broca's area is not entirely necessary for the production of language. Moreover, it has been shown, from experiments, that primate calls are genetically pre-programmed in their acoustic structure, which supports Pinker's claims, while speech production rely on learned motor functions. However Jurgen is careful about the implications of these finding:

In conclusion, we may say that human speech and monkey call production do not rely on completely different neural systems. There is a set of brain structures common to both. What distinguishes the central speech system from the monkey vocalization system is that on top of the common set of brain structures, a number of additional structures known to be concerned with motor learning come into play. Their involvement seems to be established during phylogeny by strengthening connections between learning-related and more elementary motor structures in the sense that areas connected only indirectly in the monkey become directly connected in man. -Jurgen, 2000, p.11-

or sharp? Is it along the history of humans and hominids or is there a sharp delineation between other primates and humans? These are a few of the questions that still require answers.

1.3 The Underlying Facts of Language: Evasion and Vocables

R.E. jennings is a philosopher who is intrigued by some of the ways we use language. He believes that misunderstandings in the *scope arrangement* of some functionalized relational vocabulary (such as or, and, but, until, unless, while, without, etc.) go unnoticed and lead to new uses of this vocabulary. Jennings refers to these occurrences as *scope evasions*.

The length of the following quote, by Jennings, is understandable in view of the fact that the ideas it contains were the catalyst for this research.

The construction of an explanatory theory [32] raises thorny problems of vocabulary and representation. Space does not permit a detailed justification of the adoption of biological language or the use of descriptors like species or inferential effects. The theory must accommodate many developments that are omitted from this discussion. Suffice it to say that a species is understood as a union of morphologically identifiable populations temporally ordered by an engendering relationship (earlier populations engendering later ones) and that *effects* are understood as neurophysiological effects occasioned by uses and available to later processes of speech production.

The uses of the word *or* provide a good example of the need for such a theory, especially as the assumption that *or* is always disjunctive has befuddled almost all twentieth-century logic text authors and infected the understanding of several generations of theorists brought up on their teachings. Contrast the sentence “It will be useful or it will be harmful” in which *or* is disjunctive, with the sentence “It could be useful or it could be harmful” in which *or* could be replaced by a semi-colon or even an *and* or *but* without serious change of meaning. *Or* has many other non-disjunctive uses distinct from this one. English also affords many examples of connective vocabulary having among their various uses *or/and* pairs of meanings and *and/if* pairs. For an example of the latter, consider the uses of *without* in “She will die without fear” and “She will die without help”. To

see how this dualization comes about, consider the word *unless*, a contraction of the expression on [a condition] less than that. The longer expression ought to mean something like *and not*; however *unless* means something more like *if not*. The reason for this seems to be that the gradually contracting construction that eventuated in *unless* always occurred within the scope of some negative sentence element. But an increasing segment of the linguistic population misinterpreted the scope of the negative element, reading “not (A unless B)” as “(not A) unless B”. Since the two population understood the whole construction in the same way, the *unless* had to be construed as an *or* by the one while being construed as an *and not* by the other. Eventually *unless* migrated to non-negative environments carrying with it the new meaning. In this case the older meaning died out. In many other cases, both meanings survived, generally with the new meaning marked against ambiguity. The development appears to have been a case of syntactic metanalysis, a phenomenon hitherto noted [Jes22] only at the level of word-formation, which has yielded *an umpire* from *a numpire*, *an apron* from *a naperon*, *a nickname* from *an ickname* and so on.

At the syntactic level, the structural characteristics of this kind of mutational development are repeated in all of the many other cases of connective dualization by scope evasion, including cases where negation is not implicated. Five stages can be distinguished:

1. . Stage one: The initiating scope misapprehension (by members of population B) takes place, typically involving negative and modal sentence elements (in the speech of population A). The resulting combination of approximately correct (or at least undetectably incorrect) apprehension of satisfaction conditions on the one hand and undetected incorrect processing of syntax on the other forces an incorrect (or at least novel) apprehension of connective meaning. This is the misapprehension stage.
2. . Stage two: The novel meaning in the B use of the connective is nurtured by the fact of its uses coinciding in satisfaction conditions with A uses. Since at this stage the new meaning may be thought of as lying hidden beneath the old, it is called the succubinal stage.
3. . Stage three: The connective in its novel meaning appears in environments

in which it is not hidden by the older one. But there is a sufficient B population that the use is not corrected, and members of the A population read it as satisfaction conditions seem to dictate. This is the migratory stage.

4. . Stage four: The B meaning is sufficiently established that when the connective occurs in the environments that spawned the new meaning, it is ambiguous as between A and B readings. This is the ambiguity stage.
5. . Stage five: The ambiguity is removed by a marking of the new meaning either prosodically or by the addition of elements that cue the B reading where it is wanted. This is the marking stage. Thus we have for all (compare “I am steady, for I am sober” with “For all I am sober, I am none too steady”); just any (compare “If anyone can join, I can” with “If just anyone can join, then I don’t want to”); even if (compare “If the Queen asks me, I’ll do it” with “I won’t do it even if the Queen asks me”) and so on.

-Jennings, 1998, p.1-

The specific idea that Jennings’ [48] work inspired is how the loose use of some vocabulary leads to unnoticed mistakes, mistakes that have observable consequences in the syntax of language. From this idea has stemmed another; the imprecise way in which we use this vocabulary hints at a dynamics that has profound repercussions on the shared aspect of language. The first question we ask is: Why is it so difficult to enunciate a straight-forward semantic theory for connective vocabulary? Can this question be asked about all vocables of the language?

By vocables we mean all vocabulary of a language and all of its ancestors distinguishable and different from the sort of noises that are usually associated with non-human primates, such as *primate calls*. Not all vocables are as *semantically difficult* as connective vocabulary or become functionalized, but the apparent erosion of semantic space is evident in all diachronic accounts.

Jennings’ work has also involved the tracking of change in *lexical* vocabulary. One of the most dramatic changes that lexical vocabulary can undergo is *functionalization*. Functionalization is the process that generates *functional* vocabulary such as *connective* and *auxiliary* vocabulary. We restrict our definition of *functional vocabulary* to include vocables which come to have a new syntactic role that is almost entirely divorced from their ancestral uses.

Have is such a word: In its original use, *have* is found in the context of ownership: *I have a dog*. But in its functional use, *have* has lost its connection to ownership and has adopted an auxiliary role in the composition of past tense form of verbs, as in, *I have done it already*. Notice, as well, that, in the spoken form, the change in structural role has been accompanied with a morphological truncation, from *have* to the contraction 've. So one would say: *I've done it*.

From observing many such examples, mostly in English but also in other languages, (such as the Indo-European family of languages, sign languages, pidgin and creole languages, child and adult forms of language, speech pathologies) a pattern of behavior slowly emerges. The pattern is one of change and transformation in the use of all vocables. These changes can be historically tracked and we can give a formal account of the dynamics of these changes. Our formal account involves a theoretical description of a system of linguistics that propagates through a population. The linguistic dynamics, from which we derive rules that define our language, explain the examples of language change that surround us.

1.3.1 Versions of a Language

Language is fluid. Like the river Heraclitus stepped into, it is constantly changing and we cannot speak twice in the same language. However, in order to understand linguistic changes, we impose artificial parameters on language. For example, linguists recognize at least three distinct phases (thus far) in the history of the English language: Early English (EE), Middle English (MDE) and Modern English (ME). As anyone who has tried to read *Beowulf* will know, EE is extremely difficult for one fluent only in ME to understand. The time span for these major divisions of English use (beyond which it is difficult to recognize it as an immediate relation) is about a millenium.

It is difficult to pinpoint the exact point when one version of a language changes into another. Changes are always occurring within a language and one must identify which changes are critical to distinguish an epoch change or, alternatively, how many changes need occur before an epoch change takes place.

Moreover, these changes appear at various levels and include phonemic changes, morphological changes, changes in the order and complexity of syntactic arrangement and ultimately, changes that can bring about the emergence of grammatical functions from proto-linguistic forms. At minimum, for a language version to be said to exist, some conditions must hold for a while, and maybe just long enough, for the language system to be shared

in a population.

The rate of language change is difficult to establish, but we assume that smaller changes probably occur often and do not profoundly disrupt the shareability of language. In fact, we will suggest later on that changes *promote* the shareability of linguistic exchanges. Overall drastic restructuring in a language probably happens less frequently. Whether some of these changes happen incrementally or abruptly, the shareability of language is not hindered. Indeed, it is this shareability that defines the specific version of a language as *a* language and not another.

If a verse of *Beowulf* [42] is quoted in a conversation, it probably would not generate any specific effects, except perhaps for those who are familiar with the poem. In fact, it is very difficult to read without a translation. Nonetheless we can still recognize it, especially when heard, as a relative (if distant), of a ME. There is a clear boundary between our understanding, or lack thereof, of English spoken over a thousand years ago and now.

We inherit vocables generationally from our parents and our community; If we follow a diachronic account of these vocables, we notice that even though they may survive throughout several generations, their uses in a conversation may change substantially over time. Some changes in the language are quite dramatic. The *Beowulf* poem is a good example that much of the vocabulary is now incomprehensible to us. But we also observe that the survival rate of vocables is not shared equally amongst them. In *Beowulf* some vocables are more familiar to us than others. That is, the use, the morphology and sometimes the phonetic aspects of some vocables will remain somewhat unchanged for longer periods than others.

The Canterbury Tales are a good example of MDE. Though the text is readable, to a ME user, it is clearly not written in ME. When we look closely, most of the connective vocabulary has morphologically survived until now. But in examining some of the constructions it is clear that the instances of connective vocabulary in MDE are different from those of in present times.

And certes, if it nere to long to heere,
 I wolde have toold yow fully the manere
 How wonnen was the regne of femenye
 By theseus and by his chivalrye;
 And of the grete bataille for the nones

Bitwixen atthenes and amazones;

-Goeffrey Chaucer, The Canterbury Tales, The knight's Tale Part I- [12]

The use of *for* in *And of the grete bataille for the nones* is not the same as the use of *for* in *say, this gift is for you*. If we were to translate the first instance of *for* we would probably use *because* in its stead. The second use of *for* is a use of present times and is used in a context of *benefactive*. While the morphology of this bit of functionalized vocabulary has not changed, nor has its structural role in syntax, we can see that the *benefactor for* has its origins in the *causal for*. Sometimes subtle changes in language lead to more dramatic changes in the syntactic organization of sentences.

Bernard H. Bichakjian has reported changes in language in the order of *subjects, objects* and *verb* use in the family of Indo-European languages. Bichakjian describes a reversal in syntactic constructions from *head-first* to *head-last*. He considers the *heads* of a sentence loosely as verbs and prepositions while their objects are the modifiers such as attributive adjectives. *John's FOOT* and *the FOOT of the bed* are examples in which the noun *FOOT* is the *head*. For each proposition, the *head* finds itself, respectively, *last* and *first* [7]. Bichakjian uses his observations to hypothesize that *head-first* language are more efficient than *head-last* ones, largely because they allow more complex constructions, particularly the use of embedded sentences.

We consider the occurrence of syntactic complexities, such as embedded sentences, a convergence of the erosion of specific effects generated from the use of lexical constructions. As the use of specific vocables becomes less rigid, we need to use more complex sentences to express ourselves. For all its seeming inconvenience, there are at least two reasons for the erosion of lexical uses: First, it can be, as Bichakjian suggests, exploited to increase linguistic efficiency. Second, language needs to adapt to environmental constraints. We will use the language of *exploitation of incidental effects* to describe the emergence of all kinds of vocables.

These are just a few examples that demonstrate a dynamic that starts in the everyday use of language where small unnoticeable changes occur. These unnoticeable changes are cumulative and eventually scale to very noticeable changes to the point of generating new versions of languages. Despite all these changes, thence or perhaps even because of them, the shareability of a language is never compromised.

1.4 Propagation and Efficiency

It is the relative crudeness of language that allows its propagation. It seems likely that the everyday use of language will generally tend to be efficient. That said, any process has multiple ways to produce specific effects. A certain amount of effort is associated with each way or configuration. Efficiency refers to the likelihood of achieving rapidly, or reliably, specific effects with the least amount of effort. Language is an efficient process, having evolved a broad range of specific effects with just sufficient effort that linguisticity has propagated to the whole species. The use of language requires muscular effort and synaptic effort. Some people will do it with more or less detail depending on different architectural constraints - such as being a child or an adult - and, possibly, individual priorities - such as being a writer of novels or a sales person.

Let us try a metaphor to illustrate. Many people have, at one time or another, relied on a *Swiss Army* knife to help in daily tasks. The portability and number of accessories that composes the knife make it ideal for a variety of typical needs. It is often *good enough*.

However, for more involved work, or for professional work, the lack of precision that is inherent to the Swiss Army knife limits its performance. Specialized carving, for example, requires several different sharp tools, spoon gouges for example, for the many levels of detail. Since most of us do not require that level of precision, it would be cumbersome to carry several specialized tools. Carving is just one possible specialized task - by extension, it would become impossible to carry all specialized tools for all possible tasks. Though the Swiss Army Knife cannot perform very specialized work, it is sufficient for most of us that do not need to achieve exacting levels of adequacy and thus, the effort of carrying it is appropriate to our needs.

Efficiency is about striving to reach a balance between effort and consequence. Language, like the Swiss Army knife, is portable, adequate, reliable enough and readily available across a population. However, because of its relative performance, language is not immune to misuse or misunderstanding. In turn, some of these faults will generate new uses for old vocabulary.

Lisa Simpson's astute observation, *Only kids are that incoherent*⁴ as she approaches and hears a group of ten years old speak, points to a particular way of speaking that generates effects within the ten year old population but not in others. An utterance such as *Stuff*

⁴From *The Simpsons*, animated T.V. series on Fox Cable

*sucks*⁵ or *my mom is so gay*⁶ would probably not be used by anyone over the age of ten. The vagueness of the first example and the inappropriateness of the second would not fit the requirements of adult speech, partly because, as we get older, the complexities of speech are more accessible and possibly because the requirements of the everyday life of adults demands more specificity. In the first example, the lexical vocabulary is almost entirely stripped of perceptual context. In the second example, children have coopted *gay* without the complete awareness of its usual use in adult speech. We can imagine that their use of it is somewhat derogatory, as it is often used in adult speech, but the full perceptual context of the word has been lost. In adult speech, a similar erosion of specific effects occurs for all vocables, but, in more subtle ways.

Vocables are entrenched in perceptual cues of a neural kind and uttering them generates specific effects, in hearers, of a similar nature⁷. But the fact is that, for various reasons, there are a finite number of vocables shared by a population of language users and a significantly smaller number for each individual. While the day to day activities of most people are probably routine, small differences do arise which may require specific vocal interventions. With a limited set of vocables, the strategy is to combine them in such a way that the changes in new contexts can be reflected by the refurbished use of old vocables. This linguistic *hooking* - connecting to the past - of vocables ensures that our language remains shareable.

Our capacity to discriminate between similar uses does not scale infinitely. Fortunately the everyday requirements of life do not require us to make very fine discriminations. The discriminations we need to make can be made easily by *hooking* new onto past experiences. This process of generating new, shared, vocables continues inexorably, even through changes so dramatic or numerous to signal a change in language version, as from Early to Middle

⁵idem

⁶From some child in my building

⁷The perceptual cues we are referring to are the kind reported by G. Rizzolatti, L. Fogassi and V. Gallese [53] from studies done on chimps. Chimps are made to perform a manual task while the stimuli are tracked through the ventral premotor area. In turn, the researcher performs a similar manual task while the chimp is watching and while the stimuli are still being recorded. Results show that there is a mirroring effect in neural activity involved with the activity of performing and witnessing a task, hence suggesting a kind of empathic response in being exposed to the actions of others.

We assume that effects produced by speech are empathic effects. What is triggered in observing an action is not an exact mapping of what is triggered for the individual that performed the task, however what is triggered are the individual's own synaptic structures involved in doing and observing the task. But we also assume that sensorily immediate actions such as walking can be correlated more closely in terms of synaptic mapping in individuals than less sensorially immediate actions such as the use of particular vocables.

English.⁸

Why do we not generate new vocables for each new situation? Simply put, it is not efficient. Efficiency can be casually defined as a measure of the difference between energy input and work output, where energy input is always greater than the result. The difference is wasted effort or loss. The more efficient the system, the less the difference. A perfectly efficient system can use its energy input without any loss. For example, a perfectly efficient engine will use its fuel to propel its vehicle forward without generating heat or friction. This ideal degree of efficiency is not possible in this universe because of the second law of thermodynamics [51], so efficiency is something to maximize not to perfect. For most systems maximizing efficiency is not a matter of conscious choice - we find that self-organizing systems that sustain themselves for a prolonged period of time use strategies that maximize resources and longevity. In this sense, humans and their use of language are no different from most self sustaining natural systems.

If we can see it, taste it, smell it, hear or feel it, then it is part of our readily accessible lexical vocabulary. We can imagine that its sensorial richness co-opts many neural resources and also its perceptually bound nature limits the amount and kind of work it can do. Because different people have different perceptual experiences and histories, the use of lexical vocabulary does not scale across a diverse population very well. So, while lexical vocabulary is efficient for certain tasks, its efficiency is not very high across a broad range and requires many neural resources. While generating new vocables to *fit* each new situation might seem to increase work output, we must remember that language becomes shared on the basis of common past experiences. The shareability of constant novelties is questionable. Also, the sheer amount of neural energy (effort) that would be required to invent an infinite number of vocables is itself infinite.

If lexical vocabulary on one hand does not scale and if on the other hand novelties are too expensive, how do we come to have a shared language such as English that is spoken by millions of people?

⁸The observation of a change in language version can only be made in the future, when language users are no longer intimately connected to the version in question.

1.4.1 Attenuation in vocables

Ironically, the shareability of language is safeguarded by the erosion of associated perceptual structures. We call this process *attenuation* and the result is attenuated vocables. We define attenuation in theoretical terms as a widening of the extension of the relationship of particular vocables which in turn may lead to *transcategorical changes*. A relationship can be defined in terms of subject/object such as the lexical uses of *have*;

I HAVE a house.

A relationship of *ownership* is established with *have* between the subject *I* and the object *house*. *Transcategorical changes* may lead to a widening that includes functionalization changes in the syntactic role of vocables. The new syntactic role becomes divorced from its early uses. Again *have* is a good example; in its functionalized use

I HAVE gone shopping,

have has become an *auxiliary* and has lost its relationship to its *ownership* uses.

We come to this definition from examining the dynamics and the resulting changes in behavior of specific vocables along their history.

As mentioned, lexical vocabulary does not scale very well as the spontaneous generation of vocable novelties is not efficient. The solution then is to generate partial novelties from our pre-existing vocable stock.

As vocables are extended in their use to include new contexts, there is a successive loss of the early context and its perceptual cues. We suppose that the loss of specificity reinforces certain specific features that are common across larger sets of circumstances. For example, consider the perceptual structures associated with the experience of one specific cat. These associations may involve features characteristic of that cat, such as say, black and white spots, short hair, four legs, small, meows, big eyes, loner, and more. These same perceptual structures may also be rich in associations of features that are not specific to that cat such as the room it is in, the time of the day, smells floating about, and so on.

As the experience of cat is extended to include many more cats, some of the richness of the initial perceptual structures will be shed, such as the specific-to-this-cat features. Some of the specific features of the initial cat, such as short hair, black and white, etc. will also be shed. The resulting cat description may only involve perceptual structures that are much less specific, such as say, four legs, loner and meows. The consequence of this dynamic is to increase the efficient use of synaptic resources. However, as specificity diminishes, the

ease of use for particular vocables is impaired. In our example, the set of cat items that may be included in the category will become increasingly difficult to identify. Do we include lions? They do not meow and they live in a pack. We know that lions do belong to the cat family, but only after biologists have defined more detailed criteria to give us a specialized description of what constitutes a cat.

Cat is a noun that belongs to a set of lexical vocables that become *attenuated* but not to a large extent. The extension of its relationship to biological creatures becomes wide enough to include lions or even to refer to someone as being *catty* or a *cool cat*, but none of these uses strays much from involving cat-like features.

In the case of some other types of vocables - such as our example *have* and much of the relational vocabulary, which include *over, without, as far as, in fact, etc.* . . . and logicalized vocabulary such as *and, or, if, not etc.* - the syntactic structure of their use can change dramatically as the result of *attenuation*. In fact, we suggest that all the examples of changes in language that we have provided so far are the result of a process of *attenuation*. However we have yet to explain the role of this process in the shareability of language.

1.4.2 The dynamics of shareability

At a very basic level, we recognize each other's speech by *tuning* to sounds that are particularly suited for the human *supralaryngeal vocal tract* (SVT) to produce ⁹. Unlike other species our capacity to discriminate particular phonemic streams is very elaborate. Our *speech detectors* seem to be genetically transmitted. Infants respond to basic vowels and stop consonants shortly after birth [41] ¹⁰.

⁹The SVT includes the nasal cavity, oral cavity, the velum, the tongue, the pharynx, the epiglottis, the hyoid, the larynx and the trachea. The particular shape of it in humans allows movements of the tongue that can produce abrupt changes in the *cross-sectional* area of the SVT allowing for vowel sounds such as [i], [u] and [a], and velar consonance such as [g] and [k]. Moreover, the capacity for the velum to seal the nasal cavity allows non-nasal speech [41]. These sounds are not matched in other primates, though we do share with them innate neural functions that are particularly sensitive to the sounds that we produce; this feature is shared by other species such as crickets, frogs and monkeys. Much like these species, we have neural functions that are able to discriminate and categorize automatically into discrete patterns of *formant frequencies* and *phonations*. [p] will belong to one, so will [b] and [a] etc.

¹⁰One of the most important ways by which we access each other's speech patterns is with the oral tuning around the vowel [i]. The non-nasal utterance of the vowel [i] can be recognized across different lengths of the supralaryngeal air way that changes the pitch contour of speech. The vowel [u] is also used in recognizing speech, though not as efficiently as [i]. Both of these vowels are present in all human languages. The key in recognizing speech is in the non-nasal enunciation of vowels. Nasalization increases the error of vowel recognition from 5 to 50 percent [41].

However, sounds do not occur in the absence of other perceptual cues. Consequently, we assume that phonemic streams are entrenched in the experience of other sensory experiences. In most adults, sensory maps are synaptically segregated; an oral stimulus will not activate, say, the color area of the visual cortex. This being said, the angular-gyrus is a crucial area in polymodal convergence of sensory stimuli which provide coherence between the senses in apprehending situations.

Studies in people with a peculiar disorder called *synaesthesia* provide more detailed insight into how, synaptically, we come to associate sounds and shapes, colors and graphemes or even sounds and activities. *Synaesthesia* is a condition in which people report being stimulated in a second or third modality though only receiving stimulus in one. The most common form of *synaesthesia* is the association of graphemes and color, that is, synesthetes will report seeing specific colors when presented with specific letters or numbers. So, for example, one sees the color purple when perceiving the number five. There are also cases of *synaesthesia* that have been reported in which synesthetes see shapes along with some sounds or music.

In *Synaesthesia - A Window Into Perception, Thought and Language*, V.S. Ramachandran and E.M. Hubbard suggest that *synaesthesia* is a hyperconnectivity in the fusiform and angular gyrus; they say it is the result of a mutation that causes defective pruning of connections between sensory maps [55].

We identify different subtypes of number/colour synaesthesia and propose that they are caused by hyperconnectivity between colour and number areas at different stages in processing; lower synesthetes may have cross-wiring (or cross-activation) within the fusiform gyrus, whereas higher synesthetes may have cross-activation in the angular gyrus. This hyperconnectivity might be caused by a genetic mutation that causes defective pruning of connections between brain maps. The mutation may further be expressed selectively (due to transcription factors) in the fusiform or angular gyri, and this may explain the existence of different forms of synaesthesia. If expressed very diffusely, there may be extensive cross-wiring between brain regions that represent abstract concepts, which would explain the link between creativity, metaphor and synaesthesia (and the higher incidence of synaesthesia among artists and poets). -Ramachandran and Hubbard, 2001, p.3-

Ramachandran and Hubbard suggest this hypothesis because the fusiform gyrus hosts both the areas of color (V4 and V8) and the visual grapheme area, most specifically in the left hemisphere adjacent to V4. They also suggest that immature brains are significantly more connected between and within these areas and that a process of pruning eliminates many of these connections. Some suggest that infants are born with synaesthesia and that segregation is achieved by the age of four months[5].

Maybe more interestingly, Ramachandran and Hubbard propose a theory of evolution of language and the emergence of proto-language based on their discoveries on *synaesthesia*.

We suggest, also, that the study of synaesthesia can help us understand the neural basis of metaphor and creativity. Perhaps the same mutation that causes cross-wiring in the fusiform, if expressed very diffusely, can lead to more extensive cross-wiring in their brains. If concepts are represented in brain maps just as percepts are, then cross-activation of brain maps may be the basis for metaphor and this would explain the higher incidence of synaesthesia in artists, poets and novelists (whose brains may be more cross-wired, giving them greater opportunity for metaphors). Our speculations on the neural basis of metaphor also lead us to propose a novel synesthetic theory of the origin of language. We postulate that at least four earlier brain mechanisms were already in place before language evolved; a non-arbitrary synesthetic link between object shapes and sound contours (e.g., *bouba* and *kiki*), a synesthetic mapping between sound contour and motor lip and tongue movements (mediated, perhaps, by the recently discovered mirror neurons system in the ventral premotor area that must represent the movements of others, including vocal movements), a synesthetic correspondence between visual appearance and vocalizations (e.g., *petite*, *teeny* and *little* for diminutive objects mimed synaesthetically by a small /I/ formed by the lips and a small vocal tract), and cross-activation between motor maps concerned with gesticulation and vocalizations. This would have allowed an autocatalytic bootstrapping culminating in the emergence of a vocal proto-language. Once this was in place other selection pressures could kick in to refine it (through the combined effects of symbol manipulation/semantics and of the exaptation provided by the syllabic structure for syntactic deep structure). -Ramachandran and Hubbard, 2001, p.28-[55]

Unfortunately, a theory of the emergence of linguisticity in individuals does not entirely describe how language is shared and how we come to have all this vocabulary. It also does not explain how we recognize grammatical structures from non-grammatical ones or how language changes over time.

If we believe Ramachandran and Hubbard, we may associate sounds and objects or activities similarly across the species but we must assume that there are some differences. Even synesthetes do not necessarily agree on the color for particular graphemes; for example, two synesthetes may not see the same color when exposed to the same grapheme.

1.4.3 Shareability of language in a population

In addition to the tendency to increase efficiency that occurs in the use of vocables in individuals, we propose a model that relies on short range interactions that we refer to as *nearest neighbour interaction* - and associated dispersive dynamics - to explain how differences can be more or less standardized across a population. It also accounts for how this dynamics furthers the process of *attenuation*. We will explain, in this thesis, the specifics of *nearest neighbour interaction dynamics* but for now, let us define it as the influence that linguistic neighbour have on each other, in this case, on their use of language. As we have previously mentioned, the erosion of effects in the use of vocables is the result of a tendency toward efficiency as we reuse specific vocables in association with similar contexts.

The loss of original perceptual cues, that we have mentioned earlier, make vocables susceptible to reinterpretation. In doing so, it augments their portability. As the perceptual cues of early contexts are lost the participants in a linguistic transaction are forced (or have the opportunity) to provide part of their own individual past experience to the new context. This furthers the process of decontextualization even more. What may have been shared originally by a small group of people using a contextually bound lexical (maybe even indexical) set of vocables, may have been reflected synaptically (as Rizzolatti, Fogassi and Gallese suggest). However, as vocables become extended beyond their original context and as original perceptual cues are forgotten, these synaptic resemblances may disappear as well. Hence vocables become *attenuated*.

There must be an aspect of language in which individual associations of sounds, objects and activities are compared and shared. It is not clear how this would happen, but negotiation and coordination between activities also establish common context. This would work

as gradual synchronizing, akin to rhythmic improvisations in human tribes. The tiny adjustments that individuals do to keep up with a neighbour's beat soon give rise to a unified rhythm. This kind of synchronization is common in nature. The simultaneous firing of fireflies is one example and the synchronization of women's menstrual periods over prolonged contact is another [21].

There is an enormous leap between tribal rhythmic expressions and the kind of coordinating we achieve purely syntactically. However, we do establish common context by *explaining* to each other *what we mean*. The presence of jargons in different areas of work is a good example of how a negotiation and a coordination of activities can lead to specialized vocables and novelties. The *Internet revolution* has provided us with many such novelties; *to go on-line, e-mail, google search*, have all entered our daily speech.

We assume that all novelties arise out of a pre-existing base. Each new novelty *hooks* onto a familiar syntactic vocabulary. Thus, our *individual* experience of vocables and all they entail is in a sense compromised when they are shared. When we extend our use of particular vocables to other contexts we are actually involved in a linguistic transaction with another or a group of others in which we all bring forth our linguistic specificity in a *quoting* [33] of previous experience. A kind of competition may occur between individuals' own account of the uses of vocables. The competition resolves itself with the group's perceived overlap of common features. A crude account of *neighbour interaction* that leads to common context can be found in the following:

- *I've been "watching the paint peel",*
- *Do you mean you've "been passing the day"?*
- *yes, I've been "wasting time",*
- *so, time has been "passing" and you have been "inactive"?,*
- *yes, that's what I said.*

This example uses cultural metaphors but it may just be that cultural metaphors are the product of quoting past experiences. There are thousands of these idioms in the language: *gone the way of the dodo, burning the midnight oil, nose to the grindstone*. All of these are supposed to conjure up vivid imagery that a majority of the population can relate to, but eventually the events that are described by metaphors will not be shared. Generationally,

it is doubtful that cultural metaphors can survive very long. What child today knows what *midnight oil* refers to? It is also doubtful whether cultural metaphors can be extended to other ethnicities. This is evident from the amusing, and not-so-amusing, problems that E.S.L. speakers have before they've learned English idioms.

The process of *negotiation* through *neighbour interaction* results in a furthering of the process of *attenuation* in idiosyncratic uses of vocables in individuals. This in turn, augments the portability of vocables. The requirements of efficiency and the propagation of vocables within a population lead to morphological changes in the use of these vocables. They tend to become shorter. There are many examples of this in English: *or*, for example has come from *other*; *but* from Old English *butan*; *have* and *will* in their auxiliary role take the form *'ve* and *'ll* and *must've* becomes *must of* in popular speech.

We suggest that the process of attenuation also erodes other aspects of speech such as morphology. It is reasonable to view the vocabulary of present languages as the result of the erosion of much longer strings of phonemes and morphemes that were culturally shared in previous times. It is not difficult to imagine the slow process of erosion of specific sounds over millions of years in the utterance patterns of hominids and how proto-language resulted from this. The next step involves offering a description of the dynamics that gave rise to modern languages, that is, how grammatical forms emerged. So far, we have confined ourselves to a conversational account of the dynamics that occur with language evolution. In order to describe the event of grammatical forms, we will need to invoke the idiom of statistical mechanics. We also use the event of functionalized vocabulary in present English language as a representative sample of the larger scale dynamics that gave rise to natural languages.

1.4.4 Transcategorical changes

The kinds of changes that we are tracking are the kinds of changes that we have described with our example *have*. We showed how its lexical use of possession, in some instances, is shed, to give rise to a functional auxiliary role, as in *I've gone shopping*. This is considered a *transcategorical change*. Two more such changes are the adverbialization of verbs (*I have gone RUNNING*) and the prepositional use of adjectives (*The shop is DOWN the street*). The *functionalization* of vocables occurs mostly with *relational* vocabulary, such as *even if, without, while, over, and so on*. Their lexical ancestors often bore spatial or temporal uses. These are usually shed by their functional counterparts. There are a few exceptions,

however, such as *in fact* whose ancestor is a noun.

In his paper, D.K. Johnston [34] traces the natural history of *in fact* through its first uses as a noun. It started out resembling the use of *in action* as *gracious in fact, not in word*[2]. It is now used as *based on observed events* as in *the fact that this happened...* It is also used functionally, as in *In fact, the chance of recovering the stolen items is unlikely*. Despite its original use and role as a noun, *in fact* has been extended to adverbial uses and is rarely used as it originally was in its early uses. This is an uncommon example: most nouns do not develop functional uses. However, it demonstrates that there are no *type* of vocable immune to functionalization given proper circumstances, circumstances that we will later describe.

We have said that a common context can emerge from a *neighbour interaction* negotiation. We have also said that individual metaphors are in a form of competition that resolves itself with a perceived overlap of features. Now imagine that such a transaction is successively repeated over time by every individual in a population. As vocables are negotiated, the specific effects associated with vocables will become eroded because the *empathy effects*, of the kinds reported by Rizzolatti, triggered in different individuals will increasingly become difficult to correlate between them. The increasing discrepancy between the individual use of vocables leads to a loss of common features which in turn leads to *attenuation*.

One consequence of attenuation is well illustrated by Jennings example of *scope misapprehension* [48]. These are the mistakes that go unnoticed in the use of vocables that eventually allow them to take on new syntactic roles. We may ask: What is the advantage in the functionalization of lexical vocabulary or in the grammaticalization of linguistic forms? It may be related to Bichakjian's claim of efficiency. Like word order, functionalization allows for flexibility and complexity of expression. We can imagine that, as human activities require more precise coordination, linguisticity will also have to retain or augment a certain level of precision across an increasingly large population that shares it.

The fact is, there are vocables in the language that are semantically difficult and their use is not lexical. *And, or* and *but* are such vocables. As Jennings explains, it is not entirely clear what the semantic role of *logicalized* vocabulary is - vocabulary such as *or* defined truth functionally - but its role seems purely syntactic. We also know that a diachronic account of these vocables shows that functional vocabulary descends from lexical vocabulary and that, in the course of its history, some lexical vocabulary will undergo a transcategorial change of the functionalized kind.

Our theory suggests that there is a critical point at which attenuated lexical vocabulary will no longer generate specific effects. It has become too attenuated. A consequence is the disappearance of this vocabulary. If it has no specific effects, why use it? Another is a new syntactic role, accompanied by a morphological change.

1.5 Phase Transition in the Evolution of Language

The sudden emergence of features in physical phenomena is spectacular but not a rare occurrence in Nature. The spontaneous magnetization in ferrous materials and the spontaneous structural change of a liquid to a solid state are well documented phenomena and mathematically described facts. An everyday example might be H_2O and its clearly distinguishable states as ice, water, and vapor. In the general parlance of phase, H_2O is understood to exhibit characteristics associated with different phases, and moreover, to have identifiable transitions between them. An important feature, that is well-defined mathematically, is the point at which one state of a system changes to another - this is referred to as the *critical point*. For example, as the temperature of a system of liquid water is decreased, the state will spontaneously change from a liquid with no structure into a crystalline configuration - ice - at the critical temperature, 0 Celsius. These apparently spontaneous changes are transformations that are well illustrated by the physical concept of *phase transition*.

A phase transition describes how a system moves from one phase to another, relative to a *state variable* that is unique to the type of system under observation. In the case of the water/ice transition, the state variable is a coherence length - a measure of the strength of a defined set of relationships between the objects under observation [51] - that reflects the presence of inter-molecular structure. There are several types of phase transitions, chief amongst them are the *first-* and *second-order* transitions. A first-order phase transition demonstrates a discontinuity in its state variable at the critical point - there is an abrupt change from one state to another. A second-order phase transition describes a smooth variation in the state variable. Water/ice follows a first-order phase transition.

First-order phase transitions is a well developed area in physics. Statistical mechanics has developed approaches that give a detailed *microscopic level* description of interactions amongst a large number of individual constituents, say H_2O molecules, and their repercussion at a *macroscopic level* such as the structural change in a system, say, from a liquid configuration (water) to a solid configuration (ice). Such a system can display a phase

transition if some intrinsic macroscopic quantity variable is varied. This variable - temperature in the case of H_2O example - is defined at a macro level. Elevating or decreasing the temperature corresponds to a comparable variation in the mean energy in the (equilibrium) distribution of energies. For example, a phase change from liquid to solid for H_2O involves a decrease in temperature. This is equivalent to a decrease in the average energy available to each water molecule.

The modeling of phase transitions involves the definition of the ensemble of energy states of the molecules and the macroscopic consequences for the entire system. Any related dynamical behaviours are assumed to take place on very long time scales so that the state of the system can always be viewed in terms of equilibrium or at least quasi-equilibrium. Any dynamics are then accounted for as slow, independent processes. We will describe more specifically how the model is implemented in the section *The Ising Model*.

We suggest that the dynamics of linguistic transcategorial changes leading to functionalization can be successfully compared to the processes that leads to a structural change in, say, freezing water. We suggest that these changes occur spontaneously. In this suggestion we are borrowing from physical models, used in physics, to formalize our language-related claims. In our model, transcategorial changes constitute the macro feature of the language system and vocables are the micro constituents.

What we model are the interactions that lead to attenuation and critical structural changes in vocables in a population of language users. The *macro* variable is the *activity* that exists in the system. In other words, it expresses how linguistically interactive the overall population is, both in propensity and in number of individuals. The concept of activity is a variable and is similar, in role, to the temperature in our H_2O example. The *micro* feature that is monitored in our model is the capacity to generate *effects* in a population of language users. The capacity to generate effects is akin, in role, to the energy available to the physical system of H_2O . However, though these effects are observables of a population of users we presume that, ultimately, they are of a neural kind and can be measured as physical energy.

1.6 The Ising model

To explore the phenomenon of transcategorial changes and functionalization, we use a model in which a finite population of language users exchange *meta* vocables¹¹ with their immediate neighbours. An increase in activity in the exchange of these meta vocables determines whether a transcategorial change will occur. In our model the meta vocables are either in a lexical state or in a functionalized state.

There are many ways to simulate the statistical nature of a system, including phase transitions. Since we are in uncharted territory where the simulation of phase transition in language is concerned, we will make as few assumptions as possible in the modeling process. Statistical mechanics provides an appropriate model that is simple, well-known and mathematically tractable. It is called the *Ising model*[51].

The basic two-dimensional *Ising model* demonstrates a second-order phase transition. As we have explained, a second-order phase transition is characterized by a smooth change from one state to another. With a slight modification, (adding a *degeneracy* feature), it demonstrates a first-order phase transition. Using this version of the Ising model, we show that the nature of transcategorial change between a lexical state and a functional state can be described by a *first-order phase transition*.

The Ising model consists of a number of individual constituents, usually referred to as *spins*. Spins can be understood as simple vectors with discrete orientations, either up and down, but computationally can be modeled as binary states, either 0 or 1. This simplicity is what makes the Ising model so attractive as a model. A first-order phase transition is implemented by introducing two new features to the system that are mathematically simple. One is implemented macroscopically as a *bias field* that causes one state to be energetically favoured over the other - that is, one of the two states is lower in energy than the other. The second is implemented as a *degeneracy* in the spin states. This requires further explication.

Physics informally defines *degeneracy* as a loss of feature. We have mentioned that constituents in our model can be in one of two states. In simple terms, if each state is treated similarly (bias field aside), then each state has the same statistical likelihood of being present - in the case of two states, that is 50%. If we now ascribe three possible states for constituents rather than two and each is equally likely, then the probability is reduced to 33% for each given state. But if two of these states are effectively identical and

¹¹by *meta* we mean a representative for vocables of a type rather than a specific vocable or utterance.

indistinguishable as far as the model is concerned, then a bias of 66% (33% each for the two identical states) is generated in the state potential. Though the underlying statistics do not change, two of these states are interchangeable. Such identical states are described as being degenerate because of their lack of discriminating features. In the case of our language theory, the preferred (or biased) state of vocables is the lexical and the degenerate state is the functional. We will elaborate later on how vocables exhibit degeneracy.

In a physical system such as a system of H_2O molecules, these biases will result in preferred states such as solid/liquid/vapor depending on the circumstances. H_2O molecules will tend to a solid state when there is little energy available to the system while they will tend to a liquid state when more energy is available to it, and finally a vapour state at even higher energy levels. Each state has a structural coherence associated with it that represents an energetically preferred configuration (e.g. crystalline for the solid state) and plays the role of a biasing force. At certain system energies (i.e. temperature), the H_2O system tends to a particular structure. Within the Ising model, the energy is specific to the temperature of the system and can be controlled explicitly by tuning the system relative to a critical point where the two biases balance. By then coupling the spins energetically to their neighbours via a *nearest neighbour interaction*, the system is thus equipped to demonstrate a first-order phase transition.

As mentioned, these descriptions rely on the assumption that the entire system is at least at quasi-equilibrium (not changing rapidly). The rate at which the constituents interact must be much shorter than the rate at which any changes take place. The time scales involved with structural changes in vocables range from a few generations in the cases of functionalization or in the emergence of creole languages, to a hundred years or more for the emergence of the natural language phenomenon. Compare this with the timescale for vocable use, something on the order of seconds or minutes. Clearly the applicability of the Ising model in this regard is strong.

Our Ising model hosts meta-vocables as constituents. The nearest neighbour interaction in our model is a simplified version of what may occur in a real population, as individuals influence each other's use of vocables within linguistic transactions. The *activity* (comparable to energy) variable in the system regulates the flow of exchange while the *capacity to generate effects* (comparable to degeneracy) regulates the individual influence of language users on each other. The activity in the system is an abstract concept and can be understood mainly as the level of exchange but also as propagation through a finite population.

As the level of activity increases in the system, it is more likely that functionalization will occur. While the nature of the activity assumes that the exchange of specific vocables is homogenous, we think that in reality some vocables will be in use more often than others.

We might be tempted to assume that some vocables require less use over time to become functionalized because of what they *mean* or because of their initial syntactic role. But we doubt the validity of this assumption. For example, *have* will become functionalized while *running* will not despite the fact that they are both verbs. The difference lies in the frequency of use in linguistic interaction. The attenuation of *run* beyond *running* in say, *running a fever*, is not likely to happen because the occurrence of the particular *run* vocable in linguistic transactions is not as frequent as the occurrences of *have*. To compensate for this over simplification, we have implemented a feature, comparable in role to *potential degenerate states* - which we call *potential attenuated states*. In a standard Ising model the number of potential degenerate states is equal for all constituents. This feature illustrates the idea of *attenuation* in vocables. Since, in our model, we lack a certain control over the level of activity that each meta vocable may be engaged in, we have implemented various potential attenuation states. These illustrate that vocables are individually potentially more or less susceptible to functionalization.

With this experiment, we aim to demonstrate that it is the number of instances in the lifespan of a particular vocable that determines structural changes rather than its semantic or syntactic type. Moreover, we will demonstrate that the use of the simplest model that describes first-order phase transition - the Ising model - is relevant in the explanation of such a phenomenon, though somewhat limited. More sophisticated models will subsequently be explored in which phase transitions will also be considered as altering phenomena. But more importantly, if the process of functionalization is scalable to the larger phenomenon of grammaticalization, it would seem that, as an important structural change, grammaticalization can emerge spontaneously from proto-language, given a certain level of proto-language use across a critical number of individuals. Of course, the time scale of what we consider “spontaneous” in the evolution of language remains to be clearly established. We will also point out that these findings are consistent with Bickerton’s in suggesting that there are no transitional steps in the emergence of a natural language from a proto-language.

1.6.1 Calvin and competition

Despite the simplicity of the Ising model, we can use it to demonstrate the relevance of certain features responsible for changes in language, such as the propagation and the level of linguistic interaction and attenuation. However, there are several features that we have ignored in this model, such as long-range neighbour interactions and the aspect of competing uses of vocables that we described earlier (*The dynamics of shareability*). We have sketched a dynamic in which a population of users are engaged in linguistic transactions. Each brings forth a unique use of vocables, which are then modified by a process comparable to a competition between individual effects generated from perceptual cues. The effects are, we assume, of a neural kind. However our population dynamics account is too coarse-grained to give a description of what they may be. We suggest that the dynamics of competition at a population level may be a reflection of what goes on at the synaptic level.

William H. Calvin has created a model that offers a detailed description of what neural effects may be. He describes a context of competing synaptic patterns that are responsible for our capacity to abstract including our capacity to schematize and possibly for our use of language. Calvin's refers to his model as a *Darwinian machine*. His inspiration is based on the structure of the superficial layers of the neocortex and their behavior. Once stimulated, the neural arrangement in these areas, generates synaptic patterns that are subsequently replicated across pairs of neurons that consequently form a whole structure. Calvin suggests that the replication process is not unlike putting down tiles on a floor as we try and fit the edges of the tiles to cover the surface as well as possible. Several features of the tile, such as its shape and color, perhaps its thickness, will define the overall pattern of the flooring. This metaphor illustrates how a perceptual structure occurs in synaptic functions. Calvin's model is not unlike the Ising model, except for its capacity for long range interactions (as opposed to nearest neighbour interactions) and a few other less crucial features.

In Calvin, linguistic activities are perceptual structures, that is, neural activations that occur from our interaction with our environment. These neural activations generate patterns. For Calvin, a complete pattern requires three pairs of neurons configured as a hexagon. The strength of activation of this pattern may activate an entire structure of many hexagons. These structure may be in competition with each other in a context of partial or degraded perceptual cues. The competition is usually resolved, often with inaccuracies that can give rise to several of our thinking processes such as generalized schemas.

The idiom of physics is not a large part of the elaboration of Calvin's model, but he does explain how *attractors* bias and promote the propagation of spatio-temporal neural structures. An *attractor* can be understood as a constituent locked temporarily in a particular state. Constituents locked in that state can potentially entrain neighbours into a similar state. This locked state can be a periodic vibration of sorts, a resonance. We have mentioned an example of entrainment from *resonance* with the synchronization in the firing of fireflies and the synchronization of women's periods. *Resonance* is what causes neurons to fire rhythmically for limited cycles. There are many types of periodicity in *resonance patterns*. Some can be compared to a simple melody while other are more akin to white noise.

For Calvin, multiple attractors can be in competition with each other for the propagation of their own pattern. A resolution is achieved through a *first-order phase transition* the result of which can be a merging of attractors. For example, the experience of the object apple will generate an attractor and so will the syntactic experience of the word. The merging of both *attractors* will promote the propagation of the specific neural pattern of combined perceptual and linguistic experience.

We are interested in Calvin's account because it offers a finer grained description of linguistic interactions, one that is compatible, and even complements, our more coarse-grained physical, population dynamical, account of linguistic interactions. Calvin's model is a good illustration of how a fine grained description of language does not have to rely on a semantic theory and how it can be compatible with a wider family of physical theories including physics, biology and neuroscience.

1.6.2 Integration of Calvin and Ising

We will demonstrate, through the use of the Ising model, that a *phase transition* can correspond to the propagation of transcategorial changes. Though we have limited our inquiry to the emergence of functional vocabulary in a population of users, we assume that transcategorial changes are reflected in the dynamics of synaptic structures as well.

Percolation is of particular interest because it is a process in which stable structures can emerge in a low *connectivity* environment. Consider an array whose sites can either be occupied or not. This array may be one or several dimensions. Sites may be occupied at random or according to some rules. One of which can be that there are neighbouring occupied sites. Percolation theory deals with the properties of clusters of occupied sites that

form in this lattice. The word *percolation* reflects the possibility of having a cluster spreading over the whole lattice. Percolation usually demonstrates second-order phase transition in state as defined by the connectivity between sites. This state change is usually the result of a critical number of connections between clusters.

This is a relevant concept in optimizing resources.

For Calvin, most neural structure is the result of spatio-temporal cues which provide many stimuli for a high number of neurons to be entrained and connected in a specific synaptic pattern. We assume that this requires much synaptic space. But the question is, can neurons be entrained into patterns without the benefit of spatio-temporal cues, that is, in times in which connectivity is low and the possibility of entrainment seems unlikely?

A percolation problem is well illustrated by the 60's game *Kerplunk!*. In this game, marbles are suspended in a tube by many toothpicks running across the diameter of the tube. As the toothpicks are removed successively, the marbles may shift slightly. Eventually only one crucial point of support will remain. Once the crucial last toothpick is removed, marbles will start falling until the tube is empty. Marbles will follow, one after the other, a path traced from the architectural constraints imposed by the removal of the crucial toothpick.

Percolation is a *second order phase transition*. The point of our Kerplunk! example is to show how percolation can entrain a system in a low connectivity environment. Imagine that the toothpicks are attractors of sorts that are too weak individually in a high connectivity environment - many toothpicks - to influence the rate of firing of neurons around them (to direct the orientation and path of the marbles). In a low connectivity environment (few toothpicks), these attractors will be strong enough to entrain neurons around them, and those neurons will entrain the ones around them and so on. Soon enough a new structure will emerge. Now, because of the low connectivity, we assume that these structures may be stripped of many features that direct perceptual cues would usually provide. We suggest that some schematization is born from low connectivity. These are exploited in a strategy to maximize neural resources. This may apply to many kinds of abstractions especially those that do not seem to have any ties to spatio-temporal cues such as the grammatical features of language.

1.7 Description of Chapters

Chapter One is the introduction. In it, *language* is defined as a physical phenomenon. As such we define the boundaries of this phenomenon by using a classification of important features much as biology classifies the features of living organisms using categories such as species, kingdom, family. These categories are then set within both a proper spatio-temporal frame and the specific scales at which these categories can be established.

Chapter Two reviews the work of R.E. Jennings. His diachronic account of functionalized vocabulary offers a theory that a semantic theory is not to be had for most of the vocabulary and, in particular, for logicalized vocables.

With Chapter Three, we present the concept of efficiency as the driving factor for the evolution of language. We describe the concept of efficiency through the metaphor of tool making. Many other fields also consider efficiency a fundamental feature that drives language evolution: neuro-psychology, neuro-biology, linguistics etc. We give a brief account of their views.

So far, our discussion relates language to physical systems in a conversational way. In Chapter four, we turn to a better defined physics idiom to understand in more detail the formal implications of our claim. We explore the definition of a system, the concept of propagation and dispersion, elasticity, and degeneracy and phase transitions, on our way to defining a formally tractable model for the evolution of language.

In Chapter Five, we discuss several models that are used in non-traditional ways, especially in the description of language dynamics. We first establish the validity of our approach by comparing how other fields such as biology, neuroscience, artificial intelligence and economics have used similar thinking. We examine how biology uses models of phase transition such as percolation to explain dramatic changes in genetic diversity. Neuroscience uses models that illustrate the behavior of systems on the edge of chaos to explain how we *change our mind*. This leads to similar kinds of approaches in artificial intelligence. More specifically, the use of the Hopfield model, whose state space is identical to an Ising model: It models the behavior of associative memories. Economics has used Ising models to describe trends in the propagation of information in a market place and to describe structural changes in multi-factor systems.

Chapter Six introduces the Ising model, its mathematics, and how we use it. The Ising model has traditionally been used in the illustration of structural changes in material

such as from solid to gas. We have modified the Ising model to illustrate more accurately the process of attenuation in language. We use the Ising model to track the dynamics of structural changes between lexical and functional vocabulary. We demonstrate how the dynamics of language evolution can be illustrated using a physical model, to offer a more formally tractable description.

Chapter Seven goes beyond Ising to introduce the Darwinian machine of William H. Calvin. His model is akin to the Ising model, though more complex. He proposes a model of the superficial layers of the neocortex and describes the relationships to linguistic and representational concepts. We then introduce our model in the context of Calvin's. We illustrate how the concepts of physics that we have introduced, such as phase transition, can provide additional detail and establish compatibility between our approaches.

Since Calvin does not emphasize the notion of *phase transition* in his model, we propose, in Chapter Eight, ways in which his model can be enhanced. We suggest that a phase transition such as *percolation* can entrain patterns in an environment that is low in connectivity as a strategy to maximize resources.

Chapter Nine concludes with an integration of all the previous ideas that we have described to offer a complete picture of how we can consider all aspects of language in a wholly physical idiom.

We hope that this thesis will help answer some of the fundamental questions associated with the study of human language.

Chapter 2

Loss of Meaning

2.1 Evolution of language

The inspiration for my research comes from R.E. Jennings' research on the evolution of functional vocabulary. His interest stems from the fact that semantic theory does not offer a satisfying explanation for the occurrence of this kind of language. Jennings has come to this conclusion from studying the dynamics of functional vocabulary in the English language. One important discovery, in his research, is that functional vocabulary is particularly vulnerable to *misunderstanding* that leads to *misuse* which, in turn, gives rise to *repurposed* uses.

Jennings' research describes patterns in the dynamics of the evolution of functional vocabulary that, we find, are existent in the evolution of all vocabulary.

We describe with some detail Jennings' motivation for his research and the diachronic history of functional vocabulary [47].¹

2.1.1 Logic and understanding

Philosophers of language often assume that our understanding of language resides in our having an implicit semantic theory that enables us to compose sentences whose meanings are constructed out of the meanings of their component words. To account for the fact that we infer more from one another's utterances than is strictly said, consider these two

¹The following is taken from "Language and Intelligent Understanding without Semantic Theory" that I published with R.E. Jennings in the "Canadian Artificial Intelligence" (CAI) magazine in Autumn 2000.

examples:

Our lecturer was sober today or

The department chair has not yet been sent to prison.

They invoke a secondary device.

This device, also implicitly understood by us, is called implicature [26], and is based upon rules that are supposed to govern what we say and when we say it. In the former cited example, the words *sober* and *today* retain their fixed meanings; the fact that we infer that sobriety is not her usual state is accounted for by maxims of conversational propriety, not by some variability in the meanings of the sentence-elements. In particular, the so-called *logical* words *and*, *or*, *not*, are supposed to have fixed meanings that can be specified in truth-tables. These set out the conditions under which sentences containing them are true or false. In fact, the logical vocabulary of natural language has long been supposed to provide a sort of truth-conditional bedrock upon which a full semantic theory can eventually be built.

Even introductory logic texts say things that are in direct contradiction to the intuitive notion that words have single meanings, at least about the word *or*, even those written by eminent logicians, such as Tarski [59]. He says:

The word *or* in everyday language, possesses at least two different meanings.

-Tarski, 1941, p.21-

This is an exclusive/inclusive distinction:

In the so-called non-exclusive sense, the disjunction of two sentences is true if at least one of the sentences is true... When people use *or* in the exclusive sense to combine two sentences, they are asserting that one of the sentences is true and the other false. -Suppes[58], 1957, p.5-6-

Almost invariably, the logic texts that make this point go on to claim that one of the meanings of *or* coincides with that of the familiar *xor* function. If we are all supposed to be possessed of a semantic theory, this is a very curious fact, for a string of sentences composed with *xor* will be true if and only if an odd number of its component sentences are true. So a sentence of the form $A \text{ xor } B \text{ xor } C$ will be true if exactly one of its component sentences is true, but it will be true also if they all are. So it seems that our semantic theory is

inconsistent. We can test our own semantic theory on the following sentence, adapted from a common type of example in the textbooks. Consider:

You can have soup or you can have juice.

Is the quoted sentence an inclusive or an exclusive disjunction? Most undergraduate logic students, asked this question, will respond that it is an exclusive disjunction, and most textbook authors will either agree or will argue that since it would not be false if you were allowed to have both, it must be an inclusive disjunction. Which is correct? In fact neither is, since the sentence is not a disjunction at all. If a waiter said this to you, you would be correct in inferring from what he said that you could have soup; you would also be correct in inferring that you could have juice. It cannot, therefore, be a disjunction, since from a disjunction neither disjunct can be correctly inferred. It must in fact be a kind of conjunction. If, as the customer might assume, it is taken to exclude one or the other starters, then it is being taken as the conjunction:

You can have soup; you can have juice; you cannot have both soup and juice.

We could add numerous independent uses to the list of uses of the word *or*, none of which is adequately represented by the disjunctive truth-functor. These conflicting prejudices - that *or* has just one meaning, that *or* has two meanings, and so on - and the confidence about our possession of an accessible semantic theory appear to have become prevalent only since the invention of the truth-table in the twentieth century. Earlier logical theorists were not so confident of their understanding. Venn [60], famous eponym of the diagrams, himself confessed bewilderment at "the laxity, the combined redundancy and deficiency, of our common vocabulary (*and* and *or*)". Our use of so-called "logical language" does not in general rest upon any underlying semantic theory, accessible or inaccessible, and our various uses of the word *or* do not rest upon an implicit understanding of truth-functions. Given that this is so, any explicitly formulated semantic theory must be regarded as suspect. The fact is, a semantic theory is not required in the transmission of language. Our daily use of language is not driven by an underlying semantic theory. Indeed, children use language long before they are introduced to any sort of linguistic theory and are able to use it adequately. In fact, the transmission of language requires very little theoretical understanding. If it did, languages would not change beyond recognition within so short a span as a thousand years.

But, we may ask, if the transmission of language does not require much understanding, even of the logical vocabulary, how does a language come to have any logical vocabulary at all? It is partly because so little understanding is necessary for the transmission of language

that languages acquire the vocabulary that we think of as logical- likewise, the vocabulary that we think of as psychological, or ethical, or religious. In short, for certain kinds of vocabulary, we may say that in using it, we literally do not know what we are talking about. Its use does not require that we do. How is this possible?

2.1.2 The process of delexicalization and functionalization

The answer is that all such vocabulary descends from vocabulary that, in its more primitive uses - lexical - would have been capable of relatively straightforward dictionary-style definition or ostensive demonstration. Descendent vocabulary sheds its lexical connections, and sufficiently late descendents may be incapable of being understood at all, except through an historical explanation of their descent. The linguistic uses of a language user of one generation are in part engendered by the linguistic uses of previous generations. But in part, as in biological evolution, vocabulary pre-adapted to one role may be co-opted or exploited in another. ²

All, except the earliest, linguistic practices are lacking an accessible semantic account, with vocabulary denoting simple, sensorially immediate items - objects, physical relationships, actions - vocabulary whose use can be conveyed directly by ostension and simple definition, and ultimately, vocabulary that is intimately tied to specific perceptual cues. Such relatively simple vocables are the ancestors of all of the semantically difficult vocabulary of the later stages of a language. All connectives, for example, evolve by various describable stages of logicalization from the sensorially rich, specific vocabulary of physical relationships between individuals, to the sensorially poor, with extremely versatile uses linking whole sentences. For example, the Modern English word *but* is the descendant of Anglo-Saxon *butan* (by outan, i.e., outside); *or* is the descendent of the comparative *other* (second, as in *every other day*), and so on. We can now say in some detail how the transformations come about, and corresponding stories can be sought for all of the semantically challenged vocabulary of folk psychology, ethics and religion. At every stage of linguistic history the process of logicalization, and more generally, *delexicalization*, is in progress. The semantic childhood simplicities of today will engender tomorrow's philosophically adult difficulties. Nonetheless, small children continue to acquire language ³ and manage linguistic

²Ultimately all linguistic practices must be traceable to non-linguistic practices through such exploitation of incidental causal features of pre- and proto-linguistic structures.

³notice that we do not say *the language*: they do not acquire ours, but their own.

intercourse with their parents and grandparents. However, by slow degrees, what was the simple vocabulary of childhood in earlier generations passes into less simple linguistic roles within the adult language of later generations. At each stage there is a balance, though in each case a different balance, between the semantically rich and the semantically poor. As applied to the corresponding elements of human language, Immanuel Kants [36] remark is born out by the facts: percepts without concepts are blind; concepts without percepts are empty. Language maintains a dynamic balance between what we must directly understand and what we need not understand in order to participate in its practices.

We are not denying that, in the ordinary way, any competent speaker of a language understands that language; there is such a thing as conversational understanding. A master of a language, a good novelist, say, has this sort of understanding to a very high degree. But conversational understanding does not confer any other sort, and depends upon something approaching a semantic understanding only for a portion of the material of speech. It is not the sort of understanding that we strive for in mathematics or physics, biology or history. To put the matter bluntly, much of the understanding exhibited in human conversation is a simulation of understanding. Moreover, it is a sufficiently good simulation to have sent many generations of philosophers haring after semantic theories, even long before truth-tables conjured this late illusion of success.

2.1.3 Scope misapprehension

As we have explained, there is an assumption that the semantic theory of the logical portion of language is now nearly complete and will form the foundation upon which some larger semantical structure will be built. However, having taken no account of linguistic change, philosophers of language have ordered their building materials the wrong way round.⁴

Connectives are not at all suitable for a lasting foundation. Historically, the connective vocabulary has shown itself to be extremely fragile, while under similar conversational pressures, lexical vocabulary has proved itself comparatively resilient. Although compositionality - the syntactic composition of sentences - is widely claimed as a key condition, both of our capacity for novel speech production and of our capacity for speech comprehension, this condition represents only a synchronic point of view. Language as a phenomenon must

⁴ideas and examples in this subsection are from Jennings' *The Semantic Illusion* paper, to be published in A. Irvine and K. Peacock. *Mistakes of Reason*. forthcoming.

also be viewed diachronically. When we look at language in its temporal dimension, we see matters differently. First, for all our compositional potential, we actually compose very few sentences each day, either in production or apprehension of speech. This actual composition of speech is the main vehicle of language change. But the compositional forces that can bring about compositionally significant changes of uses in functionalized vocabulary have less dramatic effects upon lexical vocabulary. This particular dynamic is well illustrated in Jane Austen [2].

Have you had any letter from Bath? [Henry Tilney to Catherine]

No, and I am very much surprised. Isabella promised so faithfully to write directly.

Promised so faithfully! A faithful promise! That puzzles me. I have heard of a faithful performance. But a faithful promise - the fidelity of promising! It is a power little worth knowing however, since it can pain and deceive you. -Jane Austen, [1] p.209-10-

The construction arises from a misconstrual of scope: in the word sequence, *promise faithfully to α* , the adverb *faithfully*, in the ancestral uses, modifies the infinitive. The *promise*, on the ancestral construal, is *the promise faithfully to write*. The other construal - reserved by Austen for the likes of the naive Catherine and the feckless Lydia -, which takes *faithfully* to modify the main verb, permits these intransitive constructions, and must therefore give to the element *faithfully* either a new meaning, or no meaning at all. Of course the language finds an idiomatic use for the construction as a whole, a use which is suggestive of earnest, hand-on-heart asseverations and undertakings, but it is not one that relies upon composition of autonomous meanings. The word *faithfully* has never migrated with any such meaning to other environments.

Negation-raising verbs such as *believe*, *think* and so on present a similar phenomenon. The common construal of *I dont believe that α* used instead of *I believe that not- α* has not spawned a new meaning of the verb believe; in biological terms, idiomatic uses do not generally propagate except artificially. A non-English speaker who says *I dont hope you slip on the ice* ⁵ by analogy with the negation-raising idiom, is more likely to be quoted than to be imitated. We mention, in passing, that the verb *doubt* may be an exception to

⁵The example is owed to Charles Travis.

this general claim. There is good reason to suppose that the present use of the verb is a mutation of an earlier weaker use, one that is exemplified frequently, for example, in Pepys Diary [61]

There I found as I doubted Mr. Pembleton with my wife - 1663 05 06

occurrence in KJV. It may well be this earlier weaker use that persists in such constructions as

I do not doubt but that the Viet Cong will be defeated. - Richard Nixon

Functional vocabulary is less impervious to the effects of such scope misconstruals. Of the many instances we now know about, we will mention only two here. The first involves the word *unless*, which is a reduced form of a longer construction; *on [a condition] less (than that)*. In its earliest inter-clausal uses, such a construction would be conjunctive in character. α *on a condition less than that* β would be representable roughly as $\alpha \wedge \neg \beta$. But now suppose (as seems to have been the case) that the construction is never used outside the scope of some prefixed negating item. Schematically, we can represent this as *Not α unless β* (the underlining representing the original *and...not* reading. On that reading of *unless* the sentence as a whole will have its present-day reading if the scope of the *Not* is taken to be as in *Not (α unless β)*. Suppose that an emerging portion of the linguistic population agrees, with its complementary portion, agrees on the occasions of use of all such sentences, but takes the scope arrangements as *(Not α) unless β* . That emerging portion must give to the *unless* element of such sentences a new construal, as it would do in the case of *or*, for purposes of new compositions. Since the two portions of the population are unaware of the ancestor of *unless* unaccompanied by the preceding negation, the difference in their syntactic construals will never become apparent, and therefore the new construal never corrected. But under such conditions, a natural bias in favour of short-scope construals of negatives or simpler syntax more generally will eventually tip the balance of construals in favour of the innovation and, if *unless* migrates to other un-negated environments demanding an *or* construal, a sufficient portion of the population of language-users will have already accepted the *or* construal of *unless* ensuring its survival. Since the *or* reading will do for all instances, both the original (on the the new syntactic construal) and the new un-negated one, the *and...not* construal, is eventually extinguished. We have expressed this in the language of construals, but the phenomenon clearly has a neural substrate involving some form of imperfect replication of structure.

One element of the apprehension of speech (or written text) involves the neural rehearsal

of the motor sequencing involved in its production [53]. We assume that the rehearsal of a spoken sequence under the auspices of one syntactic scheme is different from the rehearsal of the same spoken sequence under the auspices of another.

In the case of functionalized vocabulary, the principle governing changes seems to be occasions of identical instances, accompanied by novel syntax that give rise to entirely novel uses. For lexical vocabulary, such novel construals produce idiomatic constructions, but seldom new independent uses. We take as a requirement that an independent use is stable through some range of distinct environments. We may speak of a *meaning* restricted to a single environment type, but within the theoretical framework we are presenting, this would just be a redescription of what we ordinarily call an idiomatic use.

2.2 How Physics and Language are Related

In comparing the uses of functional vocabulary and its lexical ancestors, we primarily notice that functional vocabulary is structurally syntactically different from its lexical ancestor. This statement requires a more detailed description but, for now, let us just say that functional vocabulary adopts novel syntactic roles that its lexical counterpart does not, such as the auxiliary role of *have*.

We think that the rearrangement in the use of functional vocabulary is not unlike the rearrangement of constituents in the phase transitions of material constituents such as H_2O molecules in water or ice.

With Jennings' account of *scope misapprehension* we also conclude that the misunderstanding in the use of *functional vocabulary* is a powerful motivator in the shareability of language. As we demonstrate in the *The Physics of Language*, the dynamics that lead to the misunderstanding in the use of functional vocabulary is present for all vocabulary. This feature is also responsible for increasing the flexibility in the use of vocabulary and, in turn, for its propagation across a population. Loosely, we think that this kind of dynamic is similar to an increase in the available energy of a system such as H_2O . The energy available to molecules in water allows them to move about more freely than in its solid state of ice.

In subsequent chapters we describe how the dynamics of language evolution parallels the dynamics of phase transitions in materials.

Chapter 3

Efficiency and Language

3.1 The Swiss Army Knife

Tool making can provide a metaphor of an evolutionary account needed in the description of the dynamics of language evolution. Without adopting the language of intention that is usually involved in the description of tool making, we describe evolutionary dynamics that are mirrored in language evolution. Tool development is an evolutionary process that complies with architectural constraint inherent to any evolutionary process. These evolutionary dynamics involve concepts of energy conservation and propagation. The way to understand these concepts is in considering the physical effects that they yield. The story emphasizes the subtle balance between using what is already there and generating novelty in the context of increasingly various tasks.

The invention of the *Swiss Army* knife might have come about in a situation in which several different tasks require an array of dissimilar tools. However, it may be that only bladed tools are available. One-tool-for-one-task is a *high energy state* because it requires a redesign and remanufacture as a new task emerges. However, the energy put into developing the tool pays off in the coherence - the focus the new tool brings - of the work it was designed for. Some of the energy invested in the development stage is recouped in the ease to perform the task since the tool is perfectly designed for that task. This is a *stable state*, as a minimum is reached between the input and output of energy.

On the other hand, there is an increasing variety in the tasks to be performed. New tools are constantly required. There is a critical point at which the energy recouped in the ease of performing the task is not worth the energy investment required at the new tool

development stage. What to do then?

So what if tool availability is restricted to a bladed tool - a knife - with a handle, an edge, a tip, a width and a length? We can use these features in developing a new tool. The width of the knife can be increased and a depression can be added to perform a scooping - edge forward - action. Hence is born the spoon. As cutting and scooping tasks are encountered it may be useful to have both edged and scooping features on hand as to save time and energy. Most of us are familiar with *sporks*, which are both a spoon and a fork in one, with which one can both spear and scoop. However, the ease with which it can do both is questionable. It may be more useful to add an additional tool to one handle than to modify one tool to perform several tasks, but we will come back to this dilemma later.

A strategy to add a spoon to a knife handle can be adopted. The spoon features can then be refined while retaining the use of the knife. The work that the new tools - or collection of tool - can do compared to the work of the knife alone has been increased without investing too much effort into it. However the handle is now a little larger, the knife and spoon may interfere with each other slightly and the complete tool is less reliable. The consequence is an increase in work for a minimal energy investment but the effects are not as specific as those originally occasioned. Originally the effort in performing a cutting task is less because the knife is designed for that purpose. The balance of the handle and the weight of the blade is in perfect harmony, so as to aid the cutting strokes. The work done is highly coherent. Now that the knife is not quite as steady because the handle is slightly bigger and does not quite fit the hand, more effort is needed to perform cutting. The coherence of work done is decreased. As more extensions, such as screw drivers and bottle openers are added to the handle, and, suppose, the size of the collection is reduced to increase portability, less effort is required in manufacturing and carrying the collection of tools, compared to creating and carrying new individual tools. However, there is an increase in effort in using this collection of tools. Eventually, each new extension becomes a problem for two major reasons:

1. Adding extensions in a coherent fashion requires increasingly more effort than originally. If portability is priority then adding more tools will only increase bulkiness and awkwardness of handling. The steadiness, in using one tool that is attached to many others, will be compromised as well.
2. The collection of tools is able to do much work but the effort required to achieve specific tasks becomes taxing because it is inadequate in its lack of specificity and

reliability.

In the next stage of evolution individual tasks are evaluated to see which ones are repeated more often. For these tasks it may be more economical to invest the effort in making a specific tool. However there is no need to discard the all purpose tool, it can simply be put in a support role for the specialized tools. With a combination of specialized tools and an all purpose tool most situations can be tackled effectively. As such, a fairly high level of coherence can be achieved.

The dynamics involved in language evolution follows a similar pattern. Language rich in perceptual associations is a specialized tool. As these vocables are generalized to do more *work*, the specificity of effects that are generated by them is less coherent. Once vocabulary has lost some of its perceptual cues and starts generating incoherent effects, it can be further exploited by gaining a support role for linguistic instruments that do generate specific effects.

We think that functional vocabulary has evolved to support vocabulary on its way to become increasingly less specific but lexical nonetheless. Functional vocabulary has lost most of its perceptual associations and seems to be restricted to a syntactic role. The *one purpose vocabulary* - lexical even indexical - is a higher energy state and highly coherent because it generates specific effects. It is easily accessible for it is entrenched in perceptual cues, but as such, requires much neural space. For example: If we refer to *THAT table in THIS room* the effects will be very specific and will be entrenched in very specific perceptual cues. The *all-purpose vocabulary* - functional - is a low energy state. We know how to use it despite not having a semantic theory for it. It is fairly coherent in the effects that it generates within its syntactic role. However, it does not generate very specific effects in relationship to perceptual cues. Jennings has supplied many examples to that affect, *but*, for example is very elusive and it is very difficult to associate it to any experiential cues.

3.2 Language and Efficiency

The language system tends toward efficiency and as such, struggles between states. Much like the making of tools, language has a tendency to keep the vocabulary that is already there however, new environments require new vocabulary so eventually, energy must be invested in restructuring aspects of syntax so that a lower energy state may be restored. Because the 2nd. law of thermodynamics tells us that the arrow of time is irreversible,

the systems may revert to a low energy state, however it may do this only by generating a different configuration. In language the struggle is between vocabulary rich in perceptual cues - lexical - and vocabulary poor in perceptual cues but coherent enough in its role - functional. They are both stable energy states and they are both structurally different. Lexical vocabulary is easy to use because it is easy to describe but it is intolerant in allowing different instances of itself so it does not scale to other environments very easily. In other words, the work it can do is very limited but it is very coherent. Functional vocabulary is difficult but it is versatile and is more tolerant of different instances of itself and as such, it finds itself in many more environments. It can do more work but it is more incoherent in relationship to perceptual cues.

3.2.1 Language as a physical intervention

All of language is a physical intervention. Moreover, language is the efficient consequence of a larger system in interaction between organisms. So how is it efficient and how are specific interventions *selected* for efficiency? We can imagine that there is some minimal efficiency required in the activity of language; otherwise gestures and grunts would be the favored approach in communicating. How is it then that gestural interventions gave way to mostly linguistic ones?

Every intervention is an exploitation. Naturally occurring phenomena serve our activities frequently. The application of launching a message into the flow of a stream, is such an example. More importantly it is the motion of the flow that is exploited and not so much the stream itself. The *registering* of that fact is exploited by extending it to other environments and creating alternatives. So if a stream is not present to carry the message, the function of motion is implemented in some other way, using a bird for example instead of creating a stream to carry a message which in fact would require a tremendous outpouring of kinetic energy for the simple task of carrying a message.

3.2.2 Exploitation of incidental effects

What are incidental effects? They are regularities. Regularities are features shared by interventions. This is not to say that these features are *hidden* in the world just waiting for us to discover them. Nor is there a talent in our nature that allow us to *pick out* these regularities for us to exploit the way Hume describes it in *Of the connection or association*

of ideas in his *A Treatise of Human Nature*.

Hume thought that the mind organizes ideas in a taxonomy. The particular organization of the taxonomy relies on how ideas are brought to mind through different types of relation, the most important being *resemblance*. The theory assumes that ideas go from particular to general through a sequence of degrees of relationships. Because the reality of an object, for Hume, is not separate from the apprehension of quality or quantity up to a particular degree, it follows that forming the idea of an object is the same as forming an idea. The way by which ideas become general is by accumulating all degrees of quality and quantity associated with an idea. This is not a perfect ability but it is the one suited for the purpose of apprehending the world. It is the task of the imagination to *sort* through the taxonomy of ideas and extract the contextual ones.

This view is still widely shared. There is an appearance of exploitation of regularities on the part of our species; however, our approach does not require the world to exhibit regularities. Registering regularities is a feature of our biology. Registering, is being conditioned to expect certain effects given stimuli of some kind, not unlike a Pavlovian [45] reflex. Simply, it may be that our capacity to discriminate is limited because of architectural constraints. Discriminating capacity does not come cheap in many ways including processing effort and time. A lengthy discriminating process may reduce the survival rate of our species. Having a *notion of potential predator* based on a limited discrimination of resemblance is a useful thing to perceive in a split second. An infinite capacity to discriminate would hinder the process of registering regularities hence making very difficult categorizing based on similarities. How could we have syntactic rules if features of utterances would be overwhelmingly different as opposed to having anything in common. We couldn't refer to *noun phrases* or *verbs* because *object* or *state* would never become apparent as a shared feature. In the case of language, earlier stages of exploitation are of non-linguistic kinds but later stages of exploitations involve mostly linguistic regularities. There are many evolutionary reasons why linguistic regularities would be exploited over non-linguistic ones. A striking example is to be found in people's ability to coordinate each other or large groups into performing complex tasks. These physical effects would not occur if not for the event of language.

3.3 Efficiency in Other Language Theories

The use of efficiency in explaining dynamics of language is not a new approach. Theories of evolutionary development in language have called upon this concept to describe causal relationships between observations in language acquisition phenomena. Good examples on how efficiency is used in explaining dynamics in language are found in the works of Terrence W. Deacon, Eliza Newport and Bernard H. Bichakjian [18] [8].

3.3.1 Neurobiology's account of language emergence

Deacon [18] offers a point of view in which language is adjusting to architectural constraints associated with children's brains. The general process of language acquisition is attributed to a *co-evolution* between the capacity of the brain to produce language and a kind of competition between different types of parts of speech that are more readily available to the human brain. The present state of syntactic structures still undergoes constant transformation, the major influence being the child's brain. Deacon suggests that children do not have an amazing gift for learning language, more to the point, language adapts to behaviors and brain development already present with children. This process is analogous to what learning psychologists call *shaping* of an operant behavior. A spontaneous behavior is reinforced and modeled to be reproduced on demand. So particular syntactic features that stay in the language have likely emerged from spontaneous behavior or noises inherent to the child. This kind of approach to tool development is not alien to our everyday understanding. For example: *User friendly* is a term that has filtered into natural language from computer developers. Computer interfaces have been made to copy office environment so that gestures that are performed in that context would not need to be learned as they are already part of a routine. Similarly these behaviors can be *shaped* in such a way that intuitive guesses as what to do in a *user friendly* environment will more often be right than wrong thereby significantly reducing learning time and effort. Similar assumptions apply to the process of learning language. This approach is not only relevant in the context of children. As highlighted in the computer example, adults use this approach in tool development frequently, mostly because complying with brain functions will yield expected results better than forcing millions of years of biological evolution to stir rapidly. It is a known fact that the time scale of organisms is slower in generating mutation and quite stable compared to the time scale in which language mutations are shown to occur. A more interesting fact is

that, as adults, we retain the ability to make novel associations hence we are able to *shape* new behavior throughout our lives.

3.3.2 Psychological parallels

Elisa Newport [43] noticed in her research with young children that while they had difficulty acquiring *conscious novel associations* because of a lack of focus and general short attention span, they displayed a remarkable ability to gain language skills. The idea that “less is more” was put forward as an explanation to language acquisition in early stages. More data was gathered pointing to that fact with research involving a bonobo chimpanzee called Kanzi. Kanzi was the adoptive child of Matata, a female chimp involved in Sue Savage Rumbaugh and Duane Rumbaugh’s [56] research on language acquisition. Kanzi was exposed to the visual lexigram that Matata was being trained to acquire, during early infancy, while holding on to and exploring areas around his mother. Kanzi showed spontaneous even vicarious ability to reproduce exercises that Matata was unable to master. When it was thought that Kanzi was old enough to be trained specifically he showed trainers that he already knew most of what Matata had been exposed to. His capacity to understand novel associations, even ones that are syntactically correct but in some way anomalous and react correctly to specific inferences, is far beyond the capacity of his predecessors. He even showed abilities in understanding complex uses of spoken English. The interesting point is that Kanzi learned speech and the symbolic use of the lexigram without specific training. Significant increase in his performance is without doubt attributed to the continuing efforts of trainers to develop specific skills in more and more efficient ways but Sue Savage-Rumbaugh reported that Kanzi seems to have a much better sense of what is relevant to symbolic and linguistic communications. He attends to appropriate cues and boundaries where chimps trained at an older age never seem to surmount the difficulty of picking a specific stimulus out of many, and reacts appropriately to specific inferences. Results point to the conclusion that Kanzi’s early exposure to language has biased his learning in a way that enables him to pick emerging features out of complex syntactic structures not unlike children’s attention to phonemes in the surrounding language.

3.3.3 An argument against structural predisposition

These findings are contrary to theses which support the idea that the brain evolved to support the development of symbolic language and that the human brain is born with a structural sensitivity to syntactic construction, such as Chomsky [13] would suggest. Kanzi's abilities contradict these ideas because none of his ancestors developed language skills in a natural setting, his synaptic functions could not have been *prepared* to receive language. Moreover, the idea that infants' synaptic functions cannot acquire language skills once a certain level of maturity is reached because of a genetic time clock, is put in perspective. A newer view suggests that as a child reaches maturity, synaptic functions are involved in many tasks that are not necessarily linguistically related. If, during growth, language skills are not developed, synaptic functions that would usually be involved in performing linguistic tasks will be co-opted into performing other tasks[20].

3.3.4 The plasticity of language

Bernard H. Bichakjian calls this particular feature paedomorphosis which refers to a biological alternative in which aspects of infancy continue to be inherent to mature aspects of an organism. For example infant facial features are recognizable in adult humans. His theory suggests that linguistic paedomorphic alternatives have selective advantages. Based on the research of LeVay, Wiesel and Hubel, Bichakjian [6] states:

It is a known fact that language plasticity like other types of cerebral plasticity is not a potential held constant during a given period of time, but a decreasing disposition during a specific ontogenetic span. Therefore the sooner the neural pathway of a given linguistic feature is laid out, the stronger it is. Better established neural pathways are not only less impervious to negative somatic and psychological conditions (e.g. fatigue or insecurity) and pathological processes (degenerative or accidental types of aphasia), but also more economical in their regular use, since their activation requires less energy. Given the importance of early-established neural pathways, the shift to earlier-acquired linguistic features now becomes a development towards more advantageous alternatives. -Bernard Bichakjian, 1995, p.53-

Basically, vocables that are readily available to children will tend to linger in the language because they are in use earlier in children speech. Bichakjian has surveyed families of indo-european languages and has concluded that the trend from old forms to newer forms in languages occurs as a selection for paedomorphic alternatives. These selections happen in stages, and there are different strategies that support these observations.

3.3.5 Head-first and head-last languages

In following transformations in the family of indo-european languages, Bichakjian noticed that the relationship between subject, object and verb has changed, allowing for a more complex syntactic structure. According to Bichakjian's research, the word order has shifted from *head last* to *head first* constructions in which the *heads* are usually verbs and prepositions and their objects are *modifiers* such as attributive adjectives. What constitutes *head* and *modifier* is not clearly defined, so modified nouns and adnominal complements are modifiers, as in *john's BACK or the BACK of the door* [7] where *back* is the *head*. This example targets specific language items but the mapping of *head* and *modifiers* can be applied to auxiliaries and past participles, comparative adjectives and terms of references, case markers and verb roots, coordinated conjunctions and coordinated nouns and in numerals, tens of units, in which the first element, say *twenty* is a *head* and in *twenty-five*, where *five* is a *modifier*.

In using this analysis, Bichakjian has investigated ancestral structure and discovered that there is a shift of the head from an end position to a initial position. His explanation is that *head-first* constructions are more efficient than *head-last*. His definition of efficiency is based on how early in life a child can acquire a construction. That is, fewer *energy resources* are spent to acquire specific language features. Head-first languages are more efficient because psycholinguistic research shows that this kind of structure is acquired earlier in life than left branching[8]. Moreover the shift from last to first position marks the disappearance of inflexional markers which are usually more difficult to learn. Also, head-last constructions are problematic when subordinates are included because they have to be embedded to the left of the sentence, thus head-first constructions offer the possibility for greater complexity in sentence construction.

The advantage in *head-first* languages are not only measurable in terms of efficiency in learning time but also in real-time decoding of speech. A *head-first* construction allows the speaker and the hearer to use less *processing* effort in coding and decoding because the

structure of the *head-first* language fixes the word order such that the sequence of a sentence can be decoded serially as it unfolds and the full sentence does not need to be uttered before the employed syntactic strategy, is inferred. For example, consider the the formally correct sentence:

THE DOG CHASED the cheese eating mouse catching cat, is problematic to utter and to understand. The sentence:

THE DOG CHASED the cat that caught the mouse that was eating the cheese, is far easier to decode than the first one[7].

Subject-verb-objects relationship

The specifics of the order change are noticeable when examining the relationship between subject, verb and object. The switch from head last to head first means that case markers and inflexions have been shed as the *subject-object-verb(SOV)* relationship switched to *subject verb object(SVO)*. For example, *We three kings of Orient are* is an example in which the more archaic form (SOV) is used. *(I think that) Peter has sold his house* is an example of an SVO construction. The resulting order, as inflexional markers disappear, is (SVO). An interesting fact about this increase in efficiency is that the decrease in processing energy entails the selection of a syntactic regularity, (SVO), over inflexional markers and an (SOV) relationship.

The *information compactness*, as defined in the previous chapter, for the new form is assumed to be more *compact* than for the ancestral form, that is, a description for a *head-first* language would be shorter than the older form. Moreover the reduction in energy involves a more stable, less flexible structural organization, not unlike many natural structures such as crystals in which a decrease in the energy available in a liquid, say water, will reconfigure the relationship between molecules to form a solid, ice. A crystalline configuration is information theoretically more compact, than a liquid because all of the crystal's structure can be inferred from a basic lattice that describes the organization of a few elements, repeated successively through out the crystal.

Head-last; right brain, head-first; left brain

Bichakjian's explanation for this occurrence sits in brain development. According to him, *head-last* constructions use resources of the right hemisphere rather than the resources of

the left hemisphere of the brain. The right hemisphere is involved with spatial references where as the left hemisphere is involved with analytical and linear activities, specifically temporal references. This begs the question; if language is a specialization of finely tuned right hand control, which is a left brain function, how can a *head-last* language be the product of right brain resources? Language is a function of the left brain as it is a linear production that involves temporal sequencing but in the case of *head-last* constructions, left brain resources are possibly directly stimulated from proprioceptions of the right brain. Hence, the syntactic objects of a *head-last* language are structured in a spatial manner. That is, the construction of sentences in *head-last* languages have to be apprehended as a whole before the specific syntactic strategy that is occasioned in an exchange, is inferred. In effect a *head-last* construction sets up a spatial relation between vocables.

It is not clear how the migration from right to left hemisphere has occurred. It is also not entirely clear what are the benefits of such a shift. It may be related to an increasing exploitation of left situated synaptic structure related to the fine-tuning of ballistic motions, as Calvin suggests [11]. It may be that the capacity to fine-tune a movement is extended to the fine-tuning and an increasing by expert use of language.

There are sufficient reasons to think that, as a physical phenomenon, language will be subject to environmental pressures through which it may become more efficient. As such, language must comply with the most elemental features of Nature. This is why, in the following chapters, we will describe how the language of Physics can be co-opted for the purpose of a theory of language.

Chapter 4

The Physics of Language

4.1 Dynamics of Change in Language

Dispersion is an important process in language evolution. The presence, throughout a human population, of tokens of any particular linguistic type inevitably involves some process of dispersion. Dispersion is a dynamical process that describes change of a kind. One principal instrument of dispersion is constituted by the processes by which children newly acquire language. Another important example is how technologically inspired neologisms, fashionable inventions and rediscoveries, and other linguistic innovations, such as slang and colloquial modifications, also become dispersed among the linguistic repertoires of adult or other sub-populations.

In the course of dispersion change is inevitable. In the first place, even in the case of nouns and adjectives, individual language users can acquire slightly different extensional biases and preconceptions. These will be inherited, again in slightly altered form, by further language users who acquire the use from others.

Dispersion multiplies the instances of a specific vocable, that is, this vocable will gain increasingly slightly different uses, over time, as it is uttered in an increasing number of linguistic transactions.

The English word *internecine* provides an example of this purely extensional alteration.¹

¹In its earliest uses, the adjective *internecine* referred to a struggle that is mutually destructive, such as the Peloponnesian War between Athens and Sparta. Possibly from some earlier faulty understanding of pre-Hellenic politics, possibly through some confusion involving the noun *niece*, American usage has treated the word as though it meant *intrafamilial*. To which Merriam-Webster [4] says:

From the latin internecinus, from internecare to destroy, kill, from inter- + necare to kill, from nec-, nex

Any in English interrogatives has free-standing existential import as; *Do you have any siblings?* But it also has universal import in certain affirmatives where it is mostly restricted to negative and modal constructions. (Compare *I have not read any of these books* with *I have read any of these books* or *You may have any soft beverage* with *You have any soft beverage.*). Such divergences are occasioned by specific and kinds of schisms in human speech, in which one party of language users understands a word differently, but undetectedly so, from some other party. But the mere use of such vocabulary as adjectives, adverbs, nouns and verbs increases what could be called their *actual extension*. An actual extension is the class of things or n-tuples of things to which they have actually been applied by speakers of the language. That modification, to a standard definition of extension, alters the circumstances under which linguistic descendants receive the vocabulary. Usually, the extension of the relation of a particular vocable will only include one class of items such as nouns or adjectives or verbs, and so on. The actual extension can potentially include all of those and more. We will explain further in the section about *attenuation*.

Overwhelmingly, the tendency of these changes is toward a relaxation of earlier inhibiting controls on the uses of vocabulary.

Specializations as well as generalizations, in vocables, can be products of the kind of relaxation we have in mind. Consider, for example, the word *fond*. The loss of inhibiting controls represents as well a loss of capacity to have particular sorts of effect of *fond* to create an expectation of foolishness, even in connection with the object in respect of which fondness is attributed.

Early on, *fond*, in its use, was akin to the current use of *crazy*. To have a *fond* hope is an expression that is seldom used, mostly because *fond* has been restricted to another kind of syntactic structure such as *fond of someone*. In fact *crazy over* replaced *fond of* in generating the kind of *emotive* effects that *fond* used to produce. To be *crazy over* or *about* someone does not suggest psychopathology as: *This person was institutionalized because he is crazy* does. *Crazy over* is an eccentric use of the word in an effort to convey strong feelings for someone or something. In this case *crazy* is not meant as a disorder but rather as an

violent death – more at NOXIOUS// Date: 1663 1 : marked by slaughter : DEADLY; especially : mutually destructive 2 : of, relating to, or involving conflict within a group (bitter internecine feuds)

Although Americans use *internecine* to mean *within a family or group* the use of the word with this meaning is almost entirely restricted to environments mentioning conflict, as, for example, the expression *internecine dispute*. (We hear no auspicious references to internecine picnics.) This is a dramatic example, though this odd, even mysterious restriction of use is paralleled in other sectors (from R.E. Jennings, in conversation).

occasion of a mild form of the original inferential effects. Unlike *fond*, *crazy* has retained its original use for now. We might assume that *crazy* will suffer a fate similar to that of *fond*; that is, the use of *fond* as shifted from one syntactic form to another, while slowly shedding the original use. *Fond* is further along in the widening of its extension, more specifically, it has undergone a narrowing of use through the widening of its extensions. Statistically, some instances of *fond* have been discarded while others have survived. These instances are often relaxed uses. That is, the core use has been lost, while the eccentric use becomes the norm. The extension of *fond* is thus altered.

The loss of extensions is the basic dynamics at the root of language evolution. These examples illustrate a tendency for individual vocables to evolve towards a point of transformation wherein extensions slowly disappear from a conversational use of language until all instances of a vobale no longer generate any of the early effects.

One of the consequences of the loss of extension is the functionalization of vocables from lexical uses. We pose the question: What is the use of functionalized vocabulary?

Vocables may disappear from the language or they may change in structure; that is, they may transform from the sort of vocables that we *understand* readily and use accordingly, to the sort that is not well *understood* and used expertly nonetheless. To be more specific, the extension of a vobale is a gradient of instances that ranges, early on, from lexical uses entrenched in perceptual cues to instances that, later on, may become very abstract, even vague, and in extreme cases become functionalized uses. Functionalized uses are those of which early extensions have been forgotten and take on almost purely syntactic roles.

As explained in Chapter two, functional vocabulary descends from lexical vocabulary. The distinction may not seem dramatic; however a consistent semantic theory of functional vocabulary has evaded philosophers of language. As we have mentioned in the introduction, Philosophers' semantic accounts of functional vocabulary differ widely and there doesn't seem to be much consensus. However they seem to agree that lexical vocabulary is semantically difficult while a semantic for functional vocabulary seems more available. We think that the reverse is more likely. Lexical vocabulary is more readily definable when we inquire about its use, casually in a conversation. In the history of vocables, it is the functionalized instances of some vocables that seem to lose the capacity to be well *understood*. So the question remain, why do we have functionalized vocabulary? And how do we use it?

4.1.1 The Coordination role of linguistic activity

In order to understand the dynamics of language evolution we must inquire about the roles that linguistic interaction plays in human population.

Language is a binding process causing specific effects in a population. These effects must have a long history.

We share a linguistic system that binds us in many ways, one of which is a role of coordination in a population. Vocables are uttered and physical effects occur. A person can call up a friend to set up a meeting and gather in a particular place. With linguistic interactions, events may be set in motion that otherwise would be difficult. The coordination of activity occurs with some precision. But as explained in the *Efficiency of Language* chapter, the precision in linguistic interactions is only as refined as it needs to be in order to be shared in a population.

However language does respond to the requirements of increased specificity by becoming more refined. If we assume that linguistic transactions play a significant part in coordinating human activity, which is not to say that language emerged out of a *need* to communicate, we may also assume that our non-linguistic ancestors may not have had the tools for the level of precision that humans experience with the help of linguistic transactions. The suggestion is that the level of precision, with which coordination occurs, will change as the communication tools in proto-human activity become more sophisticated. It may also be that the requirements, in activity, for coordination, applies adaptive pressure. The adaptation can occur because these transactions are coordinated as well. For example, in linguistic transactions, it is the specific permutations of words that make a sentence generate specific *effects*. Not just any combination of vocables will generate specific effects, however, vocables that belong to different categories seem to be required in order to form a sentence that is considered grammatical, that is, a sentence that can be *understood* by most users of the language.

There is a balance, it seems, between adaptability and effectiveness. Now, the role of functional vocabulary seems to have more to do with helping us in the construction of grammatical sentences than in generating very specific *physical effects*. Functionalized vocabulary is not the kind of vocabulary that we commonly use to make statements about the world; that is, as we will explain further on, its relationship to perceptual cues has been eroded beyond recognition. We mean this quite literally. Nonetheless, the stringing

of vocables is done, we think, with some amount of precision. How much precision? As we have said, as much as is needed for the language to be shared in a population as the system propagates and as it is inherited by following generations. This means that it is not as precise as some may think. This lack of precision will allow mistakes in the production of language, mistakes that will be replicated and adopted in conversations. However, language is stable enough, in the short term, to be successful in its coordinative role and to be shared amongst a population.

At some point, our guess is that linguistic capacity was exploited toward the coordination of human activity, a role that might have overwhelmed any other use *linguisticity* may have had in its early times. The evolving capacity of human beings to coordinate their activities and the consequent changes in the depth and detail of actual coordination must itself have fostered profound, perhaps fundamental, changes. We know, for example, that we come from non-linguistic ancestors. Structurally, proto-humans were different from humans. The laryngeal morphology and vocal cords, the level of motor function involved with facial mobility and dexterity are some of the fundamental ways in which humans differ from proto-humans, yet, we have assumed that at some point in time there were human ancestors that did not have linguistic capabilities. We have no examples of the types of items that coordinated proto-human activity or the kinds of structural changes that promoted or sustained the slow emergence of language.

We must make assumptions about early language, based on recent language, and project back to earlier time. We assume that social organization involved types of communicational items, items that were the distant ancestors of vocables. We have examples of structural changes in the vocables of present languages, from which we can extrapolate, that is, we assume that the dynamics underlying contemporary linguistic changes are present throughout the history of structural changes in communicative behavior. We also have a theory as to how these changes promote a more detailed and versatile coordination of human activity. Though structural changes in linguistic types may occur at an accelerated rate, compared to the crawling rate of change in proto-human communication behavior, and may be less dramatic than the structural changes that promoted the earliest linguistic transactions in humans, they may nevertheless serve as approximate models of conjectured earlier developments in the evolution of language.

4.1.2 Time scale

Language has been evolved, via minute changes, similar to the erosion process in landscapes over millions of years. It might be tempting to think of the development of language as the acquisition of function, specifically the acquisition of linguistic function by a physical type whose earlier functions were exclusively non-linguistic. It might also be tempting to think that these changes occurred somewhat abruptly, over a fairly short time scale. The view here presented is almost diametrically opposed to that. Oversimplifying the matter, we can say that early linguistic function was the residuum of worn-away non-linguistic function, and later linguistic function the residuum of earlier. The wearing away is the inevitable consequence of use. Moreover, we cannot overstate the point that linguistic function is an ever-changing process that extends over a much larger timescale than the one we are usually tempted to consider. Also, these changes that occur are minute, unnoticeable yet incremental. The cumulative effect is felt (or observed) only over this large time scale. We suspect that some of these observable consequences of change may occur comparatively quite suddenly, maybe within a few generations of linguistic activity.

Language responds incrementally and adaptively to its applicability and effectiveness.

The respect in which this is an oversimplification is this: As with many or most evolutionary developments, it could not be predicted in what manner what residuum of what earlier feature may confer advantage, but we can confidently predict that the manner of its exploitation will self-select for some subfeatures, not select for others, and select against still others. Vocables may begin to emerge as the awkward exploitation of some relatively inefficient non-linguistic residuum, but the awkwardness itself exerts evolutionary pressures: what is not needed or actually impedes we can expect to be refined away.

A corresponding account applies to functionalization within language, and accounts, we presume for the morphological alterations that accompanies, for example, the logicalization of lexical vocabulary.

We assume many facts about early linguisticity because we think that structural changes in language, across the history of its evolution, are motivated by forces that the idiom of physics has successfully described. If the dynamics of change in language can be described using physics, we assume that observations based on a present sample of language will scale to past samples since these forces, as far as we know, do not change over time.

4.2 The Language of Physics

It is a difficult matter to settle upon a suitable idiom for an adequate description of language. It is one thing to insist that language should be regarded as a physical phenomenon. It is another to choose the physical vocabulary that is best suited to its description as such. As we have earlier remarked, the language in which philosophers speak of language can be ruled out for such a role. Terms such as *meaning*, *intention*, *belief*, *desire*, and so on, by general consent, do not refer to observables.

Instead, we consider many intuitive notions about words in physics that *are* observable terms, some of which can apply to language. The vocabulary of physics is defined by the mathematics of physical models, whose application to language is not apparent. In this case, the natural language correspondents do have familiar, intuitive, non-theoretical physical interpretations, and do seem to have applications to linguistic phenomena. Terms such as dispersion, elasticity, structure and dynamics, have a conversational use that is usually associated with spatial phenomena, occurrences that describe features of experience. Children dispersing in the courtyard, the elasticity of skin, the structure of a building, the dynamics of populations: all of these are familiar constructions, readily understood, at least for purposes of conversations but readily mathematizable without immediately obvious distortion.

These conversational uses are obvious in the examples that we have provided but their application to the description of language is more challenging. Somewhat less accessible are constructions such as *the dispersion of (features of) a language*, *the elasticity of a vocable*, *morphological and syntactic structural changes in lexical vocabulary*, *dynamics of linguistic evolution*. But they all have correspondents in the idiom of physics; that is, the vocabulary finds a more precise definition, as well as a mathematics, within the various theoretical languages of physics. Our question is: Setting aside its more casual uses, can any of the physical applications of such vocabulary add useful mathematical clarity to our understanding of linguistic evolution? Beyond that, can this detailed account reveal hitherto unacknowledged features of language?

4.2.1 Dispersion

We introduce the discussion of dispersion as it applies to language.

As we have mentioned several times, language is a shared system of interaction. We

inherit it from previous generations and from our interaction with others. We have discussed loosely how the system *propagates* or *disperse* through a population. Let us examine more closely what the notion of *dispersion*, in physics term, can add to our understanding of linguistic interaction.

We define dispersion in the context of *state space* and *equilibrium*. *Dispersion* (of components of a system) is a microphysical phenomenon underlying some specifiable macrophe-nomenal relaxation of a system toward a state of equilibrium. Equilibrium is a point in *state space* where, in accordance to some state variable, a system comes to rest, reflecting the most probable place in state space for it to be. Of course, choosing the appropriate state variables of the state space is a matter for decision, rather than a matter of fact. As a simple example, consider the system defined by a child's play room, there are many more states in which toys are scattered across the room than there are states in which they are arranged in orderly fashion on shelves. Depending upon where in the hierarchy we position ourselves, that is, how many and which parameters we use to define the space, the distinguishable states of scatterings (teddy here, or over there) will be one point or a cluster of points in the space.

We introduce our perspective from the point of view of statistical mechanics, thus defin-ing equilibrium. Statistical mechanics is the branch of physics that addresses the relation-ships between energy and forces of change, in a macroscopic system relative to microscopic constituents. It is the key to understanding thermodynamics and describes almost every change in nature from cell development to combustion in car engines. We borrow the notion of *equilibrium* from statistical mechanics. It defines a closed system (a system that is influ-enced only by forces within (no loss or gain); a hermetically sealed pot of coffee is considered closed) in terms of one or more *state variables* such as temperature, pressure or relative con-centration. It then describes a *state space*, a space defined according to these state variables in which innumerable possible configurations of the system can be represented as unique points. Equilibrium then refers to a condition in which the system's current state remains proximal to some average state over time. In other words, even if other aspects of the sys-tem are changing, the selected state variables are approximately constant. Its macroscopic limits remain constant.

We use a description of the *relaxation* of a system to equilibrium and establish the relevance of state variables such as temperature, to illustrate our point. As a simple example consider an ice cube in a glass of warm water. Assuming the ice, water and heat to be within

a closed system, the global temperature can be assigned the role of state variable. Clearly it will be difficult to define the temperature spatially, as well as temporally, as the ice melts and the hotter water cools. Eventually however, the ice will be gone, the water will be thoroughly mixed and the temperature constant throughout. Thereafter, even though the individual molecules may continue to move around owing to ambient thermal, we can say that the system has come to equilibrium. Its state variable, temperature, has assumed a constant value and will not change thereafter (aside from stochastic fluctuations around that value). Dispersion, in this example leads to equilibrium.

Now that we have defined relaxation and dispersion, how do we use it in the context of language? We have used *dispersion* conversationally to illustrate some of the shared aspects of vocables across a population of users. The main feature of *dispersion*, as we have described it in our ice cube example, is the dynamics that settles a concentration gradient into an homogenous state. Is there some aspect of human linguistic practice that can be usefully described as *dispersion* in the sense in which physicists use that term? Using what state variable can we plausibly claim that there is a kind of *relaxation*, in the theoretical sense of the term, that is specifically linguistic?

We think that dispersive effects are relevant to a description of language if we treat instances of specific vocables as statistical probabilities. We know that, in the course of their becoming dispersed (in the conversational sense of the word), vocables increasingly lose the capacity to generate highly specific effects. At the earliest stages of vocable use, the effects that are occasioned are highly particular; at later stages less so. The word *but*, or more properly its ancestors, had a relatively narrow range of uses: In particular, those in which it picked out a highly specific physical relationship, that of being spatially related *just outside of*. At later stages it was also capable of picking out a relationship of conceptual and then circumstantial *outsideness*, (nothing *but* grief, *but* that he was my own son ...), however in environments in which it has such uses, the original physical relationship no longer competes. Moreover when that physical relation *is* selected, the work of picking it out is shared with other sentence elements. In these later stages, *but* has lost its ability to generate so specific an effect. In probabilistic terms, its earliest uses have become the least probable of its uses. But, by the same token, it has acquired many other syntactically distinguishable uses, each of which has a relatively low probability. Moreover, among these other uses, the syntactic type that distinguishes it is itself so general that we would be hard pressed to say what its effect in such constructions amounts to.

The loss of specificity in vocable effect is verifiable but a description of the dynamics of this phenomenon is not yet accounted for. This thesis addresses that topic.

Consider in this connection the conjunctive use of *but* wherein it joins whole sentences. It is all but impossible to formulate in semantic terms the effect achieved by *but* rather than *and* in these constructions. But the same point is seen if less dramatically in each of the many disparate classes of use. The loss of semantic specificity is an inevitable consequence of use, even within the confines of a single syntactic class of uses. In fact, many if not most vocables both completely lose contact with the earliest classes of constructions in which they were used and also eventually find themselves used throughout a much broader range of constructions through which there is no consistent semantic thread. R.E. Jennings [32] has demonstrated that unified semantic theories for connective vocabulary are not readily available. *Or*, like *but*, has both disjunctive and conjunctive uses, though the evolutionary processes that brought about these broad semantic disparities are different for the two cases. It is a central claim of this thesis that this kind of functionalization is a particular instance of a mathematically similar process pervasive throughout all stages of the evolution of language.

4.3 Attenuation

Generalizing vocable uses becomes a process of *attenuating* its effectiveness.

Attenuation, as a process, includes more than the extension of certain relations, it also includes changes in the type of relations. The notion of *attenuation* is not the familiar notion of temporal increase in the extension of a relation. From a model-theoretic point of view, we think of an n-adic relation as a set of n-tuples of objects, and if we choose, we come to think of a relation as the union of a set of sub-relations temporally ordered by inclusion, say, and so to speak of the extension of the relation varying over time. But from the same model-theoretic standpoint, attenuation represents more than just a change in extension. In the first place, it includes changes in the semantic types of objects, and therefore changes in the semantic types of relations. Ultimately attenuation includes the introduction of uses of the vocable which, though n-adic, are not relational (that is, joining nominals) but rather connectival (that is, joining main clauses). The relatively primitive *but* of

No one BUT Harry spoke
eventually gives rise to the *but* of

Harry spoke, BUT Hilda did not.

Thus the notion of *attenuation* requires resources for its explication beyond those required even for temporal variability in the extension of a relation. Extensionally speaking, an n-ary relation is just a set of ordered n-tuples. To say therefore that the extension of n-ary relation R varies over time is either to utter an oxymoron or to speak elliptically. Conversationally R acquires members, but within the standard semantic representation, that conversational idea is accommodated by speaking of the n+1-ary relation *R* in which the n+1th position is a temporal index. What attenuation and its eventual outcomes require is more radical than this, since attenuated R begins to acquire n-tuples of vocables of different semantic types. What was relational vocabulary may acquire pre-connectival uses; that is, it may find n-tuples of vocables in its extension that are the semantic representatives of noun phrases and noun clauses rather than of nouns (terms).

At this point, it would be semantically, even conversationally, more convenient to admit defeat and introduce systematically variant notation.

4.3.1 Eccentricity and elasticity

The eccentric use of a vocable, and the attendant effort to understand it, is eventually diminished until it is normalized with its normative use.

A secondary characteristic of attenuation can be expressed in the language of *eccentricity* or *elasticity* or *felt crudity*. Uses that would be eccentric or crude at an earlier stage of the language are banal or perhaps elegant efficiencies at a later stage. We can imagine that the earliest uses of *going to* (or rather its ancestral vocable) all involved *motion toward*. That is, there was a stage in the development of the language at which disappointed expectations of motion toward would have been sufficient evidence of a false or eccentric use. The extraction from such a use at such an early stage, of expected outcome without motion as in

I'm going to be sick

as distinct from (certain occasions of)

I'm going to buy groceries

would require a little extra neural effort. At a later stage, it takes no unusual level of effort at all; whether the use is a *motional* or *motionless* use is simply supplied by context and filling. At the latest stage, the default reading is one corresponding to no expectation of physical motion toward. Similar remarks apply to the *have* of

I have finished for the day.

There is no expectation of possession, only of attainment. Unlike the *have* of *I have a table* which creates expectations of possession or at least of acquisition.

This change has a neural correspondence. Put into simple terms the use of the language of *elasticity* can be justified on purely neural grounds. With reference to the particular example above, the eccentric, motionless use does not diminish the automaticity of expectation in later uses; after such a use the expectation is triggered as a default effect. In later stages, There is no such default state.

When a use that was eccentric is eccentric no longer, the vocable in use may be described as having lost, in some dimension or other, a degree of elasticity. As in the case of lost physical elasticity, the level of neural effort required to maintain the extended or stretched state is reduced or eliminated: the extension of the object in its stretched state has become the normal extension. The usage implies that at an earlier stage a linguistic item has had a greater degree of elasticity, and this implication requires some brief comment. This is its intended import: a commonplace word used in a commonplace way requires some baseline energy expenditure on the part of an interlocutor to play the conversational role that it has. Somewhat novel uses require expenditures of energy that are somewhat above the baseline. How far above depends upon the degree of novelty. Very minor extensions of use require sufficiently low levels of additional expenditure as to put little strain upon elasticity. Such little resistance as there is is easily overcome.

To put the matter another way, it is unclear, at a given stage in the life cycle of a linguistic construction which of its uses are normal, and which are mildly eccentric, for sufficiently mild eccentricities. That range of indeterminacy must accommodate some of the expected variation in the linguistic experience of the particular linguistic population. An earlier rest state of a linguistic item may include few instances that stray from its earliest uses; that is, the item may persist in a relatively quiescent state through many generations of users. But in radically altered circumstances, such as can be occasioned by technological innovations, sudden relatively eccentric uses can find currency. Consider, for example, the innovative application of the verb *drive* to the operation of a car. The *drive* of *drive a car* which is simply an extension of the *drive* of *drive a horse or horses* but involves no element of urging on. In fact someone reported as driving her horse to a rodeo would likely be engaging in an activity akin to driving her children to school, which in turn, again receives no competition from the earlier driving of driving horses to school. In this case, though one can still speak of driving horses, the vast majority of language users remain unaware

of the connection when speaking or hearing of driving a car. The word "drive" has lost its elasticity, the tendency or capacity to be pulled back to its earlier set of effects in the absence of environmental cues.

Uses of vocables resemble the physical nature of elasticity in a number of respects, transitions in currency of usage can signal a radical change in the use of a vocable.

The generation of an eccentric use is akin to the action of stretching a rubber band. The successive generation of eccentric uses relaxes the use of a vocable. As an elastic band becomes easier to stretch after a few pulls, so are a vocable's eccentric uses that become increasingly less eccentric as these uses are multiplied. Eventually, however, the capacity for the elastic band to regain its initial shape is lost and so is the capacity for the vocable to regain its early use. It is at rest in a stretched state. The point at which an attenuated vocable is at rest is a point at which its attenuation is maximal and its extension can no longer include new types of instances. At this time in its lifespan a vocable will no longer change, it has reached a kind of equilibrium state. Some of its earlier uses will disappear from the language, but overall, although changing circumstances modify the extension of the vocable, it undergoes no further categorial shifts.

4.3.2 Elasticity in physics

Elasticity as a mathematical concept finds its origin in mechanical physics. Mechanical physics describes macroscopic features of the universe, how forces such as friction and gravity can morphologically alter the structure of physical items. It defines stress, strain, and other forms of mechanical deformation and their properties. Elasticity refers to a property wherein a body can be distorted to some degree and still retain some salient characteristic, however that characteristic is defined, and be able to return it to its original state. This suggests a maintenance of relationship. Typically, elasticity applies to spatial distortion that nonetheless maintains some relationship between two points in the system being distorted. Jelly, rubber and spandex offer everyday examples. However, there will be a point, as a material endures multiple or extreme distortions, at which it will not be able to return to its original shape and either remains in a distorted state or breaks at the point of most strain such as an elastic band that snaps at the limit of its stretching capacity.

Loss of elasticity may affect the system as a whole. With the loss of elasticity, there is a global change in a structural state (so the concept of a global state variable would apply) that changes the nature of the material. However, it does not always strictly apply to the

whole system. It may take place locally, as with a sharp edge, area of impurity, or other site. However, in that case our focus is on a system within a system. In metal fatigue a small crystallized area will have, within the metal matrix, a modulus of elasticity different from that of the rest of the system. However, with certain qualifications (e.g. a very short length scales such as the context of molecular structure) it is fair to speak about the crystallized area as an entirely separate system and to ignore the surrounding metal.

Language can be shown to behave similarly. Elasticity in language is lost over time. Early instances of a vocable can be thought of as the stage of the system before its extension is stretched. Later instances correspond to a system that has stretched to the point of structural change and has lost its elasticity. Because we are interested in changes in language, we focus on the aspects of the system that become structurally different. We turn to physics as we enquire about the specific ways in which these changes occur in physical systems.²

4.4 The Effects of Attenuation

The process of *attenuation* involves intricate interactions that manifests themselves differently across the lifespan of different vocables. Here we explore those differences through several example of specific types of vocables.

4.4.1 Lexical to functional

Lexical vocables are produced; functionalized vocables survive. We can give a positive account of the conditions under which a vocable remains lexical, and in the absence of those conditions lexical status cannot be sustained. Functionalization is a special case from the consequence of attenuation in which lexical language assumes a new non-lexical structural role. Sometimes an additional use does not interfere with its earlier role and indeed is morphologically and prosodically indistinguishable from it. Consider *going* in its use in progressive tenses, which coexist with its uses as a verb of process and motion. Consider the two presentationally indistinguishable uses

²There is room here for a metaphor. A first hearing might suggest something along the lines of Quine's so-called fabric [49]. Quine describes an interconnectivity between experience and logical vocabulary; how one affects the other in an intricately "woven" manner. Refer to appendix A for a short comparison between Quine's fabric metaphor and our general approach.

I'm going to shop one of which (lexical *going*) characterizes a present departure or motion, and one of which (functionalized *going*) registers an intention. Contrast the distinguishable uses of *have*

I have already shopped and

I already have my outfit, or those of *used*

I used my time wisely and

I used to use my time more wisely

These examples demonstrate that particular instances of verbs can acquire auxiliary functional roles in certain syntactic environments, without eradicating early uses.

4.4.2 Attenuated non-lexical vocabulary

Vocables can be attenuated to a point at which it is no longer lexical and not functionalized either. We make the distinction because these vocables remain in the language but they do not become structurally different; that is, no changes in category occur. *Will, ought, rather, sooner* are example of this vocabulary for which a semantic theory is very difficult to account for. So is *mental* vocabulary such as *intention, mind, truth, moral, and so on*. These vocables might have had lexical ancestors that are rooted in other languages. However, in present English, it is unclear what anyone means with much detail when using this vocabulary. In fact, we can only give a detailed account of attenuated vocabulary by searching for its lexical roots; otherwise, in the event that the genealogy is not available, we must rely on some other account that redefines attenuated vocables in more lexical terms. For example, a more detailed account of *intention* can be given using the idiom of neuroscience. However, this exercise would modify how people use *intention* and so we may consider the redefined term as an entirely new lexical vocable.

4.4.3 Attenuated functional vocabulary

We have already described functional vocabulary but now we look at its dynamics in the context of our newly coined physics idiom.

Functional vocabulary is the type of vocabulary most attenuated. This is because its instances multiply rapidly over time. Functional vocabulary is not immune from further attenuation, in fact, because of its already attenuated nature. Functional vocabulary can

give rise to different principles of generalization that lead to mutations in the scope arrangements of some syntax, as R.E. Jennings has demonstrated in his research. We assume that the level of attenuation of functional vocabulary is such that its syntactic role can be easily misinterpreted by users. The error can go unnoticed or it can be identified as belonging to a different class of generalization.

For instance; *but* is dualized through scope evasion; that is, as a connective *but* can be found in a connective role or a disjunctive role. Often these distinctions will come about in the presence of negation, most often because it is difficult to perceptually entertain negative states.

I won't go for a while is usually understood as
it is not true that I will go for a while rather than

I will be a state of not going for a while, hence changing the scope of *while* in the environment of a negation. *For* is another puzzling case that give rise to different principles of generalization.

I'm doing this for you(because),
this hasn't worked for the longest time(duration) and
this gift is for you(belonging) are examples of the use of *for* in which it is very difficult to give a semantic theory, moreover it is very difficult to ascertain how *for* could have gained so many uses.

Relational vocabulary tends to become functionalized because of its pervasiveness. Once a vocable is functionalized, its instances can potentially find an infinite number of environments because it is no longer restricted to relating nominals. Its extension has been widened to include verbs, adjectives, adverbs, prepositions, propositions, even whole sentences. The additional connectival roles are not discriminated between. Equivalent instances of *but* are used interchangeably whether it connects adverbs or sentences. The features that differentiated early instances of *but* have been lost. In some rare cases such as *or* and *and*, both can be applied interchangeably without much difference, within certain constructions. In

you can have tea or you can have coffee and
you can have tea and you can have coffee, both entail that the two choices are available.

It is only in the resolution of the statement

I will have coffee or no thanks or I will have both, that *or* will be defined. If the answer is *I will have both* we cannot assert that *or* is playing a disjunctive role. In this case, *and* and *or* can yield equivalent effects. That is, users will hardly differentiate between the use of *or*

and the use of *and* in many constructions. The connectival role of functionalized vocabulary is somewhat indifferent to early lexical use and will tolerate several different vocables for a connective role in a given syntactic construction. So *or* can be found in several connectival environments in which other vocables could also be found such as *and* and would occasion similar effects in users. We can say that *or* has lost the features that differentiate it from the uses of *and* in some of its connectival roles.

Statistical physics describes this phenomenon as *degenerate* states of a system.

4.5 Attenuated Language and Degenerate Physical States

Roughly speaking, repeated usage of vocables leads to dispersion. Dispersion brings about a loss of elasticity in the widening of a vocable's extension. The consequence is an attenuation of the lexical form that biases a vocable's use against its original instances.

Functionalization is one of the effects of attenuation that can lead to a loss of feature between functionalized instances of some vocables. An extreme consequence of this is that morphologically distinct vocables, such as *or* and *and* which had only wholly disjoint uses at an earlier stage may eventually find themselves interchangeable in some constructions. This increasing loss of feature is what we call *attenuation* and what physics calls *degeneracy*.

4.5.1 Degeneracy in physics

In simple terms, it is understood that a statistical system at equilibrium tends towards its most probable state.

Then *degeneracy* is used in reference to any one of a class of distinct but equivalent physical states of elements of a system: equivalent, say, with respect to their energy or some other intrinsic value.

A simple example would be a die with six sides, five of which have a single dot and the sixth has six dots. The one-dot state would be said to be five-fold degenerate since there are five distinct sides which are otherwise equivalent.

Even though the die is seemingly composed of only two different states, it obviously has six distinct sides or states, and each of those has equal probability of coming up when the die is rolled. Since we can't tell the five one-dot sides apart, it appears as though the one comes up five times more often than the six.

Degeneracy in states can be understood as a loss of feature such that the multiple states of an item become indistinguishable from each other. In the die example all but one face of the die has lost some of its distinguishing features, that is, an individual number of dots that is used to differentiate the faces. What remains is the one dot feature that makes five of six faces indistinguishable.

4.5.2 Degenerate states in language

Attenuation in language is related to degeneracy. As we have mentioned, the process of attenuation in vocables involves a widening in the extension of a relation. Like degeneracy, attenuation is a process in which specific features are lost, features that allow vocables to be described as *semantically easy* that is, features that are perceptually bound. As we project forward in time these features will eventually disappear and so will specific vocables. But in the process, attenuation has for effect the structural transformation of some types of vocables that become *semantically difficult* - perceptually emancipated.

The relations associated with lexical vocabulary are bound to perceptual cues that are relatively restricted in their extension. Lexical instances of a vocable relate objects of similar type because lexical vocabulary rely on perceptual cues to be used. That is, the environment defined by perceptual cues also defines the instance of lexical vocabulary. Similarities in the perceptual cues of various environments define the kind of relations all instances of a lexical vocable will have in its extension. For example; *have* in a relation of ownership is defined by the perceptual cues that are available for that relation. *I have a dog, I have money, I have a book*, are all example of a relation of possession that rely on perceptual cues. Once all the instances of such a relation can be accounted for, or if the extension of a relation can be accounted for by a rule, then the extension of a relation of a lexical vocable is maximally extended.

Attenuation can also be considered in terms of energy level. In Chapter Three we have described lexical vocabulary as more efficient because of its dependence on perceptual cues. The effects generated by lexical vocabulary require less energy because it is environment bound. That dependence also defines the level of energy to be preserved as new objects are included in a relation. Therefore, the similarity relationship between instances that will be included in the extension of the relation of a lexical vocable will tend to be restricted to a certain type such as nominals and nothing else. Therefore the only configuration allowed for a lexical vocable is one that includes only one type of vocable in its relation, confining

the extension of the relation to one category.

We should mention that this is also true of kinds of attenuated vocabulary, such as *true*, *indeed*, *mind*, *belief*, and so on, cases in which uses cannot be described as lexical but also not as functional either. This type of vocabulary may not be as efficient as lexical vocabulary but it has not crossed over into a new syntactic role.

Relations, in functionalized vocabulary are far less restricted because of the absence of perceptual constraints. In some cases, some occurrences of vocables will emerge that seem related to some lexical vocabulary but have in fact become structurally changed and are now used functionally. The extension of the relation of a functionalized vocable is different from the one for its lexical ancestor. Moreover the objects that a functionalized vocable connects are potentially infinite in type because a functionalized vocable can be attenuated to a point at which the perceptual cues that were associated to its early uses, have mostly disappeared. As a consequence of the loss of perceptual features, its new function is to coordinate syntactically other vocables such as verbs and adverbs, prepositions and sentences, etc. The instances in the extension of a relation that define a particular functionalized vocable may include many categories of syntactic constructs, unlike the extension of the relation of a lexical vocable.

We consider these instances to be in statistical *equivalent states* because the *energy level* feature associated with the effects produced with the use of specific functional vocabulary is similar whether it involves a relationship connecting adverbs, prepositions or sentences. Because of this, functional instances of a particular vocable are more likely to occur than its lexical ancestor. One example of this phenomenon, described in Chapter Three, is the dynamics of *scope evasion*. It is the relative similarity between the use of certain instances of specific functional vocabulary that leads to the unnoticed scope evasion of this vocabulary. The consequence is the widening of its extension to include many different categories of syntactic constructs. But the effects produced by functional vocabulary are equivalent across most of its functional instances. We do not think this is true for lexical vocabulary. The extension of lexical vocabulary will include instances that produce different effects which will be associated to different levels of energy.

As we have mentioned, not all vocables become functionalized. The deciding factor is in the frequency of use of a vocable. Vocables must reach a certain level of attenuation before they can become functionalized. The level of attenuation is a function of the number of instances that the extension of a vocable can hold and that is determined by the number

of environments in which a vocable can find itself. So we can assume that vocables that are frequently used across many different environment have a greater potential to become functionalized as compared to vocables that are seldom in use, such as proper nouns. Based on this fact, we say that attenuation levels are unevenly distributed across vocables. In comparison, this is a somewhat unusual case for degenerate states in a physical system. Degeneracy levels are usually even for all constituents, however there are some cases, such as defects in a crystalline structure, that could be compared to uneven degeneracy levels in a physical system.

4.5.3 Structural changes and degeneracy

Degenerate physical states are often associated with structural changes in systems. Physics uses the concept of degeneracy in the modeling of the dynamics of state changes such as solid/liquid/gas.

In a two-state model, say, liquid/solid, degenerate states are favorable statistical states for the constituents of the system to adopt. These are offset by an energetically favoured state that constituents will tend to be in, say, the case of a solid state. This bias, generated by the energetically favoured state, will hold the system in this most probable state unless some condition, such as the energy available to the system, is modified. An example is water molecules that tend to stay in a crystalline formation (ice) in a sufficiently cold environment. As more energy is introduced in the system, such as a heat increases in the environment, a global shift in the system is more likely to occur. Interestingly enough the liquid/solid transition takes place suddenly at a critical temperature; an equilibrium system³ is solid slightly below $T_c = 0^\circ$ in the case of H_2O , and liquid slightly above. Liquifaction or crystallization occurs suddenly.

This kind of radical shift in the state of a related system variable is referred to as a *first-order phase transition*.

We think that a model capable of *first order phase transition* can be constructed using lexicality and functionality as states.

As we have discussed, lexical vocabulary, defined as more efficient requires little energy to use. In chapter three we argued that functionalized vocabulary is less efficient; that is, it requires more energy to use but does more work and as such, can propagate further.

³a system that has had sufficient time to settle to a stable state

Lexical vocabulary provides a bias in the system as it represents the lower energy state. In order for the system to change structurally, energy must be made available to the system. In our language model, it is the effort involved in linguistic transactions that increases or decreases in a population of language users. An increase in effort has for effect an increase in the level of linguistic activity that will favor the propagation of vocables. We imagine that constraints such as a growing population or an increase in the sophistication of the coordination of human activity are responsible for an increase in linguistic activity. The natural tendency for lexical vocabulary to remain lexical, because its ease of use is greater, and the tendency for vocables to attenuate, increasing its portability but diminishing its ease of use, generates a tension in the dynamics of the system. If enough linguistic activity occur in the population of language users, this tension will influence how structural changes happen. Given these conditions, a *first order phase transition* is a likely eventuality. This means that a lexical vocable, or more likely an attenuated vocable, can suddenly acquire functionalized instances in a very short timespan. This is a very particular instance of what may happen in the dynamics of language, but we assume that changes in the structure of vocables can occur in several different ways. We can also imagine changes occurring in a more progressive way.

There are at least two kinds of *phase transitions*: *first order* and *second order*. A second order phase transition is typified by a gradual change in structure or order. We will give an account of both phase transitions because their descriptions are dependent on each other, also, both accounts are relevant for a physical description of the behavior in the structural change of vocables.

4.6 Phase Transition

We present a basic background on phase transitions.

Physics describes many emergent features as the result of *phase transitions*, that is, matter transforming, according to a certain order parameter, say, from one structural spatial state to an other as in the case of solid/liquid/vapor. Phase transition occurs in large composite systems. These phenomena are emergent features, emergent because they are observable at the macroscopic level and not at the constituent microscopic level. Phase transitions have been described for centuries but the consistent microscopic description has only been developed in the past three decades [3]. The microscopic description of

phase transitions, though deterministic in its model, accentuates the emergent character by articulating clearly the conditions under which macro features emerge. Experimentalists have also observed that disparate microsystems have nearly identical macrolevel behavior for certain aspects of phase transition [3]. It is within this perspective that the language of critical phenomena can apply to the dynamics of language.

But first let us describe phase transitions in the context of physical phenomena. There are two basic kinds of phase transition, first and second order. We have already established the nature of *first order* as a discontinuous phase transition, while a *second order* is a continuous phase transition. In the condensation of gases the phase transition can be discontinuous under low pressure but continuous at a *critical point*. The *critical point* is the specification of macro variables at which the transition occurs. In the first case, the system changes its state from one phase to the other in a spontaneous way, that is one phase is replaced by another with an obvious interface between phases at a critical point. The second case describes a gradual change from one phase to the other such that there is no distinct phase, at a critical point, between liquid and gas. Phase transitions describe systems going from a disordered state to an ordered state or vice versa. Order in a system is a macro feature that has been characterized by an *order parameter* as introduced by Landau [51]. Order is a macroscopic variable that describes the variations between macro states, that is, a finite value in the ordered phase and zero in the disordered phase. The ordered phase usually occurs at low temperature and is destroyed at high temperature such that the order parameter vanishes. The order parameter is sometimes difficult to establish; for example, order in the case of the liquid-gas transition is based in the difference of density.

Here are various physical examples. Phase transitions occur in many material and describe phases other than the condensation of gases or the freezing of liquids. Iron is a good example: the ferromagnetic ordered phase is below 771° and is characterized by a net spontaneous magnetization that produces a magnet bar. Above 771° , iron is in the paramagnetic phase which does not display macroscopic magnetization. Another example is the order-disorder phases in binary alloy such as brass. The copper and zinc atoms occupy alternate sites in a regular lattice in the ordered phase of brass and arbitrary ones in the disordered phase [3]. We will elaborate on the modeling of phase transition in the *Ising Model* chapter.

4.6.1 Phase Transition and Language

Language, although not spatially defined, has a definable structure that demonstrates transition-like behaviour. We suggest that language, like natural materials, can undergo structural changes as the result of phase transition behavior. The structural change from lexical to functional vocabulary has features that resemble structural changes that occur in materials. Structural changes in vocables are syntactic and obviously not spatial but ultimately we assume that these changes have a correspondence in synaptic functions which are spatial in character. At some level of description, our theory of change in syntactic organization is compatible with a physical theory of behavior in synaptic organization. The comparison is of importance since it demonstrates that language shares crucial features with many other physical systems. First and second order phase transitions are striking phenomena that explain spontaneous state changes in multi-constituent systems. They are well understood phenomena that are formally tractable. Also, an explanation that includes the dynamics of phase transitions for a description of change in language may mirror fine grained observations in the dynamics of synaptic functions.

William H. Calvin [11] also describes a model for language involving some first order phase transitions. Calvin has developed a model that suggests, at a meta-level, how schemas are generated within synaptic functions, how concepts, abstractions and metaphor are formed and sustained through a replication process. He also suggests that partial perceptual cues generate competition between schemas and that resolution is reached through first order phase transitions. Calvin's model will be explored in more detail later in this thesis. It is possible that the fine grained description of phase transitions in the dynamics of synaptic functions may have some influence in the generation of linguistic structures within individuals. It also follows that such features may be reflected at a coarser grained level of language evolution in population dynamics.

First-order phase transition is of particular interest because the consequence of the process is irreversible. The irreversibility of the process, that is, the energy required for the transformation, dissipates after the transformation has taken place so it is non-reusable, much like friction in an engine. The transition of lexical to functional vocabulary can be viewed as irreversible as well, which supports the notion that the structural change in vocables may occur through a first order phase transition.

4.6.2 Phase transition and the evolution of language

The evolutionary viewpoint is, so far, poorly defined. It is useful to say that the description of evolution of language is complex beyond the crude description of general movement from pre-linguistic to proto-linguistic to linguistic behavior, as we have hinted earlier in this chapter. In the introduction, we have defined proto-language as a language without syntax; this does not offer a very detailed description, but even in descriptions of present languages, the categories that define syntax are also ill-defined. Linguists do not all agree where boundaries should be set: Is it between types of words? Types of morphemes? Types of phonemes? Should it be according to semantic types or physical types? Is a table not always a table whether we are taking about a kitchen table or a table of content? Both expressions are used in very different environments but *table* in its morphology, prosody, and so on, is used similarly throughout. The assignment of such boundaries relies mostly on observation of present languages and rarely on the dynamics of evolution, typically because historical cues that could instruct us on how to classify vocables, based on an evolutionary dynamics, are not readily available, and are at best partial.

Our historical account of dynamical changes is weak. Though we may assume that such boundaries exist, a detailed description of where to discretize along the history of language evolution will not be furthered by adding or inventing similar boundaries such as say, *proso-linguistic*. That is, any assumptions on our part that rely on a overly crude account of historical linguistic evolution will not further our understanding of it. There are too few distinctions with this description, mostly because that part of the history is lost. There is no historical physical evidence to support speculations about what prompted the emergence of a proto-language and none about what a proto-language is even like. What we are left with is the chasm between what we know about a relationship between pre-linguistic and linguistic capacities. It is not clear how many millennia are involved in the history of language evolution, but we can assume that an innumerable number of transactions of many types have occurred, transactions that have led to linguistic innovations. These types of transactions have belonged to several classes that would have blurred the distinction between pre-linguistic and linguistic activities. Those transactions that, in the present, would not be considered linguistic activity, could include hunting and gathering activities, commercial transactions, group vocalizations, rituals of all sorts, or any kind of coordinating activities that promoted the specialization of fine tuned motor function, laryngeal transformations,

and specialized neural functions.

Physical processes offer some non-arbitrary source of theoretical motivation. For a dynamical account of language evolution that would lead to a classification of linguistic and related non-linguistic items, we can study in detail the dynamics of physical systems in hope that they will point to some place or time in history where a possible description of the boundaries between pre-linguistic and linguistic activities, and between types of linguistic items, can be applied. But before we can place boundaries we must have a theory. In this respect the particular effects of phase transition may give us some important details of how to describe transcategorical changes in the evolution of language from non-linguistic ancestors. Again, we focus on the dynamics of less dramatic items; the transcategorical changes from lexical vocabulary to functional vocabulary, but let us remember that the types of transaction that occur in human interactions have less-than-clear boundaries. So we can confidently say that the dynamics in present linguistic interactions is descendant and contingent on the dynamics of past linguistic and non-linguistic transactions.

4.7 Self-Perpetuating

The process of language evolution is self-perpetuating. What we mean is, as a species, we do not intend to pass on language; we just do. The most important consequence of attenuation in vocables is the promotion of linguistic behavior. Somehow, as we inherit it from previous generations of users, some of the history is lost and forgotten and somehow the *residuals* give rise to more linguistic behavior. In the loose terms of the first law of thermodynamics; nothing is lost and nothing is created. Thus we can say that we reuse all that we already have in our linguistic interactions and as it becomes more and more attenuated we have more building blocks to generate more linguistic behavior. Moreover, as a language becomes older, its constituents are further along in the attenuated process so linguistic combinations are significantly increased as to involve complex sentences. But most of language is used in reference to something else, that is, without being conventional, linguistic behavior can coordinate and generate actions much beyond simple deixis can do. The fact is, linguistic behavior is an emergent feature of the attenuation process, much like flocking behavior is an emergent feature of a group of birds flying. But linguistic behavior is more like flocking behavior in which, somehow, the birds have been forgotten. If we were to copy successively a video of flocking birds, each generation would lose some information

possibly leaving a blur in the place of birds, nonetheless, the blur could still trace the patterns that are emergent.

Similarly, with attenuation, we have the behavior but we have forgotten what gave rise to it. However, attenuation differs from our example because the resulting linguistic behavior furthers the process of attenuation which in turn gives rise to more complex linguistic behavior.

Chapter 5

Phase Transition Models in Biology and Economics

We are not the first ones to find that statistical descriptions of systems are applicable to systems other than the ones studied in physics. Moreover, applying statistical physics models to systems outside of physics is something that the fields of economics and biology have been doing for a while. In particular, we report the use of models that explore phase transition in systems studied by those fields.

5.1 Biology and Phase Transition

Emergent behavior, features observable at the level of an entire system that cannot be reduced to the description of their smaller parts are a common occurrence in many fields of study. This is true of biology where phase transition has been suggested in the explanation of evolutionary novelties.

Traditional inquiry in the field of biology is split between microbiology, which studies genotypes that define physiological makeup, and macro-biology which concentrates on phenotypes or the typical behavior of specific species. It is assumed for both parts that phenotypes are resultant from the accumulation of micro-changes in genotypes, that have survived environmental constraints. Genotypes and phenotypes are described as discrete units that take small adaptive steps individually from the whole organism. Interconnectivity is viewed as conflicting constraints that may hinder or help evolutionary steps but are

otherwise incidental. Microbiology and macrobiology are essentially two different sciences that have a difficult time describing the connecting steps between microscopic and macroscopic. Many macroscopic facts cannot find a convincing explanation from the *resultant* theory. The idea of continuous and gradual change is incompatible with the observation of certain stable traits in many life forms and the disparate rates of evolution. Comparative morphologists have found that a large number of species share a small number of basic anatomical design, *baupläne*, that differentiate various animal *phyla*. All these *baupläne* have appeared suddenly in the Cambrian Explosion, shortly after the appearance of multicellular organisms, 500 million years ago[3].

Many vestiges are still present long after the functional usefulness have gone. Somehow they have escape natural selection. Embryos of birds and mammals still have gill arches, a feature that has not been in use for 400 million years. Other evolution features, such as sexual reproduction, are puzzling. It is an expensive feature providing only half of the fitness that asexual reproduction provides; asexual reproduction has not proven to be an inferior strategy for the propagation of specific genotypes [17]. It does not explain why there are, for example, specialized flies that have evolved within a thousand years in the presence of non-indigenous banana crops in Hawaii, and why biological structures such as snapping turtles and alligators have changed little since they first appeared hundreds of million years ago. Fossil examination does not agree with the *gradual evolution theory*. Eldredge and Gould have showed, in 1972, that morphological characters are stable for long periods of time and dramatic changes occur in short bursts.

These observations are reminiscent of *emergent feature* theory developed in statistical physics. Sudden changes in the state variable of a physical system that give rise to emergent features are well understood phenomena in statistical mechanics. However this particular approach is not widely used in the description of evolutionary steps. Kaufman [38] and a few others are suggesting that many of these incongruities in empirical facts can be described using models of multi-bodied systems.

5.1.1 Kaufman and percolation

To make a convincing argument, Kaufman takes the super macro-biological point of view which involves, not the study of *fitness* in specific phenotypical traits, but fitness based on the interrelatedness of potentially infinitely many phenotypical traits. This is not a new point of view; Dawkins[17] has suggested, from this point of view, several strategies used by

genotypes and phenotypes that influence each other through a complex ecological web. The novelty of Kaufman's approach lies in his wanting to demonstrate that the interconnectivity of phenotypes give rise to emergent features that constitute evolutionary novelties. He suggests two computer models, one of which is not unlike the *Ising model* that is discussed in this thesis.

The *random nk boolean network* treats an organism or genome as a composite system with n constituents that are regulated through k other elements. A constituent n has two possible states determined by its relationship to the states of its k connections, at a preceding moment. This kind of model is usually referred too, in physics, as percolation. Percolation leads to a second order phase transition. Both concepts will be described in more detail in subsequent chapters.

Kaufman suggests that, in the face of limited resources, a strategy for minimizing cost and maximizing benefits can imply spontaneous order that will stabilize particular phenotypical or genotypical features in organisms. This point of view implies that,

1. profound structural changes may occur in fairly short periods of time, compared to the usual assumption that novelties are resultant and require lengthy time scales to occur,
2. that novelties are not necessarily the result of selective pressures of the kinds described by Darwin,
3. that the implication of infinite connectivity between and within organisms increases the complexity of systems to a point in which it is difficult to establish any kind of discreteness in a description of particular phenotypes and genotypes.

Moreover *phase transition* seems to offer a description that convincingly reconciles micro-observations in the study of genotypes and macro-observations in the study of phenotypes as it describes how phenotypical traits can influence micro-structures in fundamental ways and how the interactions in micro-structures give rise to radically different features from the ones that define any individual micro-constituent.

Phase transitions may have been first adequately described in statistical physics, but the feature of spontaneous structural changes in overall systems broadly applicable. Kaufman understands its repercussions in biological fields; however biology is not the only field to benefit from this approach. Economics also turns to statistical physics to model economical trends.

5.2 Economics and Phase Transition

Economics is an old field that has understood the complexity of its subject for a long time, but, like biology, the chasm between micro- and macro-economics is difficult to bridge using traditional theories. Macro-economics attributes variations in market trends to few variables that affects the entire economy while assumptions about the influence of individual is determined by the conflict between need to work and desire for leisure. Micro-economics studies the situations and disposition of individuals and how choices of individuals are reflected at the level of the market. As in biology, macroscopic and microscopic observations are not entirely compatible, so many researchers have turned to statistical models used in physics to reconcile approaches.

Inflation, unemployment, economic growth, balance of trade, fiscal and monetary policies, are the variables considered by macro-economics. The statistics taken from these variables show how the economy is balanced from interactions between these features, and in good times there is a direct linear correlation between macro-variables and the set of micro-interaction, but during hard times this relationship is no longer linear. Results from micro-analysis does not support the claims of macro-economics. Moreover, macro-economics fails at explaining events behind an economy's constant cycle of thriving and declining. The concern of macro-economics is the form of behavioral relations, whether its slope is positive or negative. Behavioral relations are functions describing production, consumption and investment. Equilibrium is a question of market clearing. Markets are subject to supply and demand and are interrelated since they share sets of variables.

There are many models developed in the study of micro-economics. Micro-economics is busy with the concerns of human individuals in a population and their impact on economic trends. There are two main approaches; new classic economics reduces macroscopic phenomena to incremental shifts in microscopic constituents' states, and the new Keynesian that focuses on microscopic behavior consistent with macroscopic phenomena. New classic economics postulates that rationality is the driving mechanism in individuals and that this is how a utility function can be defined as based on the rational choices of individuals who will invariably make choices that will maximize their own happiness. A utility function is a set of commodity bundles that have been assigned rules of desirability. It is also the objective function that a household tries to maximize. New classic economics does not allow for structurally emergent features such as unemployment during the *Great Depression*.

New Keynesians reject the idea of idealized household that has no internal structure for a more complex analysis of interactions within households that can lead to emergent structural changes in large scale economics. They consider explicit relations and realistic competition that occur between firms and other institutions and delve deeper into interrelationships between constituents of the economy.

5.2.1 Economics and the Ising model

This kind of approach relies largely on models that have been developed in the context of statistical mechanics. The Ising model is also prominent in the study of economics in explaining phase transition or spontaneous structural changes in multi-factor systems. This model has been used in the study of dynamics in the propagation of information for marketing purpose. The field is usually referred to as the *Social Percolation* model. The team Abhijit Kar Gupta and Dietrich Stauffer [27] are some of the researchers that are using the Ising model to understand correlations in social structures. The Ising model in its traditional physical uses involves variations in energy levels for constituents and temperature levels, that influence the entire system. The context of economics renames these variables to mirror features important in the study of social impact. *Word of mouth* in neighbour interaction and the effects of *media* in a population result in percolation and phase transition behaviors in the propagation of information, behaviors that are typical of the Ising model.

5.3 Neuroscience and Phase Transition

Neuroscience recognizes the problem of the integration of components into a functional whole.

Since 1964, the work of W.J. Freeman has promoted the use of non-linear models in the study of neural interactions in the neocortex, more specifically in perceptual processing involving the olfactory bulb. From his electroencephalogram (EEG) model, Freeman observed that the wavelike process revealed linear and near-equilibrium dynamics. More recently, many scientists involved in neuro-biological research have noticed that, despite extremely non-linear behavior at the microscopic scale, synchronicity emerges, involving many sub-components, in the spatio-temporal context of movement. Synchronicity dissipates as soon as the action is performed. These results hint at a system on the edge of chaos that can switch between states. This would indicate that components can be involved in various

tasks based on the type of coupling and most areas are not strictly dedicated to one task.

This type of research has led to the use of computational models that are reminiscent of the ones developed in statistical mechanics, namely the Ising model, to explore the impact of phase transition in artificial neural networks (ANN).

5.3.1 Neural networks and the Ising model

Traditional inquiry suggests static representations for structures of language despite the obvious facts that language-processing involves historical aspects. Researchers in the field of artificial intelligence (AI) are very aware of dynamical and temporal connections involved in the production of natural language models (referred to as performance models). But the grammatical theory that underlies the competence model and provides the context of comparison for the performance model, does not take dynamical features into consideration. The introduction of a statistical approach to compensate for the failings of static rules has produced positive results; however, context evaluation must extract features defined by conventional grammar. It is inadequate in not acknowledging the dynamically generated features of natural language use, features that co-opt pre-adapted forms for new uses, and thus put them beyond the reach of any fixed grammar [52].

The first model to involve non-trivial dynamics and predictable useful behavior in neural networks (NN) was introduced by J.J. Hopfield. He presented a model that was biologically probable and tractable to formal analysis. A neural network is a set of constituents, usually referred to as nodes, coupled according to some rules. These nodes are neuron-type units that are usually arranged as sets of input and output nodes. More conventionally, a layer of hidden nodes interconnect with input output nodes. These hidden nodes have weight that are adjusted to perform a specific computation. In all cases the elements are interconnected in such a way that the input of each node is determined by the state of some or all of the other nodes. The interconnectivity forms a whole that can perform useful computations. Moreover, these networks are arranged in layers that *feed-forward* its result to the following layer in performing more complex computation. This kind of NN does not feedback any of its results to achieve its computation. The XOR boolean algorithm originally created an insurmountable problem for NNs. Years later, the creation of *feedforward* layered networks ultimately answered the problem so that an NN could resolve an XOR algorithm. Many *feedforward* neural networks are used in modeling natural language production, however, with limited success. Hopfield's approach is significantly different.

The Hopfield model interconnects nodes with feedback, that is, each node serves as input and output. Additionally the nodes are weighted so that they can only be in one of two states. Mathematical and simulation analysis demonstrated that this kind of system evolves to a stable fixed attractor. Furthermore, by manipulating the link strength, it is possible to encode any set of node states as attractor states. The flexibility of the system allow for the encoding of various patterns.

The model was suggested as an explanation for the mechanism of associative memory. When the system is presented with an input patterns that are sufficiently close to the state space of the stored pattern the input patterns will invariably evolve towards the state space of the stored pattern. The input pattern may be incomplete or noisy. Nonetheless, the pattern will dynamically evolve, travelling from its initial state to a specific basin of attraction.

This application has also been used in word disambiguation, that is, in the automated process of differentiating the different instances of a vocable across contexts. The process is statistical not semantic and uses a network of Hopfield models [57].

Since the formal description of the Hopfield model is identical to an Ising spin glass ¹, the field of neural network attracted many physicists from statistical mechanics to study the impact of phase transitions on the stability of neural networks.

¹We define Ising spin glass in “The Ising Model” chapter

Chapter 6

The Ising Model

In the Physics chapter we have shown that we can co-opt some physics frameworks to describe the dynamics of language. To go further: Can the changes in the state of vocables (from lexical to functional) be specifically described in the language of the Ising model? This model has a two-dimensional solution,(the one we will use), and was discovered by Onsager in 1944. It is now a familiar tool of statistical thermodynamics.

Consider the developments of language as the product of the millions of individual linguistic transactions. These transactions convert the raw physical energy expended in the production of speech into minute neuro-physiological alterations in other members of the linguistic community, alterations that contribute to the shape of future linguistic interactions. This is the first step in correlating both the Ising model and a simple language model. This chapter will elaborate further along these lines, to demonstrate the relevance of such a physical model to a discussion of language.

If the underlying research is correct, then the eventual outcome of this process is that many vocables pass from a stage in which we typically can explain them by definition or ostension, to a stage in which virtually all of us can use them but none of us can understand them. This is the process that we refer to as *lexical attenuation*.

6.1 The Ising Simulation

The Ising model is a simple, well-understood model often applied in statistical mechanics because of its predictable behavior and its capacity to model complex behavior of various phenomena.

We have talked about the dynamics of language evolution. We have also stated that, overall, language is a fairly stable, slowly changing system. Specifically, across the history of language, the system does change very slowly but is stable (enough to be shared) on shorter time scales. We consider language to be a system in equilibrium (or quasi-equilibrium) and it is the local dynamics - linguistic transaction between few individuals - that we will study with the help of the Ising model.

The Ising model illustrates microscopic, short range local interactions - generally referred to as *nearest-neighbour* - that are influenced by macroscopic state variables such as temperature. It focuses on state changes at equilibrium that occur at critical points in the value of a state variable (such as temperature).

This is for the benefit of our physicist audience who are familiar with these kinds of models. Some of the terminology will remain less explicitly defined as it is not relevant to the scope of this research. Here is a standard textbook definition of the Ising model [62].

A simple model used in statistical mechanics. The Ising model tries to imitate behaviour in which individual elements (e.g., atoms, animals, protein folds, biological membrane, social behavior, etc.) modify their behavior so as to conform to the behavior of other individuals in their vicinity. The Ising model has more recently been used to model phase separation in binary alloys and spin glasses. In biology, it can model neural networks, flocking birds, or beating heart cells. It can also be applied in sociology. More than 12,000 papers have been published between 1969 and 1997 using the Ising model.

This Ising model was proposed in the 1924 doctoral thesis of Ernst Ising, a student of W.Lenz. Ising tried to explain certain empirically observed facts about ferromagnetic materials using a model proposed by Lenz (1920). It was referred to in Heisenberg's (1928) paper which used the exchange mechanism to describe ferromagnetism. The name became well-established with the publication of a paper by Peierls (1936), which gave a non-rigorous proof that spontaneous magnetization must exist. A breakthrough occurred when it was shown that a matrix formulation of the model allows the partition function to be related to the largest eigenvalue of the matrix (Kramers and Wannier 1941, Montroll 1941, 1942, Kubo 1943). Kramers and Wannier (1941) calculated the Curie temperature using a two-dimensional Ising model, and a complete analytic solution was subsequently

given by Onsager (1944).

To be more concrete, consider a set of N individuals arranged in a lattice. Each individual can be in one of two different states, say $+1$ and -1 . Let S be the space of all sequences or configurations

$$S = (s_1, s_2, \dots, s_N),$$

where $S_i = +1$ or -1 . Further, we define a function

$$w(s_i, s_j) = e^{-E(s_i, s_j)/(kT)},$$

where $E(s_i, s_j)$ represents the energy of interaction between two neighbours in the lattice, k is *Boltzmann's constant*, and T stands for the *temperature* of the system (in K). The probability of a configuration s is defined now as follows:

$$P(s) = \prod_{(i,j \in N)(i \neq j)} w(s_i, s_j) / Z$$

where Z is the partition function

$$Z = \prod_{(s \in S)} P(s)$$

Assuming that each individual can be in one of q states.

In its simplest form, the Ising model can be realized as a collection of *spin vectors*, localized on a one dimensional lattice site. The spins have discrete binary states. The energy of these spins is determined by the sum of its interactions with its nearest neighbour: one is an interaction value for sets of *nearby spins* on the lattice.

Individual constituents interact with neighbouring constituents. The states (spin-up or spin-down) of neighbours determine the amount of energy that a given spin needs to change its state. The energy available to determine this flip defines the notion of temperature. Temperature is related to the average amount of energy in the system. That variable is an intensive quantity of the system and increasing or decreasing it modifies the characteristics of the behavior of the system.

The two-dimensional Ising model shares the same properties as the one-dimensional model except for the fact that spins are localized on a two dimensional lattice site. Generally, both variations demonstrate second-order phase transition but by modifying a two-dimensional Ising model we can simulate a first-order phase transition.

We have defined both transformation types in the *Physics of Language* chapter but to clarify, let us describe further the nature of first-order phase transition. To start, a phase

transition is an emergent feature of a system, but there are several ways by which this transition can occur. A first-order phase transition is one of the ways by which a state transition can occur. The state space of a first-order phase transition (see figure 6.1) can be illustrated as a sharp delineation between the states of constituents. All constituents are coherent, either in one state or another. In an H_2O system this can be observed as a clear differentiation between a solid and a liquid state.

Figure 6.1 shows the meanfield solutions analysis for an interface model, that is, of a first-order phase transition. A meanfield solutions analysis is all solutions of the order parameter of a system. The red (or thin) curves demonstrate all theoretical solutions while the green (or thick) curve demonstrates the stable solutions, solutions that an Ising model would more-or-less follow as T increases. The straight line that crosses the graph illustrates the first-order phase transition. The state space of a second-order phase transition (see figure 6.2) can be illustrated as continuous between states, where constituents are incoherent; in a mixed state. Figure 6.2 illustrates the meanfield solutions analysis for a second order phase transition. Notice the abrupt drop in the stable solutions for the first-order phase transition and slow slope in the stable solutions of the second-order phase transition. The dash lines in the solutions illustrate the second-order phase transition. These figures illustrate that there is a clear boundary between states in the case of a first-order phase transition and no clear boundary in the case of a second-order phase transition.

6.1.1 First-order phase transition and degenerate states

To simulate a first-order phase transition in an Ising model, we first implement, microscopically, a *bias field* where one of the spin states is energetically favoured over the other. We also implement *degenerate* states that are statistically favoured by constituents.

Recall the dice example from the *Physics of Language* (chapter four). A die has six distinct faces - or states. But now imagine that five of those faces have one dot on it while one face has six. Though the probability for each face to land face up, once thrown, is equal amongst them, the probability of a one being face up is five times more likely. In this case five faces of the die have equivalent states and as such are degenerate. A similar concept can be applied to spin states. Though the probabilities for individual states to occur are still the same, the fact that five states are effectively identical introduces a bias in favour of the identical states. The faces of a die are akin to the states of spins. However, spins can have an arbitrary number of states and an arbitrary number of these states may be equivalent or

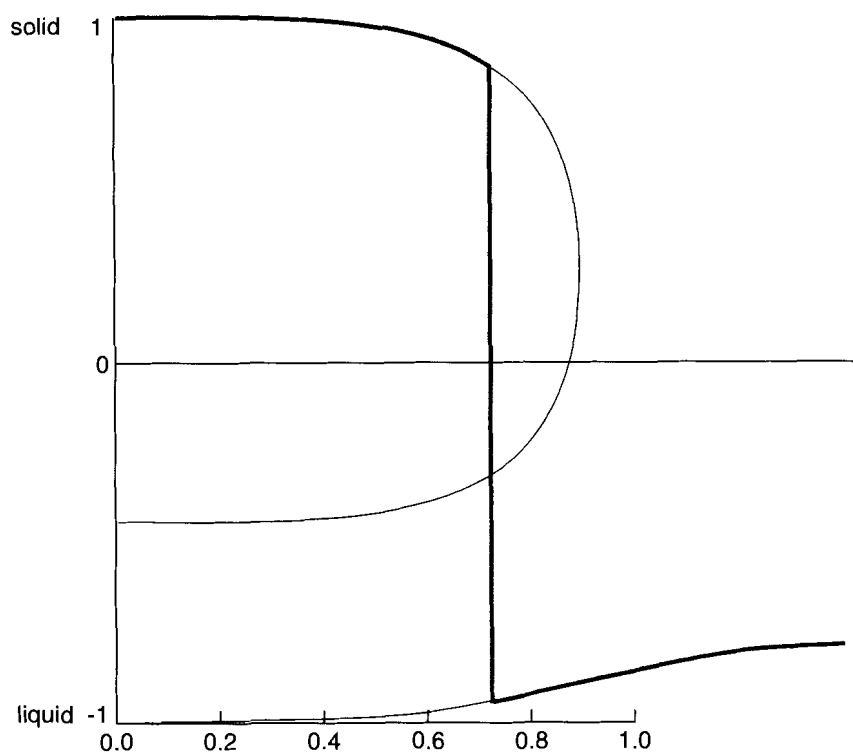


Figure 6.1: Meanfield solutions for a first-order phase transition. The red (thin) lines show all possible theoretical solutions for the order parameter, while the green (thick) line shows the stable solutions. Notice the abrupt change in spin orientation as the heat exchange increases.

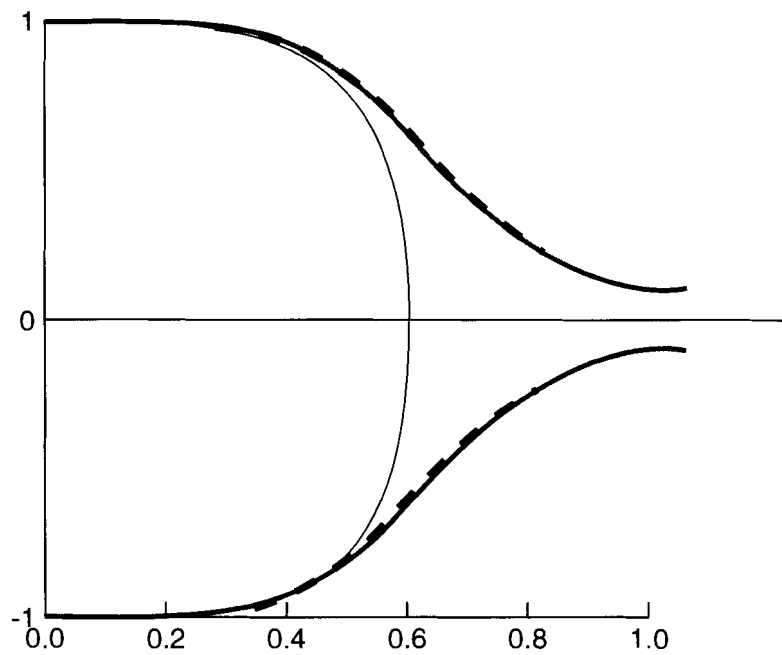


Figure 6.2: Meanfield solutions for a second-order phase transition. The red (thin) lines show all possible theoretical solutions for the order parameter, while the green (thick) line shows the stable solutions. The dash lines illustrate the second-order phase transition. Notice the slow change in spin orientation; as the heat exchange increases, spin orientation becomes distributed between up and down.

degenerate.

A competition between the *bias field* and the state of *degeneracy* is thus generated. On the one hand, the bias field is energetically favoured as it locks constituents in a low energy state through neighbour interaction, while on the other, degenerate states are statistically favoured. As energy is added to the system, local fluctuations in the state of constituents increase and influence the state of nearest neighbours. Degenerate states will be favoured despite the influence of the bias field. Constituents will hold on to their low energy states up to a critical point at which sufficient energy will be in the system that, statistically, constituents will tend to be in a degenerate state. This transition occurs suddenly.

6.1.2 Computational implementations

Several computational solutions may be developed to address a particular feature of the model. For example, there are many ways in which the energy available to the system can be implemented. One can perform a random draw to represent the level of energy available to all constituents in the system. Similarly a Maxwell [51] energy “daemon” may be employed which also distributes evenly, via a discretized transfer of the energy to all constituents. A variation in the approach may be used, and this is the one we have chosen for our model. We implement a Maxwellian demon lattice in our model, with one demon per spin site to allow for the concept of local temperature. This is referred to as a heat sink/source, since it gives up or takes in energy. The energy moves around via neighbour interaction. The energy is not randomly redistributed as in the case of a Maxwell demon or a random energy function.

Another useful computational implementation in a standard Ising is the *wrap-around* Cartesian grid as illustrated in figure 6.3. A wrap-around grid is helpful in reducing the constraints that arise from having boundaries in a system. The problem with boundaries is that they introduce circumstances that set them apart from the rest of the system. So in the case of our model, we define the interaction as depending on the nearest-neighbours. There are four nearest neighbours on a Cartesian grid. However, at the boundary of a square grid, a constituent situated on one of the corners will only have two or three neighbours hence changing the nature of the interaction. In a wrap-around grid all constituents have four neighbours because it is akin to a sphere and there are no edges on a sphere.

Several such computational strategies are used to create a simulation. The ones highlighted here are the most interesting ones. These, however, should not distract us from our

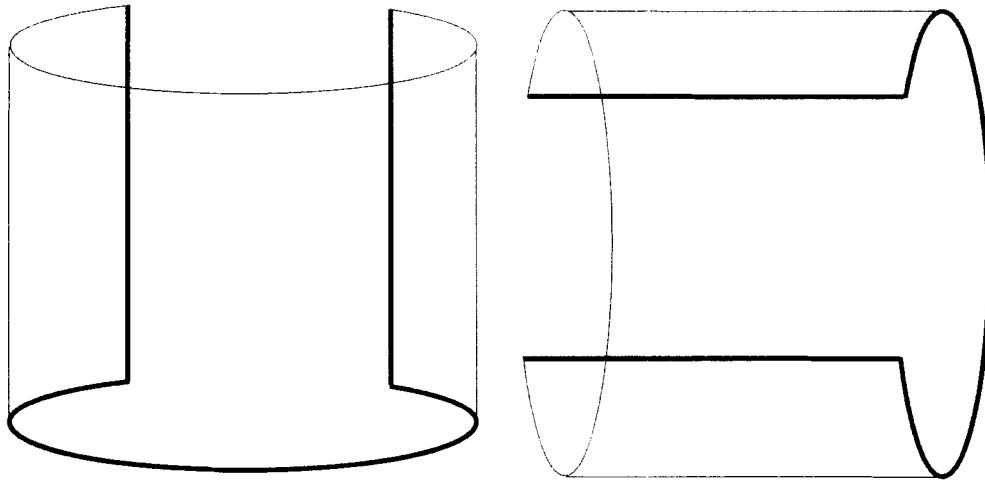


Figure 6.3: The Cartesian grid is wrap-around to reduce boundary conditions. Because the grid is square it is difficult to illustrate that corners are also wrapped so there is continuity between all edges.

primary discussion about the model and how the behavior of language can be described in its terms. There is a distinction to be made between a theoretical model and the computational implementations that can be used to *run* a simulation.

Again, in order to confirm that the computational implementations of our simulation are faithful to our model, we use a graphical representation to evaluate whether the simulation behaves appropriately. A grid of colored pixels can be used to display the states of the vector spins as binary rather than directionally. We can also display the Maxwell demon lattice as a grid to illustrate local temperature, which we have done in the case of our model. Additionally, the use of a graph is useful to illustrate the relationship between temperature of the system and the state of the system. In our model we include some modifications to the standard Ising model, modifications that we will describe in the next section.

6.2 Ising and Language

The Ising Model can be further refactored to model the dynamics of what we call *lexical attenuation* in natural language. Our model describes a possible history of the dynamics of a meta-vocable. This meta-vocable is a statistical representation of all vocables.

As in the standard model we use spin states to represent vocables on a two dimensional lattice. We also use a Maxwellian demon lattice to apply a local level of *activity* - the frequency of use - for each constituents. In addition, we introduce the concept of *attenuation* and we use it as a statistical bias. It is similar to *degeneracy* in that it can be introduced as a feature of a two-dimensional Ising model but with one difference: Usually a particular degeneracy value is assigned homogeneously to each spin. In our case we introduce a gradient of *attenuation*, that is, spins are assigned an increasing number of *potential attenuated* states, illustrated as a gradient. We will describe this process in detail a little later.

We use three Cartesian grids for our language model. The first grid represents the two dimensional lattice that hosts spins. These spins represent a set of instances or occurrences of one meta-vocable. A black (or dark) pixel for a constituent represents a functional state of the vocable while a blue (or light) pixel represents a lexical state. Figure 6.4 illustrates a typical representation of that grid. Notice the white line that travels the grid. This line highlights the ratio between functional (dark pixel) and lexical (light pixel) for all rows. A second grid represents the local potential for *attenuation*. A gradient from black to

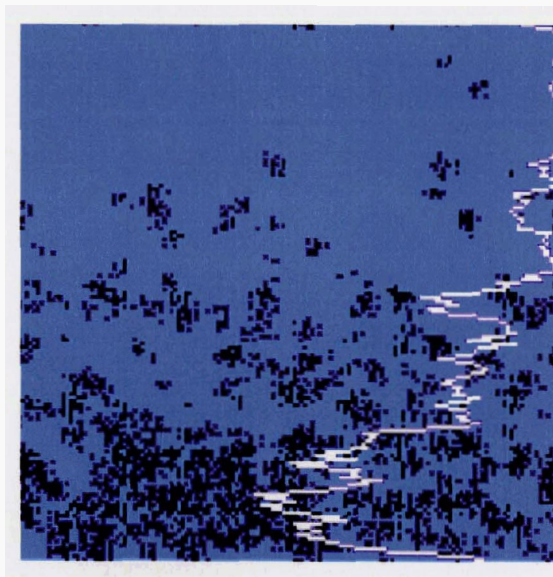


Figure 6.4: Vocabale (spin) grid. Vocabales can be in 1 of 2 states; light is lexical while dark is functional. The white line indicate the ratio of light to dark for every row.

white describes the state of attenuation each meta-vocable can achieve. The changing levels of attenuation can be thought of as actual real time-line (versus simulation time) in the

lifespan of a vocable. Early instances are found in the dark areas of the gradient while later instances are found in the light ones (see figure 6.5) This gradient affects the critical point at



Figure 6.5: Attenuation grid. The different shades represent degrees of attenuation; in this case constituents can be attenuated up to a 7 to 1 ratio.

which a first-order phase transition can occur within each area for each degeneracy. For our physicist audience, let us remember that, given a Hamiltonian $H_i = -\sum_j s_i * s_j - s_i \Delta/2$, the field value is Δ (flipping a spin changes the energy by $2*\Delta$ if spin values are $+1/-1$), and attenuation is δ . Because δ changes along the length of the attenuation grid, we observe (figure 6.6) that the critical point (green or dash line) is pushed inwards as the level of attenuation augments. Consider a simulation where there are mixed attenuated states for constituents. The attenuation gradient is up to three states. No state of attenuation is represented as black while attenuated states from one to three is represented respectively as dark, medium and light grey (as shown in figure 6.6). The Meanfield solutions in figure 6.6 illustrates that some attenuated constituents remain lexical - x at the top of the curve - at a sub-critical level of activity. Once the level of activity in the system reaches the critical point of state change for particular states of attenuation, constituents become functional - x on the bottom curve. In this graph we illustrate the critical points at which constituents will become functionalized. Notice that only constituents that have three states of attenuation

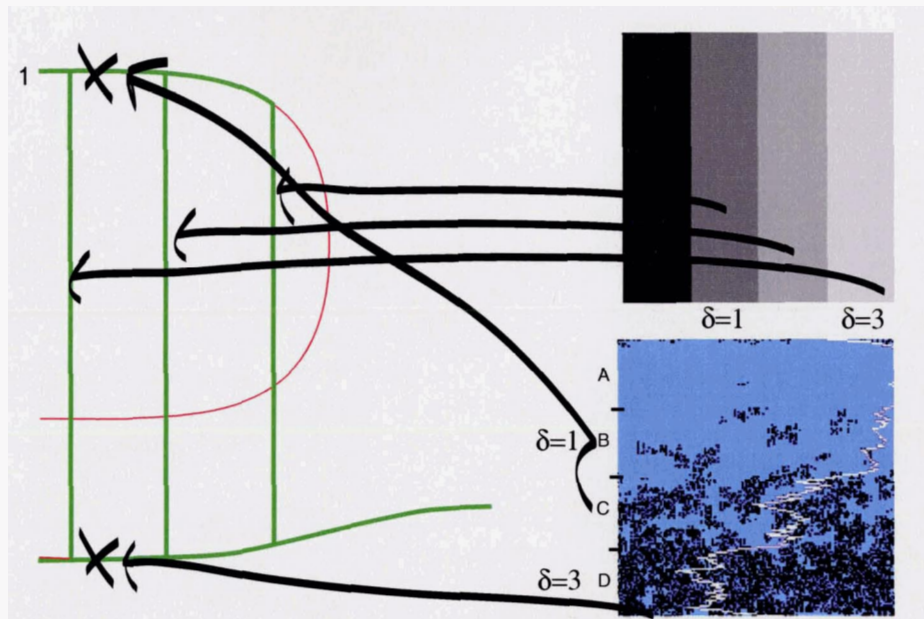


Figure 6.6: This graph illustrates how the critical point (green or thick lines) at which a first-order phase transition occurs moves inwards for constituents ranging from 1 state of degeneracy to 3 states of degeneracy. The x's indicate the state of constituents. Those with 3 states of attenuation have become functionalized because the activity level is past their critical point. The other constituents remain lexical because their respective critical points have not been reached.

have become functionalized because their critical point has been reached.

As in the standard Ising model, we also use a bias that favours a particular state energetically. Lexical vocables tend to stay lexical for reasons energetically similar to the reasons why ice tends to remain in its solid state until environmental constraints are such that a liquid state becomes favoured. This feature, however, is not graphically represented.

A third grid, corresponding to the Maxwellian demon grid, illustrates an *activity* rate. The energy driving attenuation is the physical effort expended in the actual uses of the vocables in speech. The temperature variable used in the standard Ising model finds its equivalent in the *activity* variable of our model. As the *activity* level is increased, the likelihood for vocables - spins - to become *attenuated* will also increase. Augmenting the *activity* value represents an increase in the use of vocables within a linguistic community. In the real world, linguistic activity increases with a growing number of linguistic participants in a population; however, since our model has a finite population - constituents, it is solely the frequency of use of a vocable that defines the *activity* variable. For example, given a system where there are only two linguistic participants, we can expect that some vocables can become very attenuated by the mere fact of overuse.

Consider the case of technical language used by small groups of people in the context of developing new technologies. As the use of such vocables becomes attenuated, we lose our capacity to say what it specifically means; we think that technical jargon is a product of similar dynamics. Figure 6.7 illustrates local variances in the activity levels of each constituent. Each spin site has a local *activity* value assigned to it. In the application to language, a high *activity* value at a spin site can be thought of as representing extensive use of that vocable by a single user. The underlying theory suggests that higher usage eventually entails a higher rate of attenuation. So a low distribution of *activity* in the system can be taken to represent comparative lexical richness.

6.2.1 Ising for language

Now we establish the rules by which the interactions of the model are defined. We are interested in the change from lexical vocabulary to functional vocabulary. We propose that this change can be simulated as a first-order phase transition and that enough of what we know about language will find descriptions in these terms.

Local interactions lie in nearest-neighbour interactions in which energy is transferred from constituent to constituent. Standard nearest-neighbour interaction is described in the

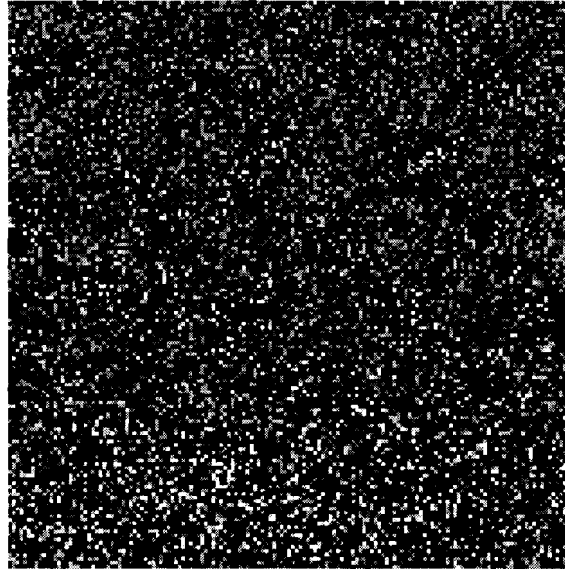


Figure 6.7: Activity grid. This is a "Maxwell" demon grid. A random level activity is assigned to every vocable constituent in the system. The overall average is the activity in the system. It is equivalent to the level of linguistic activity that can occur in a population of language users.

following; each individual spin is in contact with the spins on the four sites surrounding it and is influenced by them. When most immediate neighbours are in a given state, it is energetically favorable for a given spin to be in a similar state.

A phase transition will occur only when a critical point in the activity of the system is reached. This critical point of activity allow for attenuated states to be statistically favoured.

Consider the functional use of *have* in its auxiliary role. At a stage of the history of the language in which interlocutors use *have* as a relationship of possession, its auxiliary role would not generate the neurophysiological effects appropriate to that functional use. In every-day cases, less dramatic innovations - even if they cannot be fully understood - nevertheless increase the extension - attenuate - for the neighbours' continued lexical use, even if they are not yet in a position to use *have* in its functional role. Eventually, a critical level of attenuation will be reached. As this dynamic prevails in all neighbourhoods of the system, we observe a sudden change in the state of the system. Functional uses will have come about. The sudden change occurs because, initially one of the states, is energetically favoured by the system. In low activity, the influence of nearest neighbours lock the state of constituents by forcing each other to remain coherent. As the activity increases in the system, the rate of change in the local states of constituents augments. Because of our attenuation bias, attenuated states will be statistically favoured and eventually the energy bias will lose its hold.

Obviously there is no solid state for language but there is a state in which language cannot propagate any further. Like a crystal, the structure of word associations is semantically circumscribed and, as such, easy to understand but limited in its use. Attenuated language is much like an amorphous state, in which vocables are in a structurally loose state: Easy to use but difficult to understand. The structure of liquid water is correspondingly difficult to describe at a molecular level. Attenuated language is syntactically circumscribed and has lost most of the perceptual connections present early on. This phenomenon is a direct consequence of its propagation. The mere fact of use produces a degree of attenuation, howsoever minute. And the loss of perceptual circumscription that would be required for stability in a language, itself promotes new uses, and therefore greater instability.

Changes in the rate of attenuation can be occasioned in a number of ways; certainly through an increase in the biological population of its users; certainly by changes in the rate of environmental changes such as average age of the users and their longevity, by

immigration, by technological developments and so on. The model is not dependent upon the specifics of propagation; it gives only a mathematical model of behavior at and near equilibrium.

6.3 The Experiment

We set out to confirm that the Ising model can adequately represent phenomena that occur in language evolution.

The methods of physics' formal inquiry usually require a specific approach that goes beyond heuristics. Data is gathered in a systematic way. Running an Ising model simulation requires us to set state parameters to a certain value, let the system relax into its likeliest state under these set conditions and wait for it to settle. Once this is done, sample constituents, that are considered representative of the entire system, are chosen for formal assessment. The state space of these constituents is mapped out so a formal expression can be stated that describes the state path of these constituents. Figure 6.8 illustrates the state space of a typical Ising model that simulates a first-order phase transition. This graph illustrates a clear boundary between the two phases. The area beyond the the interface is an area of incoherence where spins have so much energy that they are in constant disordered states. In language, this may be akin to a room full of people talking and the noise level is such that no vocables are distinguishable.

Because we are modeling language and not H_2O behavior, we take a more heuristic approach. In the modeling of the phase of H_2O the values of state parameters refer quite literally to physical states. The mathematical temperature variable of an Ising model is representative of a similar measure used in the observations of real systems. For example; in the simulation of state changes of H_2O , it is relevant that changes occur around $T_c = 0C$. In our simulation, it is not. It only matters that at some critical point there is a state change. It also matters that we can simulate that change as a first-order phase transition. In our case, the mathematical measure of the activity variable does not find a counterpart (that we know of) in the real system of language. Moreover, the use of such a formally well understood model relieves us from solving it. At this point in our research our focus is in the observation of the simulation and assess the relevance of the model to language.

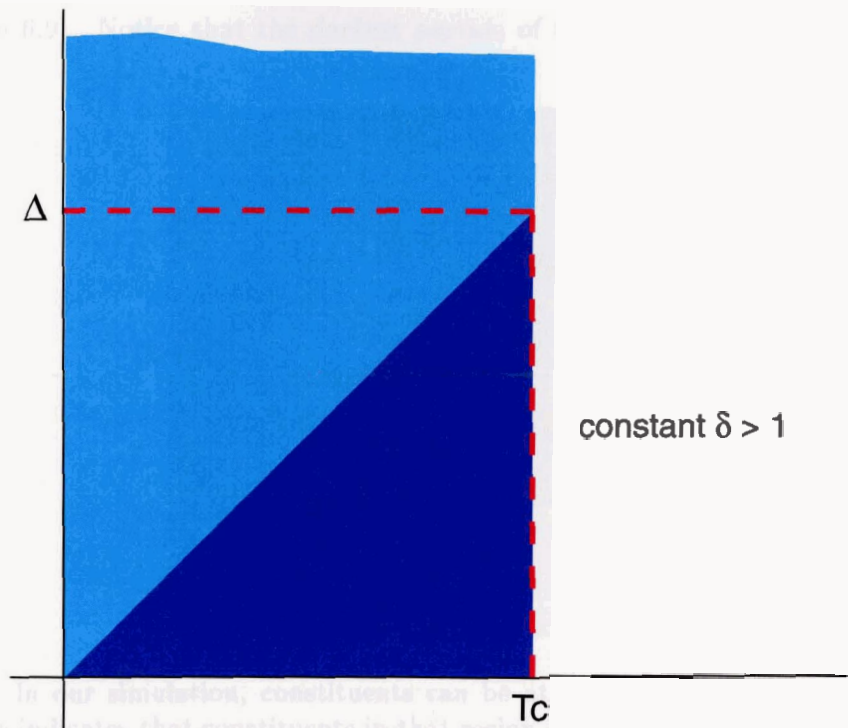


Figure 6.8: State space of a two dimensional Ising model with a first-order phase transition. The light represents the solid portion while the dark represents the liquid portion. The dashed lines illustrate the critical point and the critical temperature at which a first-order phase transition takes place. The region outside of the interface is a space of incoherence where spin activity is so high that it is in a disordered state.

6.3.1 The simulation

We are testing the effects of activity on attenuated constituents. We want to illustrate the behavior of meta-vocables as attenuation increases.

After much experimentation with the model, we have settled on an exemplary simulation that best demonstrates our assumptions about linguistic interactions.

We first set the attenuation so that constituents will attenuate up to a three to one ratio (figure 6.9). Notice that the darkest portion of the grid indicates that constituents

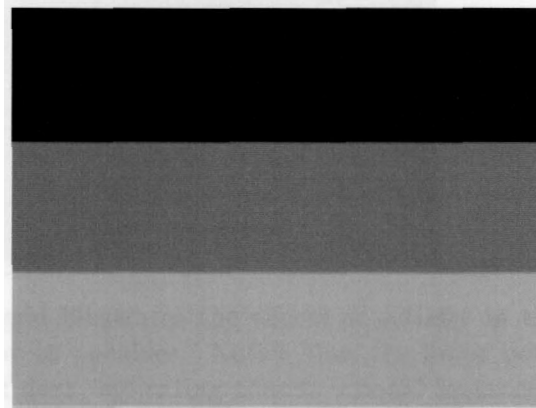


Figure 6.9: In our simulation, constituents can be attenuated up to a 3 to 1 ratio. The darkest area indicates that constituents in that region do not have an attenuation bias.

occupying the upper part of the vocables grid have no attenuation therefore cannot display a first-order phase transition.

We run the simulation around the critical point of state change for constituents with a three to one attenuation ratio but not the critical point of the other constituents. We then let it settle for many steps.

6.3.2 Results and interpretation

Preliminary results show that the model is consistent with observational data.

Figure 6.10 illustrates the vocable grid. We observe that, in its settled state, the lower

portion of the grid is almost entirely dark while the upper portion is almost entirely light. The middle of the grid show a downward increase in clustering of dark constituents. We

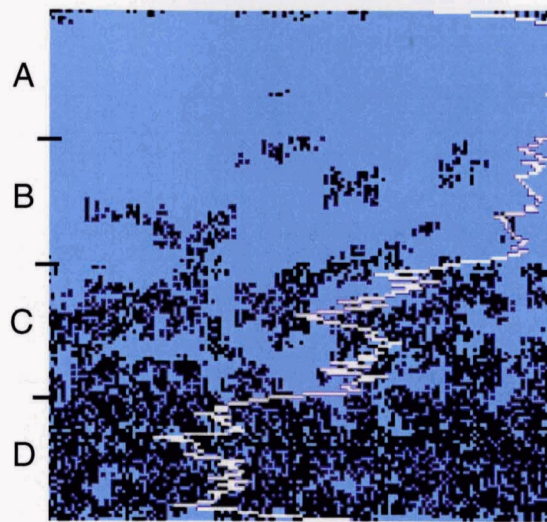


Figure 6.10: The spin grid illustrates the effects of activity in the system in the context of degrees of attenuation in vocables. Notice that the lower portion of the graph shows a greater distribution of dark, indicating that functional instances are in a majority. The middle portion illustrates some clustering which may indicate occasional functional instances however, unstable.

expect the A portion of the grid to remain mostly lexical (light) because constituents in that region have no attenuation. However they are subject to the bias field that keep these constituents in a lexical state. This is comparable to the very early uses of vocables that are entrenched in perceptual cues. In fact, some vocables do not leave that stage. (e.g. some nouns and verbs)

In the B area of the grid we observe small sparse clusters (dark areas). This is equivalent to the appearance of aberrant occasions of functional uses. These are ultimately not sustainable because the critical point has not been reached for those constituents. In linguistic interactions these occasions of functional vocabulary will not survive because they may not generate enough specific effects so they will be ignored or perceived as mistakes and corrected.

In the C area of the grid, the number of clusters augment while the majority of constituents still remain in a lexical state. This could indicate rare but stable pockets of functional uses. Again, because the critical point for these constituents has not been reached either, these functional uses are not stable. This may be equivalent to dual values in the use of some relational vocabulary. Often one of these uses will eventually disappear or be tagged for disambiguation. Consider the uses of *thanks to*. In the construction;

Thanks to Sue, we have met our goals. This lexical use of *thanks to* is commonly used.

Thanks to the weather, we had a good run. In this example, the slightly less common use of *thanks to* is equivalent to the use of *because*.

Because of the weather, we had a good run.

In the D area of the grid, most constituents are in a functional state. The critical point for these constituents to change state has been reached. This is a stable state for constituents. This is when lexical uses of a vocable become sparse and isolated and mostly functional instances are in use. It may be that some lexical instances persist in the language but they do not generate specific effects or that later lexical instances have strayed considerably from early ones. Consider the use of *if*. Its Anglo-Saxon roots, *giefan or gyfan* - related to the current uses of *given that* - has disappeared from Modern-English but can still be found in written literature. However, these written occasions will only generate specific effects with certain linguists, but not in the general population.

We have simulated the lifespan of a meta-vocable in a simple simulation as it changes because of linguistic activity. It is unlikely that any vocable in the language follows an identical path as the one we have modeled for the meta-vocable. However, we have illustrated, with a few examples, that several types of vocables display behavior similar to the ones we have simulated.

6.3.3 What we have demonstrated

With this experiment we have demonstrated that the framework of physics is useful in building a theory for the evolution of language. We have shown that the simplest model that demonstrates a first-order phase transition - the Ising model - is viable and relevant to the exploration of interactions in language evolution. However, we think that our heuristic approach to the Ising model has conveyed as much as it could offer. We also think it important to pursue more sophisticated models to explore such phenomena as phase transition. Moreover, we also think that beyond the capacity to confirm what we already know about

language we can learn new facts about its workings by using modeling techniques.

Chapter 7

Neural Darwinism and Physics

William H. Calvin has constructed a model in which Darwinian competition can lead to sophisticated organization in a chaotic environment. Darwinian competition, in this context, is the successive cloning of a pattern through a territory of synaptic connections.

His model relies upon observations of the superficial layers of the neocortex. As we shall concede, the particular properties demonstrated by the cell distribution in this region and their wiring support Calvin's theory. The details of Calvin's claim and the concession are as follows.

7.1 Superficial Layers of the Neocortex

Calvin uses the term *superficial layers* (see figure 12)¹ to refer to the three top layers of pyramidal neurons, as identified by cell size and axon-packing density: level one is situated immediately beneath the pial-glial membrane, level two, of small pyramidal cells, and level three, of larger pyramidal cells, whose apical dendrites reach upward to level one, and whose axons extend to lower (cortical and sub-cortical) regions. The distribution of the neurons (c. 5mm apart), their repetitive firing, and their mutual re-excitation and consequent entrainment provide, as he supposes, significant clues to the nature of the self-organization of neuronal firing patterns.

The consequent behavior examined in superficial layers exhibits characteristics not unlike the emergent synchronicity in a local population of firing fireflies, say those in a single

¹This picture is taken from *The Human Brain Coloring Book*. See the bibliography for a more complete reference.

tree. Individual fireflies light up in independent sequences but the proximity of the nearest firefly provides a weak entrainment influence that eventually propagates to the entire tree. Eventually all fireflies light-up synchronously. This kind of phenomena is related to *feedback* behavior; a concept that will be explored later in this chapter.

It is this kind of organization that the Calvin pattern-propagation model seeks to illustrate in the context of the superficial layers of neocortex.

7.2 Triangular Array and Hexagonal Firing Patterns

The model that, in Calvin's account, provides the stage for the simulation of such behavior, is an idealized triangular grid that hosts a neuron-type constituent at every intersection. The triangular grid models the particular way in which pairs of neurons that are in mutual re-excitation mode tend to entrain sub-threshold neurons that sit equidistant from the members of the pair. Furthermore, adjacent edges of excited triangles may co-opt a fourth and a fifth cell to complete a large hexagonal synchronous patterns of triangular arrays. This creates a kind of *hot spot* that can potentially excite entire structures through its influence over *nearest neighbour* interactions - interactions with nearest surrounding constituents. Figure 7.1 illustrates Calvin's model. These hot spots are sometimes referred to, particularly in physics, as *attractors*. An attractor is usually a small critical number of constituents in a system, that are in a particularly stable state, that is, a strongly defined pattern that can induce into its own state neighbouring constituents that are inactive or in a relatively indeterminate state. Activation in a neuron can occur in two ways: from direct sensory input and from *feedback*² or *reentry* in which initially stimulated neurons activate neighbouring

²In the shift from naturally produced synaptic effects, that is, effects generated by being exposed to perceptual cues, to artificially produced ones, that is, neural effects produced endogenously, there is a temptation to infer the idea of *deliberateness*. *Deliberateness* is related to a state in which an organism can generate behavior without seemingly being reactive, that is, without having been immediately stimulated by a related event. Most organisms can motivate a chain of events, a cat can jump on a lap without the active coercion of its owner. It is that kind of deliberateness that is discussed a kind of *self motivated* activity that is seemingly *internally* generated.

The study of synaptic effects and their workings suggests that *deliberateness* may be the consequence of a complex interaction with one feature that may be responsible for *internally generated* neural structures. To be deliberate about an action means that some synaptic effects are stimulated such that some activity will result. However, not just any activity will do, so not just any synaptic effects are stimulated. These synaptic effects are a rehearsal of the activity to come a kind of simulation before the fact. This simulation can occur because neurons are two-way units that can be stimulated directly from input patterns and from output patterns. The output pattern is a result of resonance or feedback, that is an echo of the input pattern.

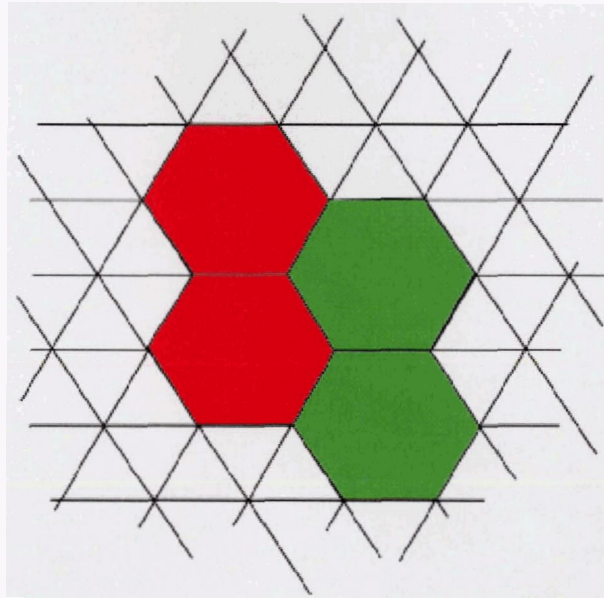


Figure 7.1: Calvin's model: The triangular grid models the particular way in which pairs of neuron that are in mutual re-excitation mode tend to entrain sub-threshold neurons that sit equidistant from the members of the pair. Furthermore, adjacent edges of excited triangles may co-opt a fourth and a fifth cell to complete a large hexagonal synchronous patterns of triangular arrays. This creates a kind of "hot spot" that can potentially excite entire structures through its influence over nearest-neighbour interactions - interactions with surrounding constituents. Here, the green and red hexagon represent competing patterns for neural territory.

neurons and then, are re-stimulated from the activity in the neighbours.

As neural cells are co-opted into a particular pattern that extends into a net-like structure, it is likely that this structure will encounter another set of synaptic patterns along its edge. The other pattern has its own attractor. Sometimes these attractors can be easily disrupted so that one structure will overrun the other. At other times, attractors are very stable so they get into competition with each other. These structures may be the result of some sensory stimulus. The competition may arise from choices to be made and may resolve with time or additional information.

The resolution may occur in different ways but it usually implies that a territory overtakes a competing territory. However, sometimes territories can somewhat overlap and the boundary of overlapping territories can develop its own attractors. Calvin thinks this resulting territory can give rise to what we would cognitively recognize as categories. An overlap of structures corresponding to *apple*, *orange* and *banana* can create a coherent superposition experientially rooted in a shared sensory experience such as *sweet*, *tree-borne*, and so on. These stable points may resonate sufficiently to generate stable attractors that can co-opt neurons throughout the overlapping territory. This new territory may be akin to a process of

Feedback, resonance, recurrence, reentry are concepts that have slightly different connotation in different fields of studies but they are all synonymous and refer to a particular feature of complex interactions. The most general definition is found in the description of non-linear dynamics such as catastrophic events. It states that feedback is a functional iteration, that is the result of a function becomes the input value for the serial iteration of the same function. In systems, feedback can rapidly get out of control as in the distortion effects sometimes generated by a loudspeaker that feeds back into a microphone. It can also generate stability as in the case of a thermostat that regulates the temperature in a house. A temperature above a fixed point leads to cooling while anything below it leads to heating[24].

Given a strong or well worn input pattern, resonance can be prolonged or strengthened from partially similar input patterns. The output pattern can then be generated given loosely similar input patterns. Indexical effects are then generated increasingly outside from their natural production. As the endogenous pattern is sustained through multiple input patterns it is possible to generate the original indexical effects on command. We can imagine that in the absence of original auditory and visual cues, resonating auditory discriminations may fade or adopt novel features. This can explain how throwing behavior give rise to pointing behaviors, how string of vocalization can become divorced from the natural context of production and become functionally extended.

Resonance or feedback is a key feature in the dynamics of systems in evolution. Feedback occurs as constituents impact on each other. Negotiation between unions and management rely on feedback to reach an agreement. An agreement pushes the system in a state of equilibrium. Prior to the accord union members may hover between a state of strike and a state of work. As negotiations proceed the system is in a state of criticality as it is not in any ordered state at all. Contracts are drawn up based on items that both party recognizes. The end product is often quite different than what both party had intended as they imagine outcomes before it happens.

Feedback enables simulation. Without feedback it is unlikely that particular synaptic pattern could be sustained for any period of time. Moreover, the retention of information would be impossible.

schematization that, one might surmise, could be the neural basis for a derivative category such as that of *fruit*.

In Calvin's model, these structures are not merely syntactic; in fact, syntax is rooted in the sensory experience of being and moving about in the world. Syntax evolves from perceptual and motor functional structures that are *activated* during vocal transactions.

The perceptual (those involved in the experience of an apple) and motor-functional structures (those involved in the utterance of *apple*) are intertwined. The neurons involved on the edge of the structural landscape tend to alternate between spatio-temporal attractors, one for the motor-functional structures and one for the perceptual structures, until recognitional competitions are resolved. Once resolution is achieved, the attractors merge into one structure that constitutes the immediate experience of *apple*.

7.2.1 Spatio-temporal patterns or active memory vs spatial patterns of lingering memory

Spatiotemporal patterns are structures of a kind that arise from the sensory stimuli of, say, a linguistic interaction. This involves syntactic aspects and non-syntactic aspects such as prosodic flow, and the sensory stimuli related to the motion of uttering. All of this takes place in the here and now of neuron firing and axon bundles of a present time and constitute a kind of active blackboard that is constantly subject to a variety of stimuli. However, to follow the metaphor, as we erase the blackboard, lingering shapes remain visible, and sometimes become so imbedded that even a good scrub will not remove the patterns. These can be compared to spatial patterns that are characteristic of long term memory. These structure rely mainly on the feedback of lingering attractors to sustain their pattern.

Calvin suggests that mixed memories or loss of detail may be the result of the reliance on certain patterns of distributed attractors that result from stimuli of similar but not identical sensory patterns. These attractors are distributed through axons that can extend laterally or beyond the superficial layers of the neocortex to other areas, not unlike pipelines that bypass local natural gas distribution to service area further away. Calvin makes an assumption that once the signal has traveled to a different area the attractor is reproduced remotely because the neuron cell arrangement is similar to that of the initial area. However, the distance traveled may hinder the signal, so that the now distributed attractors may have to rely on local stimuli to strengthen their common pattern. Upon reactivation the distributed attractors will provide a workspace in which the initial pattern can find a loose

fit.

Since distributed attractors rely on neighbour interactions to strengthen their own pattern, it is in fact temporally divorced from the initial pattern. The resulting effects constitute memories. They can be sufficiently robust to survive certain types of seizures and coma, but the general procedures by which they are produced are also the vehicles by which memory error is introduced.

7.3 Schemas and metaphors

Essentially, a schema is the consequence of overlapping patterns that intersect on common features. A metaphor is similar in the way it presents a feature that is common to several physical events. A metaphor is a *well worn* linguistic string that is disarmed of its capacity to occasion some original effects. Unlike functional vocabulary, however, a metaphor's early perceptual associations are held as a *subtext*, somewhat unspoken and not entirely forgotten.

If someone refers to a situation as being a "Catch 22"[29] it is not, we suppose, because he has just read the book and compulsively refers to it, but because a particular aspect of the story-line that was salient for some previous commentator, is salient in some aspect of the present situation. The use presents a kind of puzzle that may be insufficiently resolved even for the speaker, and may require the resolving power of an additional speaker to achieve usable focus.

I feel like I'm in a catch 22.

You mean you feel damned if you do, damned if you don't? In this example the second statement is also a metaphor but it provides a focus for the initial statement, partly from the overlap of common features but also because, although the expression *damned if you do damned if you don't* is not much clearer, and perhaps only minimally more resolved, the agreement on the term represents a negotiated common ground.

Sufficient resolution is achieved by a neural process that consults both memory and sensory data provided by the present situation.

7.3.1 Algorithmic shortcuts

One way of avoiding processes that are too expensive in terms of cortical competition in neural space is to create *shortcuts* with some of the processes that are constantly activated.

Many of those are generated in sub-cortical areas such as the basal ganglia. However, Calvin suggests that the cloning competition itself can give rise to algorithmic shortcuts.

Most spatio-temporal structures involve the coupling of sensory stimuli. Maybe an object, or a voice, a presence can be activated along with a spatiotemporal structure, and will prompt some action. In situations that require a quick reply or a certain flow, such as in a conversation, the resolution of competitions that involve large cortical territories may require too much time. A strategy in which a minimal number of strongly defined hexagons may be sufficient for an appropriate reaction structure to be gated. In a structure that is frequently activated in a large cortical territory, some of the hexagonal arrays will always involve the same cells and cell arrangement. These hexagons are usually somewhere in the middle of the structure so that the arrays are complete. The result is a high level of synchronicity, synchronicity precise enough for it to overwhelm surrounding noise from incomplete arrays, and may be *good enough* to set up a response in the absence of a more involved cloning competition. If you turn to me and say *catch* and I get a glimpse of a smallish red object, I will put my hands close together and assume it is an apple. Had I caught a glimpse of a big brown thing I might of assumed a basketball and part my hands further. Calvin even suggests that the constant activation of specific arrays may bias neuron threshold.

7.3.2 Calvin and phase transition

As we have mentioned, Calvin's model offers a finer grained description of dynamics similar to the ones we have used in our description of the evolution of language. We think that our description is compatible and must be reflected in the dynamic of neural interaction. Calvin's approach allows us to do so. We think that by using a description of Calvin's model that includes phase transitions we can add details that Calvin's description lacks, specifically in relationship to what could be the synaptic event of syntax and more abstracted forms of vocables such as functional vocabulary.

Calvin's model does not strictly model *phase transitions*, though the concept is not incompatible with it. He does mention that a resolution of competition is the result of a phase transition. However, no detailed description is given of the kinds of structure that can result from a phase transition. Phase transitions could explain much in the activation of structure in low-connectivity or "hurry up" time. I suggest that this finer grain mechanism

can help along the cloning procedure. Dynamical equilibrium can be reached in a low-connectivity system - where limited energy is available - through a phase transition.

Percolation is a good example of a phenomenon leading to a phase transition that can further the description level of Calvin's model. To re-iterate, percolation theory deals with the properties of clusters of occupied sites that form into a lattice. Percolation refers to the possibility of having a cluster spreading over an entire lattice. Percolation usually demonstrates second-order phase transition defined by the connectivity between sites. This state change is usually the result of a critical number of connections between clusters. Percolation is a description of low-connectivity structures. For example, given two metal plates separated by some physical distances, say 100 cm, that hold an electrical charge, percolation would describe the creation of a path for current flow when coins are randomly thrown down between them. The definition of connection here is a conductive path, the nature of the medium is round coins that can overlap one another, and the nature of the medium evolution is random dropping. The function of these variables give rise to a particular percolation behaviour. Also consider the *Kerplunk!* example given in the introduction.

In Calvin's model, locked states can emerge in hexagonal arrays that are highly synchronous, from constant activation. A situation that requires a fast response could be described as a low-connectivity system in which percolation may occur. The high synchronicity of stable hexagon structures, despite being minimal, will successfully clone without the benefit of competition.

7.3.3 Phase transition and functional vocabulary

Some of Calvin's examples involve spatio-temporal patterns that do not rely on a cloning competition or the activation of a entire structure in order to be gated. Further sensory input may help in defining a course of action.

Functional vocabulary may also be the result of a restructuring in a well- worn synaptic pattern. This restructuring may have occurred as the result of phase transition but it probably does not involve spatio-temporal patterns. Functional vocabulary may be the result of a restructuring in low-connectivity, spatial patterns such as schemas or metaphors.

Imagine a well worn metaphor in which *other*, the ancestor of *or* will occur. For several reasons the cortical territory that sustains the metaphor may be fragmented and co-opted into other structures. However, frozen cores or stable hexagon arrays may remain as attractors and percolation may occur to form a related synaptic structure. Because many schemas

or metaphors are sustained throughout distributed synchronous attractors, it may be that the resulting structure of a percolation is only vaguely related to the distributed structure of a metaphor. So out of a metaphor structure involving say, *other*, the *or* structure may emerge. Spatial patterns rely mostly on endogenous activation for their survival so that may be why some vocabulary come to play a functional role divorced from any physical attribution. The resulting structure of a phase transition, such as percolation may be used as functional vocables once prosody fails to sustain the causal flow in a long sentence.

So *or* may be found stringing sentences if only for the purpose of sustaining the attention of a listener;

The murder happened because he was depressed, or maybe he didn't do it and she did, or maybe I'm wrong all together.

7.3.4 From pre-grammatical to grammatical

As we have explained, functional language plays the role it does because it is structurally different. We assume that in the long run - across a millennium - no vocabulary is immune from structural changes since low connectivity involves an optimization of cortical space. Of course, not all vocabulary becomes functional but all vocabulary becomes extinct eventually. It makes sense to assume that, as the cortical space is hosting an increasing number of cloning competitions, there must be strategies to optimize limited resources. Phase transition is one way by which cortical space may be optimized. Furthermore, we suggest that phase transition in proto-language may have promoted grammatical forms without the necessity of major restructuring of neural resources at the genetic level.

The percolation phenomenon - remember the Kerplunk! example - would make it possible for grammaticallization to emerge from a system that has not displayed that particular feature at any time in the past. Moreover, constituents in a proto-language system do not have to be biased towards a grammatical arrangement in order to develop that feature. Grammaticalization may emerge (to sustain a causal flow in a proto-linguistic transaction), from the exploitation of low connectivity structures in cortical territories. The level of generalization in functionalized items may be just what is necessary to accommodate increasing complexity in vocables.

7.3.5 From infant to adult

The childish form of language may undergo a similar structural change, though, in this case, some items may be biased to prefer a functional state. The structural change would result from a first order phase transition instead of a second order phase transition such as percolation.

Children learn grammar from adults. They register sentences that involve functionalized vocabulary. It is not clear how these vocables are processed but it is common knowledge that when a child is asked if she wants an apple or a banana she will say *yes*, rather than specify which. It often requires a certain experience with language before functionalized vocabulary can be used expertly. It may be that the cortical territory must reach a point of overload to exploit low connectivity structures, characteristic of functionalized vocabulary. It may be that sentences involving functionalized vocabulary must lose the temporal aspect of immediate sensory activation before the distributed nature of the functionalization process will be fully exploited.

We can imagine that the synaptic patterns of thousands of sentences intersect on only one point, say, the sound *or*. This sound could be the vocable *or* or it could be the *or* in *ordinary*, *other*, *orbital* etc. . . .

Suppose that all structures that hold the *or* sound generate stable *or* attractors that are widely distributed. Other of these *or* attractors may be easily dismantled and readily co-opted into other structures but generate weak entrainment non-the-less. As cortical space becomes restricted the bias generated by the strong hold of *or* attractors may percolate and entrain all distributed - non spatio-temporal - *or* attractors, the stable and less stable ones. This kind of phenomenon may explain the emergence of the *connectival* use of *or* and how its functional aspects may be readily exploited.

7.3.6 The difference between metaphorical and functional language

Both types belong to the class of vocables that have been *disarmed* from its original context. The temporal context that gave rise to metaphors and functional vocabulary is increasingly absent in subsequent uses. Both are generalized to the point at which only a handful of features can apply to many contexts. Calvin suggests that the making of a metaphor involves loose-fitting structures activated from distributed attractors, layered from short-term memory to long-term memory, that can easily recall the underlying spatiotemporal

foundation classes of more literal structures. This suggestion offers a convincing synaptic explanation that supports linguistic facts. Calvin also suggests the use of short cuts in times of quick responses, may contribute to the casual use of some vocables.

It may be that some of these short cuts are achieved through phase transition. These possibly give rise to structurally altered vocables, that play an essential functional role. The functional role played by vocables is different from metaphorical roles because it is structurally different. Metaphors result from competition in which schematic features try to match a spatio-temporal context. Functional vocabulary is probably not often involved in competition as its only connection to spatio-temporal structures are non-linguistic and prosody related (we will elaborate on that subject in the next chapter).

It is clear that a simple explanation cannot be given at this time for all the many ways we use vocables and how we come to have them. However a description of language that uses phase transition as part of its story has the additional advantage of offering a point of view that can explain the emergence of a variety of linguistic types and their role outside of the intervention of a psychological agent and in a completely predictable manner. Models that include attractors tend to be sensitive to initial conditions and have to be handled in a heuristic manner because the results they yield are not exactly predictable. Critical behavior, such as *phase transition*, is, on the other hand, formally well understood and predictable. The nature of a phase transition process is also such that it is universal because it is an observable feature of a number of natural phenomena, modeled in systems, such as in ferromagnetization behavior, and in macroscopic systems such as social percolation models.

7.3.7 Institutional vs spontaneously occurring metaphor

It may be useful to make the distinction between the kind of metaphor that has become institutionalized as part of common speech (*gone the way of the dodo*, *burning the midnight oil* and so on), and metaphors that turn up spontaneously in individual speech. The distinction is important. By which feature is the spontaneous nature of one different from the institutionalized nature of the other? There must be a connection between the two since, we assume, an institutionalized metaphor does not begin life institutionalized.

One fundamental difference must lie in the manner of production by the individual brain. Newly coined metaphor draws upon elements of speech, whose primary functions have hitherto lain separately elsewhere. In metaphor that has become institutionalized, the same

elements, though they retain distinct distributed uses as well, also have a separate availability specific to the present construction. In traditional philosophical terms, the difference might be said to lie in the character of the associations of a metaphor's components. If we imagine a succession of associative images prompted by the successive components of a speech production or reception, the contrast might be considered as the difference between

- (a) distinctly represented items which are assembled into a spatiotemporal whole, and
- (b) an spatiotemporal assemblage simultaneously or nearly simultaneously given.

The reversion of institutionalized metaphor to non-metaphorical status would, then, require a kind of over-illustration, as if when we heard the expression *Lord Privy Seal*, our brains contrived successively images of Jesus, an outhouse, and a pinnipedial mammal.³

In Calvinist terms, the production of the spontaneously occurring metaphor may result from the competition of parts of spatiotemporal structures closely related to perceptual cues. Institutionalized metaphors are treated as a whole and are meant to stimulate structures that have a distributed nature, not so entrenched in perceptual cue, such as lingering memories. The difference in both is really a point of efficiency in which an institutionalized metaphor is somewhat cheap since the competition has already occurred somewhere *outside*.

The example of the *Lord Privy Seal*, well illustrates the kinds of dynamics involved in a competition of parts versus a context treated as a whole. The *Lord Privy Seal* is not really a metaphor; nevertheless it furnishes a convincing example of an expression which, while it is capable of being treated as an indivisible whole, has for its components elements having distinct, robust, and unrelated uses elsewhere. This is vividly true for metaphors and idiomatic constructions, but we suspect that in some considerable measure, it is crucially true for the whole of language.

It may be a minimal requirement that a model, such as Calvin's, distinguishes between institutionalized and spontaneously generated metaphor. However we maintain there is still an unexplained residuum of linguistic phenomena that require a model capable of representing populations of brains. In particular, linguistic constructions cannot become institutionalized other than through propagation in a population of language users. In fact

³The illustration originates with David Frost *Frost over Europe*, 1970's where, in a fictional news program, the news presenter's text is illustrated as it is spoken by a projected image that is shown simultaneously as the text is presented. The over-illustration accelerates as the reading progresses, until they are shown at a rate of one per word at the end of the broadcast. The associated images accompanying the final string *Lord Privy Seal* are, as in our example, a conventional image of Jesus, an image of an outhouse, and a circus image of the marine mammal. The example originates with R.E.Jennings.

it is plausible to suppose that it is through being *learned* as a unit that an institutionalized construction gains its characteristic independence from its spatiotemporal parts. A certain neural innocence on the part of post-inception generations of language users is required. Thus institutionalized metaphors cannot be the result of one brain's Darwinian competition. If they result from any such competition, we must assume that they are the result of a competition that involves many brains. Indeed, we may not need the details of Darwinian competitions within single brains for the purpose of modeling this kind of functionalization. This does not preclude the possibility that the mathematical description of a multi-brain model of institutionalization will be similar to that of a single-brain model, though the multi-brain model itself may be relatively coarse-grained.

It is significant that Calvin does not make this distinction: Central to his model, and certainly the main focus of its application, is the idea of a discrete linguistic subject and its brain. For this reason, Calvin's model does not have the means to explain how metaphors can become part of the common speech. The example of "Catch 22" suggests a kind of negotiated vocabulary for present perceptual cues between individuals. This suggestion helps erase an artificial boundary set by the concept of linguistic capacities *within* individuals and linguistic transaction that occur *outside* individuals. Institutionalized metaphors are negotiated, not in the sense of convention agreed upon by committee but much like neuron structures in which there is *competition or energy exchange* in neighbour interaction that propagates to an entire population. It may be that all vocabulary is *negotiated* and that negotiation is a coarse grained version of *neural competition*, both occurring nearly simultaneously and contributing to each participants.

Calvin's model suggests that a metaphor results from a competition between schema of literal spatiotemporal patterns. We suggest that metaphor may arise in speech as the outcome of a competition between schema that have been disarmed of some spatio-temporal effects, and their capacity to occasion highly specific expectations of perceptual cues. However, the competition between patterns is not resolved with one individual but requires interactions with neighbours, a behavior that is mirrored in neuron interaction.

This kind of competition cannot be supposed to be peculiar to metaphorical language unless this category is dramatically widened to include most vocabulary. Virtually all of the vocabulary of any natural language becomes institutionalized in the way that metaphor does by a gradual diminution in the capacity of a construction to occasion highly particular effects. This requires propagation through a multi-generational population of language

users. Given that all linguistic beings have non-linguistic ancestors, it is plausible to infer that uses of all vocables represent novel exploitations of previously available devices. This in turn requires that linguistic devices be to some extent disarmed of their earlier capacities. Calvin's model suggests that all syntax is the result of successful propagation of certain patterns in a neuronal landscape.

We suggest that this kind of competition extends to linguistic transactions between members of a linguistic species and includes not only metaphorical language but all vocables.

7.3.8 Population Dynamics and Calvin

It not quite enough to suggest models that describe parts of the linguistic process, it is also important to show how these models fit within the larger body of knowledge that we have about physical systems.

In the next Section we will scale the problem to include a description of the emergence of language in terms of synaptic functions and prosodic flow. While considering the dynamics of language evolution from a physics point of view we will build a finer grained argument that will include neural functions amongst many brains. The integration of the physics of language theory and the theory that describes a kind of neural competition that exists amongst brains, offers a functional description on how linguistic activities may have emerged.

Chapter 8

Population Dynamics, Prosody and Neural Effects

This chapter includes miscellaneous thoughts that are not crucial to our argument but nonetheless important for an explanation of the emergence of language. Most of these thoughts gravitate around the concept of prosody - the song quality of speech - and its importance in the shareability of language. Simply put, prosody includes all vocal aspects of speech such as the tone of voice we use, the different emphases and pauses, and so on.

Prosody is often considered a secondary feature in linguistic interaction and as such we have not given it the attention that we now think prosody may warrant. As our research progressed it became apparent that prosody may play a much more important role than we had previously thought. Here are some of the theories that we have about prosody that have led us to believe that, in fact, it may have played and still play a crucial role in the emergence and the continued shareability of language.

In this chapter, we imagine how the exploitation of noises and vocalization of the everyday life of proto-humans might have led to prosody and, in turn, to language.

We also describe how seemingly unrelated everyday sounds come to be associated to all of our perceptual experiences and may have been exploited towards a shared system of vocalization.

We finish by describing the break-down of prosody, in certain expressions, as the consequence of attenuation.

Many of these ideas have been discussed previously in other chapters however, here,

we slightly reframe them in a neural context. As previously mentioned, we think that the population dynamics that we have described in the context of sharing a language must be reflected at every step in a neural context.

8.1 Neural Competition Amongst Many Brains

The competition aspects of Calvin's model is extensible to population dynamics even at a very early stage in the evolution of language.

Imagine a time when there is no language as such, only utterances or even some type of linguistic noise that could constitute very early proto-language. This proto-language may employ previously used actions that are disarmed through a kind of quoting mechanism, such as we have described in previous chapters. These might be disarmed linguistic vocal responses. These linguistic vocal responses may constitute proto-vocables in which a kind of proto-attribution occurs. This attribution may be described as the kind of competition metaphors present. However these are not quite metaphors, despite the fact that proto-attribution represents an adaptation of previous uses of linguistic vocables. A proto-attribution also involves direct visual stimulus that differentiates the linguistic vocables from their original context of use. Hence the competition for neural space. As in the case of the metaphor this proto-attribution may require a vocal transaction to resolve the competition.

A system of schematization could have promoted linguistic novelties early on in proto-attribution. Imagine that the category *fruit* does not exist, but that the shared experience of the category *oranges* does. Imagine further that an utterance exists that is characteristic of the presence of oranges, but that there is none characteristic of the presence of kiwi fruits. Imagine further that a kiwi fruit is presented by one member of a late hominid species to another member of the species. However there is no specific utterance that exists to describe classifiable features of a kiwi in particular or fruit in general. The strategy may be to use the utterance for the experience of orange and modify it somehow. If this is an attribution at all, it is certainly a wrongful attribution, however a gesture that mimics eating presents a competition between the wrongful proto-attribution of orange and the visual cue. The other member may *infer* something about eating, perhaps sweetness and some vague aspect of shape. There may be some back-and-forth gesturing to establish a common experience of the kiwi fruit and, in the process, modify the proto-attribution of orange to include - and exclude - perceptual features that will specifically acknowledge similarities and differences

between oranges and kiwi fruits.

We suggest that competition between perceptual structures of groups of individuals bring about conventionalization. The transaction is of the sort that we might fancifully suppose would have been resolved by the overlapping synaptic structures emerging from exchanged reports of previous experiences, reports that would focus particular features. Hence the competition for neuron space may partly be a competition engaged in an arena involving two or more brains rather than one. We imagine that it is this sort of interaction that generates what we commonly understand as *conventional use*. On this view, the competition between vocal stimuli and visual cues provides a mix-and-match resolution problem that promotes a kind of generalization; or perhaps, proto-generalization since the vocal and visual cues may not yet constitute anything that we should want to label *linguistic*.

8.1.1 Early functionalization

The functionalization process is not restricted to linguistic transaction, we can also apply the concept of functionalization to other types of behavioral transactions, present in hominids. The earliest ancestors of functional vocabulary need not have had any specifically linguistic characteristics. There is no reason to suppose that for hominid species, vocal sounds - perhaps proto-musical sound streams - could not carry socially significant information. Since the character of such streams would be affected by the physiology of their authors, they would be a common individuating voice print of members of a group. Opportunities for novel exploitations of such productions arise with activities involving sudden motions or alterations of motion that have similar effects on such a stream for any member of a group.

This, of course, is not the beginning of language, nor need we suppose that it is even part of the early story of language, but it is an illustration of what a pre-linguistic precursor of functionalization might have been like: In the imagined account, non-linguistic sounds naturally produced with available initial benefits acquire a role in which they contribute to the safety of groups by providing a means to distinguishing members from non-members, and individuals from the group. Derivatively, therefore, it contributes to *social cohesion* and other advantageous conditions. This is already an instance of a process of the same general kind as functionalization, that is, an exploitation for a novel application, of a feature preselected for some distinct earlier role. In Stephen Jay Gould's [25] language, it is an *exaptation*.

Merely to have said this much hardly makes it less mysterious that such changes should

occur, particularly later. In what we have so far described, we find three stages and two transitions between them. One stage is from voiced sound production as a natural expression of physiology which transitions to voiced sound production as social tool, such as the soothing sounds of a mother, and constitutes an other stage in the use of sound production. There is also a stage in which voiced sound production as identificatory device has accidental features - exertion sounds, eating sounds and so on. These activities can be distinguished by sound alone without the support of visual cues. At some point we can imagine a transition towards the exploitation of these auditory discriminations in the artificial production of effects, that is, in the absence of physical and perceptual cues that give rise to these sounds in the first place.

8.1.2 Exploitation of auditory processes

It is difficult to imagine the process of migration of the auditory discriminations from natural production to artifact. In the first instance, sounds are produced that generate synaptic effects in members of the species. Sounds and concurrent events are registered as auditory and visual cues, and are discriminated in some way. The precise character of this discrimination hardly matters for our purposes. What is significant is that auditory cues are eventually selected and reproduced so that indexical visual effects are recreated, as it were, on command.

The point is easily illustrated. Imagine two people engaged in an activity. Their engagement is primarily kinetic, but since their contributions are complementary, their awareness, though partly kinesthetic, is also partly visual, since each sees the other's contribution, and secondarily auditory, since the activity also produces (perhaps characteristic) sounds. In consequence, their activity produces *concerted* visual and auditory synaptic effects. The coordination of the activity yields a coordinated composition of neural effects. To borrow a human term of folk-psychology, we could say that these effects are *associated* in virtue of being produced in this way. The neural details of the association remain somewhat obscure. However, the significance of this association, perhaps the substance of it, is that one kind of stimulation is capable, in some measure of functionality, of doing the work of two.

Thus, even in the absence of auditory stimulation, just the visual cues generated by the activity might be capable of stimulating a sufficient range of auditory synaptic effects for

some kinds of control; similarly, so might just the auditory cues be sufficient.¹

The suggestion that neuronal structures can be arranged such that partial excitation is enough to generate complete stimuli is convincingly explored in Calvin's theory. The details given about structural features are somewhat speculative but the account is sufficiently consistent with experimental data to make a good argument. Calvin relates the experience of one humming a few bars of Beethoven's Fifth and suddenly the entire piece is playing in someone else's head. He goes on to explain how his model can explain this phenomenon. Most of the explanation has already been explored in the previous chapter.

In daily life the phenomenology of this neural set-up is taken for granted as a feature of our common experience. However, that there is evidence for it in the most fundamental facts of language suggests that one of the earliest stages of its pre-development was the exploitation of this kind of migration. The exploitation features the artificial reproduction of such effects; the artificial recreation, for example, of sounds producing effects functionally, like the multimodal effects of a natural activity.

In any adequate account of the evolution of language, it may be presumed, some such development would underlie the earliest steps. These would describe the shift from which the sounds of human, or rather pre-human, activities are eventually conventionalized and emancipated from their natural occurrences.

8.2 The Process of Generalization in Perceptual Structures

Generalization, or rather some development that is a kind of succession of generalization occurs within three stages that can conveniently be labelled *accord*, *competition*, and *dominance*. In the first stage, visual and auditory cues are in agreement. (Indeed at the earliest stage the auditory cues may be presumed to be those incidentally produced in normal interactions with the particular objects in question, and the latest stage under consideration may still be only one that imitates those sounds.) At this stage the synaptic structures (apart from the purely auditory ones) excited by vocalizations are those that would be excited by

¹Researchers in the field of neurobiology have identified a class of neurons the work of which is to modify, coordinate, integration, facilitation, and inhibition between all sensory input and sensory output. These are the so-called *interneurons*[19]. They are a class of neurons whose cell structure has no direct contact with peripheral structures, that is, receptors and effectors. They are *multipolar* which is to say that they are characterized by one axon and two or more dendrites. *Multipolar neurons* are the most common neurons and are found throughout the nervous system. *Interneurons* are the source of the richness and seemingly endless array of responses to our environment.

particular visual stimulations.

At the competition stage, a discrepancy emerges between visual synaptic structures and auditory synaptic structures. One can be exposed to auditory cues without its usual visual associations. This does not mean that other visual cues are not present while these auditory cues are occurring. These new visual cues set up new contextual information while the auditory cues set up a past context. As the new context is resolved synaptically, the non-auditory perceptual cues are accommodated through auditory synaptic structures, as new associations, and are carried into future context. As such, auditory cues tend to dominate, since they can be processed in the absence of visual or other perceptual cues. However, the dominance of auditory cues do not match accurately any synaptic structures that could be excited visually. We usually do not have a specific image when *fruit* is uttered, certainly not images of hybrid bananas and apples. These auditory structures in hominids are not linguistic, but the dominance of auditory structures over other perceptual structures may eventually give rise to a functionalization of vocalization that has eventually produced linguistic noises.

Auditory cues are not solely noises but also stimulate synaptic structures involving other perceptual cues. The dominance of auditory cues occur because they can be shared in the absence of indexical cues. Though they are initially shaped from indexical cues, the process of competition endows auditory cues with independent capacities. Competition leads to a generalization in which perceptual structures are synaptically stimulated despite the absence of indexical cues. In imitating indexical cues, auditory cues can stimulate synaptic structures that can loosely encompass perceptual structures that originally would be stimulated in the presence of indexical cues. As such, auditory cues can free perceptual structures from indexical constraints. The incidental effect of such generalization is exploited and give auditory cues an independence from specific perceptual structures.

It is that independence that somehow enables the accommodation of other non-shared perceptual structures, structures that help define the elusive aspects of vocal cues that make them more than noises. Further generalization occur when incidental effects, generated from the independence of auditory cues, give rise to new classes of physical types, usually higher-order physical types. These new objects are used to produce certain effects on demand. They can also be used to perform different functions, functions that stray from original constructs.

The process of generalization can describe the emergence of vocables from non-vocables,

because the emergence of new physical types happens from a change in the accommodation of some synaptic structures. Because new physical types are usually of higher order, they can accommodate a larger array of synaptic structures. The accommodation of certain structures will lead to changes in function which in turn will give rise to new physical types. Auditory cues may have occurred naturally as a process of exertion, but the repetition of auditory cues on demand to produce specific effects has led to a shared phenomenon between members of a species. This phenomenon may not have, at first, been linguistic, and produced effects might have relied heavily on prosodic features, but through the process of generalization, vocables have appeared as new classes of physical types.

8.3 Prosodic Forces

The account of sound making in a population must lead us into a discussion about the prosodic forces that are present in the phenomenon of language production. We consider conventionalized sound production of natural activities exploited towards social purpose as already having prosodic features. Any vocalization that has not migrated into some other function is considered pre-prosodic.

Prosody usually refers to the aspect of language that are not adequately represented syntactically and not accounted for in semantic theory. Other aspects of linguistic interactions are also very important such as motor functions involved in vocalizing, perception mechanisms and all these aspects that are commonly referred to as body language. We often rely on these cues to ascertain the mood of a person or a situation. In this thesis, we refer to prosody with the implicit understanding that all these aspects are also included in the definition, in some fashion.

The standard definition of prosody includes *pitch contour*, which covers the frequencies of speech, stress, that is, volume varied to emphasize vocables, and *lengthening*, in which vocables are prolonged with sound. The role of prosody in uttering language is an intrinsic part of “understanding” it. Prosody affects the *pragmatics* of language, that is, aspects of language that distinguish mood such as interrogative;

You’re leaving tomorrow?, indicative;

You’re leaving tomorrow, I’m going with you, imperative; You’re leaving tomorrow (as

in, *leave tomorrow*)².

Mood provide cues that disarm a syntactic expression from assertion. For example the *if* clause can modify a sentence by removing the effects of an assertion;

If you leave tomorrow, I will go with you.

There is no inference from the sentence that someone is leaving tomorrow for a certainty. The time scale that extends between pre-prosodic and prosodic forces is not clear; we can assume that the process was fostered over a long period of time.

The difficulty in pinpointing the critical state at which lexical vocabulary becomes functionalized is evident by the absence of data on the phenomenon. However, there are numerous examples that demonstrate how the breakdown of prosody brings about mutations in scope arrangements in previously functionalized language [33]. The consequence is a change in role for the connective vocabulary involved. The change is not quite as dramatic as in the change from lexical to functional but is none the less reminiscent.

There are various ways of disarming an expression. As we have mentioned, quotation is such a device, *tense* can also be used;

If it were cold, I would dress up. There is no assertion that it is, in fact, cold. Altering *word order* will also disarm expressions;

were I younger, I would. . . . Using a prefix such as *if* can disarm an expression without the subjunctive;

if you do this, I will. . . . Prosody plays a role in all cases, however order sensitive constructions are less reliant on prosody to produce specific effects compare to order sensitive constructions. *If-then* clauses are order sensitive; the *if* usually precedes the *then*, so prosody is not as indispensable. The order sensitive constructions usually occur in longer sentences and may be an answer to compensate for the breakdown of prosody in complex sentences. Prosody is stable in short constructions.

Chinese languages, for example, are a well known challenge for non-Chinese because of their tonal specificity. These languages are usually offset by a simple grammar that is fluid and that does not rely on the order of its vocables to be understood.

Prosody plays a syntactic role, beyond a certain limit of the song characteristic of speech. In Chapter Two we have demonstrated how the construction *not (A unless B)*, in which the scope of the negation affects (*A unless B*) changes over time to exclude *unless B*. The

²This example was provided by R.E. Jennings

construction is now commonly understood as *(not A) unless B*. In the presence of negation, the scope of the construction is shortened. The sentence

I will stay unless he leaves is of the kind *(A unless B)* and

I will not stay unless he leaves is an example of a change in scope arrangement of the kind *(not A) unless B*.

We have also demonstrated that the shortening effect is also evident in the forming of functional language. Most functionalized language descends from longer lexical vocables, as we have said, *but* descends from Old English *butan*; *or* is a remnant of *other* in which *the* has disappeared [33]. The shortening of functionalized language is an attempt to preserve a causal flow, carried partly by prosody, in a sentence.

The practice of *quoting* previously used syntactic expression with the addition of syntactic elements invariably produces new elongated syntactic environment. The efficacy of prosody to convey mood, scope arrangement, and so on, is stretched to its limit as syntactic constructions become longer. Linguistic phenomena such as the shortening of scope, the functionalization of vocables, the creation of order sensitive constructions less reliant on prosody are all attempts to preserve a causal flow necessary to the shareability of language.

8.3.1 Plausibility of context

An other important role of prosody is in the *plausibility of context*. The Chomsky [14] example;

time flies like an arrow, in which prosody marks *flies* as a verb is entirely different than;
fruit flies like a banana, in which *flies* is a noun. To utter the second sentence using the initial sentence's intonation, renders it insensible. The first part of the sentence sets one up for a kind of syntactic structure and fails to deliver the rest of it, such that appropriate effects are not generated.

Sometimes the puzzle can be resolved by changing the scope arrangement;

I will have breakfast only if there is not enough food, sounds somewhat ruthless. The more likely interpretation changes the scope arrangement to;

I will have breakfast only, if there isn't enough food.

The habituation that comes from exposure to a specific language create *expectations* about a syntactic stream. When prosody fails to generate specific effects, for the reasons that have been discussed, changes occur in the use of vocables. As a result they find themselves in new syntactic environments.

One of the reasons for functionalization is the failure of prosody to sustain a long string of previously used vocables. The functionalization of certain vocables is an attempt to preserve a causal flow in which prosody creates expectations. Expectations result from exposure to a specific language. For example, a few bars of a well know song will entice someone to finish the melody. Prosody plays a similar role in language. A few *bars* of a well known linguistic construction creates expectations about the rest of the phrase. Unfamiliar constructions require new synaptic structures. Without sensorially immediate triggers to stimulate these new pathways, one has to rely on endogenous stimuli to generate effects. Prosody can provide some of the necessary triggering in the event that some of the syntactic aspects of an unusual construction fails to provide clear expectations.

Some aspects of *attenuation* can be understood in those terms, as well, and functionalized vocabulary is a result of this dynamics.

8.3.2 Prosody and dispersion

Several features are related to the failure of prosodic forces and the forcing of vocables into new functions. One feature is the dispersion of vocables in a population. A critical point is reached when the majority of a population is able to accept a novel syntactic environment for a particular construction. In its use, this novel construction breaks the bounds of prosodic functions. The compensating measure that results from this dynamic often brings order changes, morphological transformations in syntactic expression and, more often than not, transcategorial changes.

Our research understands categories in syntactic constructions as emerging from extensive use of all vocables in a population of speakers and that all categories descend from lexical vocabulary. Therefore grammaticalization does not occur if language as a system is not widely shared in a population.

This assumption is corroborated in Sonia Ragir's observations about the emergence of grammatical sign languages, from proto-sign-languages in Martha's Vineyards, Guatemalan school for the deaf and the Caiman Islands [50]. Her findings show that, without the prompting of conversation in sign language between a critical number of individuals in a community, grammaticalization will not occur. Deaf children who are not given access to one another or who are not stimulated by family members to express themselves, live with rudimentary sign language skills that are very context-bound.

This account does not provide a clear description of competition in neural space in our

ancestors, nor does it make any claim about the complexity of neural arrangements in our ancestors, however functionalization mechanisms remain in present forms of language and do facilitate its propagation. We may not be able to provide a model of what happened in hominid synaptic functions but an exploration of Calvin's model in the context of functionalized vocabulary may provide some historical clues.

The architectural constraints associated with synaptic functions interfere with prosody or prosody-related features such that the required causal flow in vocal transactions cannot be maintained. The failure may not be so much an inability as much as a taxation in perceptual cues that becomes too expensive in terms of neural space competition.

8.3.3 Exploitation of Incidental Effects in Population Dynamics

The story of functionalization is really about using what is already there. A gesture, a stare, a vocalization can be used to perform a function that differs from its original context or as an extension of it. The functionalization of features is an exploitation that can either, further the propagation of a system or stabilize some aspects of it. Both strategies result from the kind of competition that is descriptive of most dynamical systems, the ability to propagate at a low energy cost in a context of limited resources. In this case, the limited resource is synaptic landscape.

Visual cues may have been replaced by vocal cues because visual cues, being indexical, fail to provide the abstract specificity and capacity for generalization required for a propagating communication system. In fact, some degree of specificity is probabilistically guaranteed just by the facts of vision in the first instance. It is likely that organisms of common stock, similarly oriented, would focus upon the same region of their visual field, and therefore to some extent share salience relations among visible cues. Certainly, whatever such visual commonalities, as there were, would underwrite the success of early diexis. But they are in themselves too neurally taxing and possibly too entrenched in individual idiosyncrasies to play any communicative roles, even with pointing. The indexicality of pointing certainly represents a decisive limitation on its playing a role in the more general propagation of information.

However pointing, and the visual capacities that underwrite its provide, is a plausible boot strap by which vocables are brought into play. Vocables are the means that allowed us to sever the ties of dependence upon indexical devices. At the same time, to say that vocables are involved so early on is already to put the matter too precipitately. There is

no reason to suppose that the earliest ancestrals of linguistic vocables were not comfortable spontaneous social accompaniments, a kind of social adhesive. It may have been the later ancestry of pointing that drew from a relatively undifferentiated social soundscape, the later ancestors of words, simply by eliciting the relatively differentiated alterations that are the natural consequences of sudden motion or exertion such as grunts.

Even the soothing sounds of social bonding may be an extension of sounds that accompanied the care of children. In turn, the comforting sounds of mother may have been a continuation of feeding noises or an imitation of baby gurgles. A baby's satisfied noises after a good feed is probably the least threatening sound of all. The soundscape of hominid social habit must have been very complex indeed. It must have involved, as well, individual characteristics in vocalizations so that recognition was possible, variations between males and females, adults and young. In addition, the sheer number of members in a community influences the level of noise. Many species of birds use the stress level, created by the level of noise generated in a flock, to control their birth rate. Stress levels control the production of the number of eggs. As in the bird example, many features of hominid soundscape may have contributed to some function beyond its original occurrence. For example as throwing or hitting occurs, sudden exertion noises would have punctuated the soundscape of social interaction. These exertion noises would have promote changes in the rhythmic flow of social sounds. Moreover, the noises of things hit would have provided additional sounds that could have been used to further modify a rhythmic flow. Given a social soundscape, punctuated with peeks in vocalizations and clunking sounds of different textures being hit, the rhythmic sequencing of such features may sound like a kind of music. The non-portability of the hitting noises may have lead to the imitation of clunking noises in the absence of hitting tools or the *appropriate* hitting surface. This practice is still in use in music as, say, rap artists used their mouth to provide the rhythmic track that accompany the spoken words. The imitation of clunking sounds can provide the necessary contrast that is typical of rhythmic changes in linguistic noises. The sequencing of vowel-like sounds punctuated by imitations of clunking sounds is not vocabulary but may sound like a song. In fact, in many cultures, baby's first introduction to language is through the songs of its mother. For example, the Maoris of New Zealand traditionally did not talk, but sang to their young for the first six month of their lives[18]. It is feasible to think that the refinement of social soundscape provided the vowel feature in linguistic interaction and the imitation of clunking sounds provided the break in rhythm typical of the consonant feature. We can also assume that

the event of consonant type noises to social soundscape most have happened much later on, maybe much after there was a kind of sequencing of vowel type noises and pointing since many ancient languages, such as Hebrew, do not script vowels at all but only consonants. The belief might have been that vowel sounds do not provide much context for an interaction as they are *naturally* shared in a population but that the consonant sounds are a kind of previously indexical technology that provide detail. Vocal punctuation is a pre-functional use of what would usually be part of a proto-attribution. These pre-functional vocal cues are, later on, related to prosodic flow in linguistic interaction.

8.4 Attenuation in Neural Terms

It may be supposed that we ought to be able to cash out the notion of attenuation in purely neural terms. If at an early stage in its life cycle a word does one job and at a later stage another, then we should expect that the neural reception of the word should be different at these different stages. In the course of these changes we should expect the word to turn up in different syntactic roles and therefore in positionally distinguishable roles in speech. This is certainly true of logicalized vocabulary that has relational uses at an early stage and connectival roles at a later. In this case the differing neural responses ought to be as trackable, as neural receptors are syntactically specialized.³ But if this thesis is correct, then there ought also to be changes in the neurological effects within single syntactic types, changes that correspond specifically, for example, degrees of abstraction, *chair* as concrete referent and *chair* as name of academic position or organizational role, the ambiguity of *the chair is all wet* being expressible in neural terms. Again, we ought to expect there to be differences in neural effects of, say, philosophical vocabulary upon

- (a) someone who continues to rely on a conversational understanding, and
- (b) upon someone who has listened to and grasped a technical definition. In principle we can provide a measure of attenuation based on what is required to regain early neural

³We know that aphasia is a pathology that affects several aspect of speech depending on the area of the brain that has been damaged. Damage to Broca's area impairs speech production while damage in Wernicke's area will affect speech production. But these are not the only form of aphasia. Damage in the right hemisphere of the brain can affect the production and comprehension of jokes, puns, prosodic flow, and *inferential* effects. Though some damage in specific areas of the brain may yield different pathologies, it is, non-the-less difficult to differentiate between types of aphasia. This is because most of our language skills are the result of complicated neural dynamics and can be traced through both cortical and subcortical regions [31].

effects. The claim is that we need contextual material to produce these effects. The contextual material needed to regain early neural effects may include something more or different from composition. In some cases such as *the chair is all wet*, there is a loss of compositionality because the sentence itself cannot provide the necessary inferential resources to enable disambiguation of the instance of chair that is part of this sentence. Only additional perceptual cues such as prosodic variation or additional sentences can disambiguate the sentence.

In the case of prosodic clues, in general they have their effect not through differentiations at the level of individual words, but through presentational modifications of a larger grouping of words or sentences taken as a whole. For example, the normal presentation of the string *world series* (world SERIES) is such as to suggest a discrimination among global competitions (the world series, the world cup, the world championships), whereas historically the label served to distinguish among (perhaps newspaper) sponsors (the *World Series*, the *Globe and Mail Series*, the *Province Series*, and so on), and was initially prosodically marked accordingly. In some cases (such as this one) the generality of language users do not notice that the vocabulary retains only an historic connection with the semantic construal that spawned the construction. In other cases the connection is even more remote. The *or* in interrogatives is a semantically puzzling construction and it is not entirely clear what role the *or* plays in, say,

Is it in the drawer OR in the laundry: it certainly cannot be represented as an operation upon logical disjunction.⁴

The breach - loosely, the originating semantic content - is also a feature of attenuation. Not only does the extension of a term increase with use but, in successive generations of language users, it is some fragment of that increased extension - rather than some earlier semantic pulse - that guides the next use. It is also a fragment of that extension that guides an interlocutor's understanding of the term. It may therefore be expected that, changes in neural effects of occurrences of words track this loss of connection. In many cases, it may also be that attenuation carries the vocabulary beyond all but forgotten historical connection with compositionally dependable semantics.

⁴This is an oversimplification. There might be a use of that sentence which seeks only a *yes/no* response, but such a use would require a distinct prosody or some additional such modification as to

Is it either in the drawer or in the laundry.

Chapter 9

Conclusion

With this thesis, we have offered a physical theory of language that, we think, can withstand the scrutiny of scientific inquiry. We have also set out to establish the usefulness of a physics model with which to talk about the changes that occur in language evolution resulting from linguistic transactions.

We have presented the need for such an approach by establishing why a semantic theory is inadequate in describing most of language. We have also established that there are precedents for a physical theory of language in the research of several linguists and cognitive scientists such as Derek Bickerton and William H. Calvin.

In the effort of capturing salient features in language evolution it became necessary to introduce new vocabulary, such as *attenuation* and borrowed vocabulary such as *phase transition*. We have used these notions to describe the erosion of perceptual cues that correlates the use of vocables, erosion that, pushed to a critical point, will transform some *lexical* vocabulary into *functional* vocabulary.

In conducting this research, it became apparent that the small-scale dynamics of change that can be observed in the history of vocables could be mapped onto much larger-scale dynamics of change such as the changes that could bring about the phenomenon of grammaticalization.

Jennings' theory of connective vocabulary was our starting point and the inspiration for this research. Jennings has described how connective vocabulary is vulnerable to *misunderstandings* in its scope arrangements. These misconstruals may lead to mutations that give rise to new uses in functional vocabulary. He also establishes the ancestry of *functional* vocabulary in *lexical* roots and explains how these findings highlights the inadequacy of a

semantic theory to describe this vocabulary. We then supposed that the dynamics - attenuation - that brings about functional vocabulary is also responsible for many changes in the use of all vocables. Moreover, that attenuation is exploited toward the propagation of linguistic systems. This assumption has led to the introduction of some physics concepts, the first one being *efficiency*.

The concept of *efficiency* was introduced as a driving force in the dynamics of language evolution. Efficiency is a phenomenon that many living systems are governed by. We have described how generally we use what we already have for novel tasks rather than generate new tools for novel situations. We have described how people, in their use of language, use a similar strategy. For reasons of efficiency, language becomes transformed with its use in a population. The dynamics of shareability extend vocables beyond original perceptual cues while still generating specific effects.

The transformation of vocables involves several dynamics, many of which can be described using other physical constructs. We have borrowed terms from classical and statistical mechanics and demonstrated a direct correlation with the dynamics of language evolution. We have used the concepts of *propagation*, *dispersion*, *elasticity*, *degeneracy* and *phase transition* to illustrate how vocables become *attenuated*. We suggested that *attenuation* in vocables will occur as a consequence of a *negotiation* of context between the members of a group, hence generating *coherence* for the use of vocables. This negotiating process involves dynamics that can be described using the physical concepts that we have mentioned previously. In order to formalize our approach we have used modeling, particularly the modeling of phase transitions.

In an effort to support our claims we have referred to the work of several researchers to show how models of phase transition can describe successfully certain dynamics characteristic of many phenomena that can be found in biology, economics and physics.

We then have focused on the description of a standard Ising model. We also introduce our modified version. We have demonstrated that changes in language can be illustrated by a simple, mathematically tractable model such as the Ising model. Though somewhat limited, it is relevant to use such a model to explore further aspects of language, such as first-order phase transitions, that are not so easily historically tractable. We have suggested that language can be thought of as a system in quasi-equilibrium, a system that changes, however, slowly enough to allow a population to share it. Finally, we have shown that our heuristic approach to the Ising model was viable and that our results prompted us to explore

more sophisticated models, compatible with our thesis.

Formal scientific inquiry requires of a theory that it be compatible with a larger body of knowledge. We have shown that our thesis is compatible with a microscopic account of language given at the neural level.

We called upon the work of William H. Calvin to help us in this task. We have found that our account of linguistic transaction is compatible with his microscopic *representation* account based on neural firing pattern replication. We have suggested that functional vocabulary can result from certain kinds of neural pattern replication or partial replication. Calvin also uses the concept of *phase transition* to describe the resolution of competition between synaptic structures. Additionally, we suggested that Calvin's model could benefit from our approach by including additional concepts such as *percolation*. This could explain the emergence of stable patterns in low connectivity, patterns that could account for some kinds of generalizations and perhaps certain types of vocabulary such as functional vocabulary.

The requirements of a successful theory requires that our approach can be applied to a macroscopic view of language, as well, so we have tested our theory for the context of the emergence of language. Most of our theory has based its premises on a diachronic account of vocabulary. But our theory has also accounted for a time before vocables, when the dynamics we have describes may have been present in the soundscape of hominids and how they might have led to proto-language. We have pondered on how naturally occurring sounds and vocalizations may have been slowly exploited toward prosodic features of language. We also have elaborated on the role of prosody and how functionalization of some vocabulary may play a role in preserving prosodic flow.

From a generalized concept of functionalization, we have supposed a transference in priority, from visual cues, towards auditory cues. This possible ancestral dynamics may have led to shared communication systems. We suppose that vowels may be the descendent of natural vocalization while consonance may be the result of vocally reproducing rhythmic punctations that occurs with say, hitting a surface.

In closing, we have claimed that our approach is universal by suggesting that the concept of *attenuation* can correlate a macroscopic view of the emergence of proto-language in a population and microscopic view of this phenomenon, in neural terms.

Appendix A

The Physics of Language: The Fabric Metaphors

A.1 Quine's Fabric

Recall Quine's eloquent description:

The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Re-evaluation of some statements entails re-evaluation of others, because of their logical interconnections; the logical laws being in turn simply certain further statements of the system, certain further elements of the field. Having re-evaluated one statement we must re-evaluate some others, whether they be statements logically connected with the first or whether they be the statements of logical connections themselves. But the total field is so undetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to re-evaluate in the light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through

considerations of equilibrium affecting the field as a whole.

If this view is right, it is misleading to speak of the empirical content of an individual statement, especially if it be a statement at all remote from the experiential periphery of the field. Furthermore it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements which hold come what may. Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle? -Quine, 1951, p.41-2-

The reason for setting out Quine's metaphor is not to adopt it, but to set it aside. Quine's is not a fabric of sentences except indirectly; it is fundamentally a temporal fabric spun rather than woven, and a convenient image intended to apply in some measure to all of the vocabulary that natural language throws up.

Though Quine's metaphor does not apply, the general concept of a fabric or field does; related is the temporal relationship of a vocable's later uses, relative to its early uses. On the account presented here, roughly parallel remarks apply to the elements of natural language on a metric of attenuation. They can be thought of as constituting a field. The notion of centre and periphery have no direct application, but we may think of the relevant development as it affects a single item as beginning where the vocabulary originates and broadening gradually, roughly speaking in such a way as to track the growing set of items to which the vocabulary has been applied. In the case of adjectives, we may think of tracking the growth of their extension. In the case of binary relational vocabulary as tracking the growing family of pairs of items that are referred to in its attributive uses. There is no periphery of this fabric. But, we claim, there is a function whose values increase monotonically with this extensional growth; for as we move farther away from the origin, we find uses of the vocabulary that are less and less dependent upon the uses nearer to

its point of origin. However, those farther from the point of origin are more dependent upon others relatively far from their points of origin. Beneath the fabric lies all of the non-linguistic physical activity upon which the linguistic material initially depends for its effects.

A.2 Our Fabric

The fabric of our metaphor is itself evolving. As in Quine, the vocables in our metaphor are related to experience, but the nature of the relationship has more to do with population dynamics than semantics. In fact most of what we call meaning is the result of exploiting behaviors, behaviors that are reinforced in a social setting, some of which have evolved into linguistic interactions, others have not. Our fabric evolves because of human interactions giving rise to linguistic behavior from non-linguistic behavior. But the emergence of a category such as linguistic behavior is reflected, later on in the history of language, in the ever-growing number of linguistic categories that emerge from strictly linguistic interactions. Moreover, we can assume that present linguistic categories are less related to non-linguistic behavior than linguistic categories were in, say, proto-languages, and in some cases not related at all. For example, in the case of logicalized connectives, such as *or* and *and*; the vocabulary has lexical roots, but the truth functional has emerged purely from the dynamics of linguistic interactions and the ability of linguistic transactions to coordinate human activity.

One could say that, at the periphery, the distinction between language and non-language is minimal.

Quine's model follows physics, unconstrained but consistent with interactions in the core with facts (or experience, in this case) defining the behavior of the edge. Quine describes a semantic field, the edge of which is supported in a certain way by experience. An incidental consequence of such an image is that it imposes an actual *boundary* (one that we may presume him to have implicitly acknowledged) between language and non-language, and so between linguistic and non-linguistic behavior. From this perceived constraint Quine defines semantic consequences that are derived from the interaction at this *edge*. Quine's notion of bounded field and resulting behaviors borrows indirectly from physics. To assume a boundary is to stipulate fixed conditions for some of the constituents of the system (those that constitute the boundary) that distinguishes them from those of the interior of the

theoretical system. Through the defined interactions of the system, the behavior of the boundary constituents affect the behavior of the system as a whole. In Quine's metaphor, experience constitutes a dynamic but functionally fixed condition. The semantic content of sentences at the edge may change (because sentences may come and go), but they have a semantic status which is constant because of their constant relationship to experience.

The nature of the fabric (core, periphery, dependencies) is laid out. Elsewhere in the fabric and increasingly as we approach the centre, semantic status is also relatively unaltered, but, again, as we approach the centre, semantic content is also increasingly unchanging. If the vision of language that underlies the metaphor has sentences of mathematics at the centre, then one might say that the semantic content of the sentences at the centre is fixed because it is null or nearly so. It is relatively impervious to the influence of changes at the periphery, but not absolutely so. And again, changes at the periphery are relatively independent of one another. Changes nearer the centre more fundamentally disturb the structure of the fabric as a whole.

Quine's metaphor is, of course not a metaphor principally about language, but about human understanding as it expresses itself in language. The relative peripherality or centrality of positions within the fabric represent the status of sentences that are taken to be true, that, taken as a whole, constitute a view of how things are.

Our fabric is distinguishable from Quine's by virtue of the elements described in it; in our case they are positional elements referring to the relationship between vocable uses. Our fabric or field is not semantic as Quine would have it. Only comparatively and indirectly is it connected with the truth values of sentences. Our fabric registers positions of vocabulary, and therefore of sentences in which they occur, independently of whether those sentences are true or false. It records movements of vocabulary from uses reflecting the particularities of experience that give rise to them to the roles they retain when those particularities have long disappeared from the collective memory of the speakers of the language. Thus it records a loss of information. But it is not primarily synaptic either, though Deacon's account [18] of three kinds of meanings (iconic, indexical, and symbolic) could plausibly be mapped onto some such structure; iconic meanings being the elements at the periphery, then indexical meanings in intermediate position and symbolic meanings at the center. The notion that a progression in evolutionary time from an early stage of language development towards abstraction should be paralleled in developments of synaptic function in individuals is by no means incompatible with our account; however our point is a different one: the

shared nature of language in a population shapes linguistic transactions in ways that are independent of the synaptic effects that occur in individuals involved in the transactions.

Our fabric also assigns a different understanding of centre and edge. In fact our fabric has no edge. Movement is from the centre, but what is recorded is *attenuation*. As we move from the centre, we encounter no boundary, only a gradual diminution of a semantic value, a motion registrable in the closed unit interval. At the centre that value is 1; but it drifts inexorably toward 0.

The fabric is woven from relationships between vocable uses and their diachronic account. This fabric grows with succeeding generations of users, adding and subtracting uses, but its internal constitution is shaped by the collective *memory* of the whole body of its users. In fact, we may say that *the memory of the fabric* is the collective memory of its users. What we mean is: Uses of elements of the language are determined partly by previous uses and partly by whatever initiates change. But for successive generations of users of the language, there is less and less contact with earlier uses. For elements near the centre, use is determined by relatively short-term memory; for those toward the periphery, likewise. But for elements at the centre, short-term memory has the ingredients for semantic understanding based in relatively non-linguistic material. In later stages, experience and relatively earlier uses of elements play a diminishing role in the regulation of new uses. This is mostly because the history of uses is slowly lost over time, often forgotten as successive generations inherit the vocabulary from previous generations. The loss is, however, relatively smooth and slow, since the changes, though discrete, are discrete on the scale of individual births and deaths, that is, well below the scale of observable general linguistic changes or trends. As we have previously mentioned, the attenuation of uses is statistically unidirectional, the direction being that toward the growth of the fabric. By contrast, in Quine's metaphor, re-evaluations of the truth-values of sentences can occur, although against varying degrees of resistance, anywhere throughout the fabric. Moreover, it is not clear that in the Quinean interconnectedness of meanings there is any concept of *temporality*, *entropy*, *equilibrium* that govern the area from the periphery to the center. On our account, elements acquire new uses as they are further away from the centre and in most cases the interconnectivity of uses generate unidirectional attenuation towards the virtual periphery. As elements drift away from the centre there is also a gradual loss of non-idiomatic compositionality. In functionalized uses of auxiliary verbs, for example, we do not consult their surviving lexical uses in constructing tensed forms of verbs. But, for example, by the time the verb *have* was

given an auxiliary role, its non-auxiliary uses had already lost their capacity to occasion very specific neural responses independently of linguistic environment. Much the same could be said of the auxiliary *do*. A sentence composed from elements that have already moved toward the periphery cannot generate any specific inferential effects unless additional information from the periphery is added.

In fact the notion of periphery is ill-applied here. We illustrate the movement of vocabulary uses towards a *horizon*, with specific examples. We might better speak of an *horizon*, since beyond it, elements of language cease to have effects of a specifically linguistic nature. So for example at the present point in the development of English, *wycg* (horse), absent very strong contextual clues amounting virtually to definition, would have auditory effects, but is essentially foreign. Like much of the rest of the Anglo-Saxon lexicon, it has passed the periphery into complete disuse. Some of this vocabulary is superseded by other lexical vocabulary, but some has simply disappeared with the need for it. Other vocabulary approaches this horizon essentially through our ceasing to require a semantic grounding for its use. Some of this vocabulary approaches the horizon through functionalization. Auxiliary verbs (*have, is, do*) and some connective vocabulary (*or, but, if*) are examples of this. But still other vocabulary approaches it simply through a loss of connection with its more concretely grounded origins. Into this category falls much of the philosophically charged vocabulary, (*intention, mind, good, right, truth, fact*), but other vocabulary as well (*as well, indeed, very, mere, even, just* and so on.)

The dynamics are so gradual as to allow the system to be in quasi-equilibrium on timescales most reasonable for several generations of language users to interact successfully. Our claim is that movement of vocabulary toward a horizon is an inexorable feature of language-change. It is sufficiently demonstrated by the observation that, whatever the rate of change of a language, (and that itself may change) it must eventually be so altered that its practitioners can learn the earlier form only as a second language - Anglo-Saxon for example. Nevertheless, the change is a gradual one; every generation of users inherits a language in which nearly all stages of this process are represented. To be sure, some vocabulary may pass from innovation to extinction with few recognizable intervening stages. But for much of the vocabulary of a language, the process is a gradual one involving many degrees of semantic attenuation.

Bibliography

- [1] William P. Alston. Philosophy of language. In Paul Edwards, editor, *Encyclopedia of Philosophy*, volume IV. MacMillan and Free Press, New York, 1967.
- [2] Jane Austen. *Northanger Abbey*. First published 1818, Wordsworth Editions, Ware, Hertfordshire, 1993.
- [3] Sunny Y. Auyang. *Foundations of Complex-System Theory: in Economics, Evolutionary Biology, and Statistical Physical*. Cambridge University Press, Cambridge, 1998.
- [4] Philip Badcock cove and Miriam-Webster, editors. *Webster's Third New International Dictionary*. Miriam-Webster, January 2002.
- [5] Simon Baron-Cohen. Is there a normal phase of synaesthesia in development? In John E. Harrison, Simon Baron-Cohen, and Ian McDonald, editors, *Synaesthesia: Classic and Contemporary Readings*. Blackwell Publishers, 1997.
- [6] Bernard H. Bichakjian. Essentialism in language: a convenient but fallacious premise. In Stanislaw. Puppel, editor, *The biology of language*, pages 33–60. John Benjamins Publishing Company, Amsterdam/Philadelphia, 1995.
- [7] Bernard H. Bichakjian. Language diversity and the strait flush pattern of language evolution. *Jazyki i Rechevaja dejatel 'nost' (Language and Speech Activity)*, 2:18–34, 1999.
- [8] Bernard H. Bichakjian. Language evolution and the complexity criterion: Target article on language complexity. *Psychology*, 10(033), 1999.
- [9] Derek Bickerton. *Language & Species*. University of Washington Press, Chicago & London, 1990.

- [10] Derek Bickerton. *Language and Human Behavior*. University of Chicago Press, Seattle, 1995.
- [11] William H. Calvin. *The Cerebral Code: Thinking a thought in the mosaics of the mind*. MIT Press, Cambridge, Massachusetts, 1996.
- [12] Geoffrey Chaucer. *Selection from the Tales of Canterbury, and short poems*. Houghton Mifflin Company, New York, 1966.
- [13] Noam Chomsky. *Syntactic Structures*. Mouton, The Hague, 1957.
- [14] Noam Chomsky. *Knowledge of Language: Its Nature Origin, and Use*. Praeger, New York, 1986.
- [15] John M. Cooper and Hutchinson D. S., editors. *Complete Work / Plato*. Hackett Pub., Indianapolis, 1997.
- [16] Michael C. Corballis. *From hand to Mouth; The Origins of Language*. Princeton University Press, Princeton, May 2002.
- [17] Richard Dawkins. *The Selfish Gene*. Oxford University Press, New York, 1976.
- [18] Terrence W. Deacon. *The Symbolic Species: The co-evolution of language and the brain*. W.W.Norton & Company inc., New York, London, 1997.
- [19] M. C. Diamond, A. B. Scheibel, and L. M. Elson. *The Human Brain Coloring Book*. HarperPerennial, New York, 1985.
- [20] Jeffrey Elman, E. Bates, M. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking Ennateness: A Connectivist Perspective on Development*. A Bradford Book, The MIT Press, Cambridge, MA, London, England, 1996.
- [21] Murray Gell-Mann. *The Quark and the Jaguar*. W. H. Freeman and Company, New York, 1994.
- [22] Anna Giacalone Ramat. Testing the boundaries of grammaticalization. In Anna Giacalone Ramat and Paul J. Hopper, editors, *Typological Studies in Language 37: The Limits of Grammaticalization*. John benjamins Publishing Company, Amsterdam/Philadelphia, 1998.

- [23] Anna Giacalone Ramat and Paul J. Hopper. Introduction. In Anna Giacalone Ramat and Paul J. Hopper, editors, *Typological Studies in Language 37: The Limits of Grammaticalization*. John benjamins Publishing Company, Amsterdam/Philadelphia, 1998.
- [24] James Gleick. *Chaos: Making a New Science*. Penguin Books, New York, 1988.
- [25] Harvey Gould and Jan Tobochnik. *An Introduction to Computer Simulation Methods: Applications to Physical Systems, Part 2*. Addison-Wesley Publishing Company, Massachusetts U.S.A., 1988.
- [26] Paul Grice. *Studies in the Ways of Words*. Harvard University Press, Cambridge, Massachusetts, London, England, 1989.
- [27] Abhijit Kar Guptay and Dietrich Stauffer. Social percolation on inhomogeneous spanning network. *arXiv:cond-mat*, 60-12, June 2000.
- [28] John Haiman. Ritualization and the development of language. In William Pagliuca, editor, *Current issues in linguistic Theory, 109: Perspective on Grammaticalization*. John benjamins Publishing Company, Amsterdam/Philadelphia, 1994.
- [29] Joseph Heller. *Catch 22: A Novel*. Simon and Schuster, New York, c1961.
- [30] James R. Hurford. Nativist and functional explanations in language acquisition. In Roca I. M., editor, *Logical Issues in Language Acquisition*, pages 85–136. Foris publications, Dordrecht, Holland/Providence RI, USA, 1990.
- [31] Daniel H. Jacobs. Aphasia. In Joseph Quinn, Francisco Talavera, Richard J. Caselli, Matthew J. Baker, and Helmi L. Lutsep, editors, *eMedicine*. www.emedicine.com, 2002.
- [32] R. E. Jennings. *The Genealogy of Disjunction*. Oxford University Press, New York, 1994.
- [33] R. E. Jennings and P. K. Schotch. The preservation of coherence. *Studia Logica*, 43 1/2, 1984.
- [34] D. K. Johnston. The natural history of fact. to be published.

- [35] U. Jurgens. A comparison of the neural systems underlying speech and nonspeech vocal utterances. In *Becoming Loquens eds.*, pages 1–13, 2000.
- [36] Immanuel Kant. *Critique of Pure Reason*. A51/B75, Germany, (first edition) 1781.
- [37] Stuart A. Kauffman. *The Origins of Order*. Oxford University Press, New York, 1993.
- [38] Stuart A. kaufman. *The Origins of Order*. Oxford University Press, NewYork, 1993.
- [39] Saul Kripke. Identity and necessity. In A. W. Moore, editor, *Meaning and Reference*. Oxford University Press, Walton Street, Oxford, 1993.
- [40] David K. Lewis. *Convention: A Philosophical Study*. Harvard University Press, Cambridge, Massachusetts, 1969.
- [41] Philip Lieberman. On the evolution of human language. In John A. Hawkins and Murray Gell-Mann, editors, *The Evolution of Human Language*, volume XI, Santa Fe, September 1997. Santa Fe Institute.
- [42] Edwin G. Morgan. *Beowulf: a verse translation into modern English*. University of California Press, Berkley, 1966[c1952].
- [43] Elissa L. Newport. Maturational constraints on language learning. *Cognitive Science*, 14:11–28, 1990.
- [44] William Pagliuca. Introduction. In William Pagliuca, editor, *Current issues in linguistic Theory, 109: Perspective on Grammaticalization*. John benjamins Publishing Company, Amsterdam/Philadelphia, 1994.
- [45] I. P. Pavlov. *Condition Reflexes: an investigation of the physiological Activity of the cerebral cortex*. Oxford University press: Humphrey Milford, 1927.
- [46] Steven Pinker. *The Language Instinct*. WilliamMorrow, New York, 1994.
- [47] N. Prévost and R. E. Jennings. Language and intelligent understanding without semantic theory. *Canadian Artificial Intelligence (CAI)*, no 47, pp,9-16, Autumn 2000.
- [48] N. Prévost, R. E. Jennings, and L. M. Jorgenson. A visual model of mutation and stability in the evolution of connectives. unpublished.

- [49] W. V. O. Quine. *Two Dogmas of Empiricism*. Harvard University Press, Cambridge, MA, [1951], 1981.
- [50] Sonia Ragir. Indigenous sign languages do not follow the predicted course of grammaticalization. to be published.
- [51] L. E. Reichl. *A Modern Course in Statistical Physics*. University of Texas Press, Austin, Texas, 1980.
- [52] Stephen D. Richardson, Lucy Vanderwende, and William Dolan. Combining dictionary-based and example-based methods for natural language analysis. *Microsoft Research*, tr-93-08, 1993.
- [53] G. Rizzolatti, L. Fogassi, and V. Gallese. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2:9:661–70, 2001.
- [54] Bertrand Russell. *An Inquiry Into Meaning and Truth*. Bishop and Sons limited, Edinburgh, 1961.
- [55] Ramachandran V. S. and Hubbard E. M. Synaesthesia - a window into perception, thought and language. *Journal of Consciousness Studies*, 8:12, 2001.
- [56] Sue E. Savage-Rumbaugh and Roger Lewin. *Kanzi: The Ape at the Brink of the Human Mind*. John Wiley, New York, 1994.
- [57] M. Sreenivasa Roa and Arun K. Pujari. Word sense disambiguation using the hopfield model of neural networks. *Journal of Intelligent Systems*, 9-no.3/4, 1999.
- [58] Patrick Suppes. *Introduction to Logic*. Princeton, 1957.
- [59] Alfred Tarski. *Introduction to Logic and to the Methodology of the Deductive Sciences*. (Revised 1946 edition), New York, 1941.
- [60] John Venn. *Symbolic Logic*. 2nd edition, New York, 1894: First published 1881.
- [61] Henry B. Wheatley. *The Diary of Samuel Pepys*. The Heritage Press, New York, 1942.
- [62] S. T. Wierzchon. Ising model. In Eric Weisstein, editor, *Eric Weisstein's World of Physics*. scienceworld.wolfram.com, 2002.