STRUCTURE AND PROCESS:
PROSPECTS FOR THEORIES OF COGNITIVE SCIENCE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYŞEGÜL ÖZKAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN
THE DEPARTMENT OF PHILOSOPHY

DECEMBER 2010

Approval of the Graduate School of Social Sciences

Prof. Dr. Meliha Altunışık
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.

Prof. Dr. Ahmet İnam
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts.

Assoc. Prof. Dr. Erdinç Sayan
Supervisor

Examining Committee Members

Prof. Dr. Ayhan Sol	(METU, PHIL)	
Assoc. Prof. Dr. Erdinç Sayan	(METU, PHIL)	
Assist. Prof. Dr. Annette Hohenberger	(METU, COGS)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Ayşegül Özkan

Signature :

ABSTRACT

STRUCTURE AND PROCESS:
PROSPECTS FOR THEORIES OF COGNITIVE SCIENCE

Özkan, Ayşegül
M. A., Department of Philosophy
Supervisor: Assoc. Prof. Dr. Erdinç Sayan

December 2010, 83 pages

Different theories of cognitive science propose different system descriptions in their models for the explanation of cognitive phenomena. According to one view, they are incompatible and competing theories. The view is defended by theorists and philosophers from different perspectives and they all claim that the proper conception of cognition is the conception provided by the theory which they advocate. The other view, on the other hand, insists on the compatibility of those theories. According to this view which is also defended here, these different theories are not only compatible, but also they are complementary. The cooperation of these theories and integration of the conceptions provided by these theories are needed to have a full account of cognition.

Keywords: cognition, system description, decomposition, ontological categories, idealization

ÖZ

YAPI VE SÜREÇ:
BİLİŞSEL BİLİMLER TEORİLERİNE İLİŞKİN ÖNGÖRÜLER

Özkan, Ayşegül
Yüksek Lisans, Felsefe Bölümü
Supervisor: Yrd. Doç. Dr. Erdinç Sayan

Aralık 2010, 83 sayfa

Bilişsel fenomenlerin açıklaması için, bilişsel bilimler teorilerinin modellerinde farklı sistem betimlemeleri önerilir. Bir görüşe göre bunlar bağdaşmayan rakip teorilerdir. Görüş, farklı perspektiflerden teorisyen ve felsefecilerce savunulmuş ve hepsi bilişin en uygun kavranışının kendi savundukları teorinin sağladığı kavrayış olduğunu iddia etmiştir. Diğer görüş ise bu teorilerin bağdaşırlığı üzerinde ısrar eder. Burada da savunulan bu görüşe göre bu farklı teoriler sadece bağdaşır değil, aynı zamanda tamamlayıcıdır. Bu teoriler arasındaki işbirliği ve bu teorilerin sağladığı kavrayışların entegre olması bilişi tam olarak anlayabilmek için gereklidir.

Anahtar Kelimeler: biliş, sistem betimlemesi, ayrışım, ontolojik kategoriler, idealizasyon

To My Parents

ACKNOWLEDGMENTS

I wish to thank my supervisor Assoc. Prof. Dr. Erdinç Sayan for his advice and especially encouragement. I also wish to express my deepest gratitude to Assist. Prof. Dr. John Bolender, who read each and every draft and helped in the development of this thesis, though he was abroad, and Assist. Prof. Dr. Annette Hohenberger and Prof. Dr. Ayhan Sol for their valuable comments, criticisms, suggestions and support.

I also wish to thank my director, Selda Bilgin, administrators, Yasemin Koçak and Süheyla Duru, and colleagues, Dilek Çınar, İnci Temizer, Şahin Arslan, Öznur Soylu, Burcu Demiralay, Sedal Çapa and Seval Yavuz at Academic Personnel Affairs Office at Directorate of Personnel Affairs of METU for their tolerance and helpfulness during the thesis writing period. I am very grateful to my parents for protecting me from the problems of daily life.

I thank my love Gökhan Akbay for his psychological support.

TABLE OF CONTENTS

CHAPTER 1:

INTRODUCTION

Alan Turing in 1950 has started a perennial discussion with his question, "Can machines think?"[1] His answer was affirmative; however, he was defining a new thinking concept: the digital computing procedure. He proposed an abstract machine, namely the Turing Machine, to explain that procedure. After the revolutionary conception of intelligence of Turing, Computational Theory of Mind (CTM) was founded as the explanation of cognitive architecture which clearly accepts this new thinking concept which is a componential conception of cognition with discrete change. However, we have another conception of cognition – i.e., integrated conception which can exhibit discrete or continuous change. The other two theories which I will discuss in my thesis, namely Parallel Distributed Processing (PDP) and Dynamical Systems Theory (DST)[2], use the latter type of conception.

I will attempt to investigate the relations between these theories of cognitive science over their conceptions of cognition and whether one of these conceptions is enough in understanding cognition. I claim that different aspects of cognition might demand different assumptions about their nature. Thus, we might need cooperation between those conceptions and hence those theories of cognitive science to be able to have a full account of cognition. In explaining relations between different conceptions of cognition and my claims about those relations, I will focus on the following two topics in my thesis:

1. The distinction between componentiality vs. integrity and discreteness vs. continuity (C/I-D/C) of systems: I assume that being a componential system is about

---

1 Turing, p. 433.

2 'DST', in this thesis, is not used to refer to one dynamical theory, exactly; rather to dynamical approach to cognition (see Abrahamsen &Bechtel, p. 174.)

the structure (or architecture) of that system, while discreteness of the system is about the change in the system, or the evolution of the system. Thus, the distinction of componential and integrated systems (C/I), and that of discrete and continuous systems (D/C) are closely related but different distinctions.

I use the terms 'discrete' and 'continuous' in the sense that they are used in mathematics. In other words, I regard a system as continuous, if the relevant change in the system (in time, or between states) is incalculably minute. In this sense, it is typical for mathematics that a difference equation is a discrete expression but a differential equation is a continuous one.

Componential systems can be described as systems divisible into functional subsystems, i.e., they are nearly decomposable. A componential system has basic, autonomous elements with intrinsic functions (the contribution to the system is specific to that component)[3] but an integrated system does not have such *elements*. Elements of integrated systems are simple in the sense that they do not have intrinsic functions and autonomous mechanisms, and are interdependent such that their function is determined by the organization of the system, i.e., they are minimally decomposable.[4] We cannot regard those elements in an integrated system as

3 In Abler, the properties of particles in a particulate system, namely 'sustained variation' and 'retrievability', can be seen as the properties of such autonomous elements (pp. 2-3). Thus, the notion of elements in componential systems explained here is consistent with that of particulate systems in Abler. Componential systems are also compatible with nearly decomposable systems of Simon and modular systems of Fodor. However, Fodor proposes a more specific notion with more properties than componential systems explained here. On the other hand, "modules" of Fodor correspond to specific cognitive faculties which are responsible for different cognitive tasks such as language (both its acquisition and usage). Those faculties can also be componential in their structure in the sense explained here. Shortly, I do not refer to a completely different kind of system than these three systems by the notion of 'componential systems', rather I refer to a similar system with a broader description.

4 Bechtel & Richardson, pp. 26-27. According to Bechtel and Richardson componential ('component' in their terms) and integrated systems are composite systems – none of them is simply decomposable or aggregative, and they both exhibit a degree of decomposability. These notions will be clearer in the next chapter.

autonomous elements in the sense that we use the notion for the elements of a componential system.

The processes of an integrated structure can be formulated discretely or continuously depending on the conception of time or change in the structure. Although the choice of the time scale depends on the phenomena under consideration and is not an arbitrary choice, we do not have to regard the system as componential while we are studying discrete changes of the system on larger time scales. Regarding the discrete change with respect to time in an integrated system as continuous, only due to integrity of the system, would trivialize the notion of discreteness. To regard such a discrete system as the synonym of componential system, would trivialize the notion of componentiality. A system can be an integrated system, but the evolution of the system can take place discretely in time. A componential system, on the other hand, can show continuous change due to one or more components of the system whose continuous changes affect the system. Working upon discrete items would lead a componential system to behave discretely. These notions will be clearer in the first chapter.

2. Controversies between CTM, PDP and DST: The core of this topic is idealization. This discussion will take place mostly in the light of idealization types. We deal with the problem of explaining the reality by idealizing and abstracting its objects, properties and relations. Theories propose a conception of reality in which reality is not represented wholly. The tools and the structure of a theory in representing real mechanisms depend on the theory's conception of the reality – the idealized reality. This process necessitates some distortion on real mechanisms in order to describe/explain them. Idealization is "… intentional introduction of distortion into scientific theories."[5] The nature and the justification of the distortion depend on the

5 Weisberg (2007b), p. 639. The definition of the idealization implies an idealization in a model. I will, as Weisberg does, refer to model-based theories by 'theory' while talking about the idealization kind of a theory; i.e., "… the indirect representation of real world phenomena with models." (Ibid, p. 640) According to him, there are two kinds of theorizing: modeling and abstract direct representation.

goals governing and guiding idealization.[6] For example a theorist who intends to represent the core causal factors would regard less factors relevant than the theorist who intends to represent the reality as complete as possible. However, even the theorist who has the goal of full representation of reality needs to distort real mechanisms while representing them.

For example, Turing justifies his idealization while asserting that "[e]verything really moves continuously …. but for most purposes we can forget about [intermediate positions]."[7] This is clearly introduction of a distortion into the theory; any procedure in reality is not exactly the digital computing procedure. However, by this idealization Turing enables us to see common properties and underlying mechanisms of a kind of process in which intermediate steps do not affect the outcome of that process directly. In the same way, elements in componential systems are idealized entities. The integrated and continuous systems are also idealizations of reality. We simply do not observe the reality purely as componential, integrated, discrete, or continuous systems like they appear in our theories. Our theories propose a conception of the reality in which reality is represented partially with false assumptions. However, this does not mean that they are useless. The use of false assumptions and models to construct true conceptions of reality will be discussed in the second chapter.

In modeling, reality is represented indirectly by construction of models and their analysis. Theorist starts with constructing a model, then he/she analyses and refines the model and finally he/she investigates the relationship between the model and the reality. If the model is similar enough to the reality then analyzing the model means analyzing the reality indirectly. In abstract direct representation, reality is represented directly by abstraction. Relevant properties of the real objects are abstracted and organized in a theoretical framework (Weisberg, 2007a). According to Weisberg, "Volterra's explanation of the cause of the post-WWI fish shortage [serves] as … example of modeling, and Dimitri Mendeleev's explanation of chemical properties via their periodic dependence on atomic mass [serves] as … example of [abstract direct representation]." (Weisberg, 2007a, p. 210) Though, whether the Periodic Table of the Elements is a theory or just a classification is a controversial issue. Since the theories under consideration are model-based theories and the examples which will be considered are models, I will not discuss the issue further.

6 Weisberg (2007b), p. 639.

7 Turing, p. 439.

The most important controversies between CTM, PDP and DST are the relevance of lower-levels, time and environment, and innateness. These theories have different accounts for the relevance problems and innateness. It is mostly because they have different conceptions of cognition. Although I will deeply investigate these issues in the second and third chapters, to put it roughly, I assume that all three theories, namely, CTM, PDP and DST, use different levels of explanation. Though, the conceptions of cognition proposed by them do not belong to different levels of organization ontologically; all of them are theories of a higher-level phenomenon. Different sub-targets of a target phenomenon might urge theorists to assume different conceptions of cognition. I claim that we have started to admit this kind of necessity and thus recent studies imply that different conceptions of cognition are assumed to be true in single models. Prima facie, admission of different conceptions together in a single model might mean that the model includes contradictory assumptions. This concern would be proper, if different conceptions of cognition are incompatible. I will attempt to show that they are compatible and indeed complementary throughout my thesis.

There are at least two construal of the relation between higher- and lower-levels: ontologically and epistemologically which are levels of organization and levels of explanation, respectively. According to ontological construal of levels, levels of organization constitute part-whole relation in which whole systems at lower-levels are parts at higher-levels.[8] Wimsatt introduces three main ontological categories, namely, levels of organization, *perspectives* and causal thickets which will be explained in the second chapter in detail. However, describing them roughly seems useful, now. The phenomena which belong to levels of organization category show part-whole relation among levels. Physical sciences, in general, imply this kind of ontology. For example, according to chemistry an atom in a chemical molecule is part of the structure which in turn is the whole system which is constituted by

[8] Wimsatt, pp. 201-202.

particles according to quantum physics.[9] *Objects* of the second ontological category, *perspectives,* can be explained via different disciplines however, these explanations do not imply a part-whole relation, even if they are different explanations of different levels. For example, there is an organism conception in genetics which cannot be seen as an explanation of a part in anatomical explanations. They are different explanations of the same *thing*, organism. They both contribute in our understanding of an organism, though not as a part-whole relation. Although these do not exhaust our conception of the phenomenon on their own, they are partially autonomous; they have different problems and methods specific to them.[10] In the conception of the phenomena which belong to the last category, i.e., causal thickets, contributions from different perspectives are not distinguishable. Theories which deal with the phenomena belonging to the causal thickets category would include methods and concepts of different perspectives in an undetectable manner.[11] The phenomena which interdisciplinary studies deal with such as financial markets according to econophysics mostly belong to this ontological category.

According to epistemological construal, lower-level theories are sub-symbolic explanations and higher-level ones are symbolic explanations.[12] Lower-level explanations are not explanations of parts of higher-level phenomena. Like in the case of genetic explanation and anatomical explanation of an organism, these kinds of explanations imply *perspectives* as ontological category. It seems that the second construal fits better to our discussion. Thus, I will mostly refer this second construal of the levels distinction throughout my thesis. I will not refer to levels of organization while saying higher- or lower-levels, rather to levels of explanation;

---

[9] Ibid, p. 202.

[10] Ibid, p. 227.

[11] Ibid, pp. 237-238.

[12] Abrahamsen & Bechtel, p. 163.

except the second chapter. The discussion of levels pertaining to cognition will be clearer in the second and third chapters.

Time, on the other hand, has importance in the distinction between C/I-D/C because the discrete or continuous nature of the change in a system is mostly about chosen time scale in integrated systems. Discreteness of a componential system is due to discreteness of the items upon which the system works (e.g. lexical items) and idealizing time as sequences in the system. Although an integrated system is usually conceived continuously, change over time in an integrated system can be discrete (when the important differences of the states of the system occur discretely in time), or *create* discrete change (when it *seems* componential and discrete at higher-levels).

To understand the relation between componentiality of higher-levels and integrity of lower-levels (in the epistemological sense mentioned above), studying on language acquisition models would be useful. It seems that language acquisition models are good examples for the use of different conceptions of cognition for different sub-targets of a target phenomenon. Similarly, John Bolender's twofold account on social cognition is a useful example in understanding the relation between the use of different conceptions of cognition and the nature of the phenomenon under investigation. Thus, my primary examples come from learning mechanisms, especially from language acquisition and social cognition.[13] Discussions pertaining to language acquisition and social cognition will take place in the fourth chapter, in detail. However, I wish to mention roughly, now. Language, at the level we perceive and use it, has a discrete and componential structure. However, the system which is responsible for acquisition of language and that of the language which will be acquired are distinct systems. Analogously, the *producer* and the *production* have

13 In the context of social cognition I will try to explain the research program proposed by Bolender which is a refined version of Relational Models Theory of Fiske (1993). According to research program proposed by Bolender, there is an early stage in which discrete items are produced by a dynamical process (Bolender, 2007, 2008) for the use of a computational component in the next stage (Bolender, 2010).

distinct systems, though they might (or, might not) have similar systems. So, the structure description of the former system need not be componential, just because of the componentiality of language. The content which will be acquired might (or, might not) have a componential structure. Regardless of the description of the outcome of the acquisition, it is both possible that the structure of the system responsible for acquisition is integrated or componential. The knowledge of language is acquired through a *process* in which the already existing structure changes. According to CTM, the system responsible for language acquisition is also componential. However, according to PDP, the system responsible for language acquisition is integrated.

In the first chapter, I will discuss the first topic mentioned above: the dual distinction (C/I-D/C) pertaining to system description. Firstly, I will attempt to distinguish two parts of a system description: structure description and process description. Structure description part of a system description is a description of the relations between components of the system. Process description part of a system description is a description of the change in the system/evolution of the system. The first distinction, namely, C/I, is a distinction about structure description part of the system description. Although none of the theories which I will investigate (namely, CTM, PDP and DST) proposes a simply decomposable/aggregate system, I will firstly explain simply decomposable systems which is necessary in understanding componential and integrated systems.[14] The less a system has a simply decomposable system's properties the more that system approximates an integrated system. Accordingly, decomposability is a matter of degree. I mean, for example aggregative, or simply decomposable systems show extremely decomposable structure. By the decrease in the degree of decomposability, or departure from aggregativity we reach to componential and integrated systems, respectively. However, composite systems (i.e., componential and integrated systems) are still decomposable systems, though at

14 Bechtel & Richardson, pp. 25-26.

different degrees of decomposability. D/C distinction, on the other hand, is a distinction about the process description part of the system description. Thus, in this chapter I will clarify the difference between C/I and D/C, and I will discuss the relation between them. In the last part of this chapter, I will propose two main conceptions of cognition according to these two distinctions (C/I and D/C) and explore the conceptions of cognition of these three approaches which are CTM, PDP and DST.

In the second chapter, I will focus on idealization. As it is known, scientific knowledge is an approximation to the reality, or the representation of it[15], essentially because of the complexity of reality and/or the limited capacities of human minds – *even* when they are scientists.[16] Without doubt, the phenomenon under consideration, namely cognition, is a complex one. Thus, I will attempt to explore what kind/s of idealization/s we need to understand this complex phenomenon. Some questions which I wish to answer in this chapter are: Which aspects of the cognition can be neglected legitimately? Is it possible to explain cognition under one theory or one conception of cognition?

Chapter three will be about the controversies between CTM, PDP and DST. I will turn to topic of idealization in this chapter and try to clarify my thesis with help of the examples from language acquisition and social cognition.

In the fourth chapter there will be two example phenomena: language acquisition and social cognition. In my opinion, the theories of both language acquisition and social cognition support the claim that emergence processes of cognition and emergent properties of it might necessitate cooperation of different conceptions of cognition to be able to have a full account of cognition.

---

15 Cartwright, pp. 192-195.

16 Wimsatt, pp. 5-6.

In conclusion, the goals of this thesis are to investigate firstly, the relation between the two distinctions mentioned above: C/I and D/C. According to that dual distinction, there can be two conceptions of cognition: componential structure with discrete or continuous processes, and integrated structure with discrete or continuous processes. The properties of cognitive phenomena and the mechanism which produce those properties – i.e., the emergent phenomena and the emergence of those phenomena – necessitate different conceptions of cognition and hence different causal factors to be taken into account. Accepting that the only interest in understanding cognitive processes should be directed to those emergent properties does not seem enough to understand cognitive phenomena. Thus, cooperation between different perspectives of theories (namely, CTM, PDP and DST) would be better to understand the cognitive structure and processes; rather than competition between them. I think they can form a unified approach in the long-run with coevolution of all theories if that cooperation is pursued.

CHAPTER 2:

STRUCTURE AND CHANGE IN A SYSTEM

I will try to make a distinction between the structure of a system and the change in it. Prima facie, this distinction might be trivial: "Change is defined relative to a period of time, while structure is defined relative to a point in time. Those are not full definitions, but they suffice to capture the difference between the two."[17] However, I will try to provide their definitions explicitly to be able to distinguish structure descriptions and process descriptions as components of system descriptions of different theories of cognitive science. I will do so to be able to clarify the different conceptions of cognition according to three main theories of cognition, namely, CTM, PDP and DST. These three approaches propose three different kinds of system descriptions with different structure and process descriptions for cognitive phenomena. However, these three systems can be grouped under two basic conceptions of cognition: componential and integrated conceptions of cognition. "The kind of change these systems may incur – discrete or continuous – represents a second dimension according to which a system can be described, namely the process dimension."[18] In my opinion, both of the conceptions of cognition, i.e., cognition as a componential system and cognition as an integrated system, might be useful, or even necessary to be able to have a full account of cognition. In this chapter, however, I will only try to explain the differences between these two conceptions for using them to support my thesis.

Throughout my thesis, I will refer to the notion of 'system' and its components (e.g. variables) which are described by van Gelder as follows:

---

[17] John Bolender, personal communication.

[18] Anette Hohenberger, personal communication.

> Systems are … taken to be sets of interdependent variables. A *variable* is simply some entity that can change, that is, be in different states at different times. …. The *state* of the system is simply the state or value of all its variables at a time; the *behavior* of the system consists of transitions between states.[19]

There are factors outside the system on which change of the system also depends: parameters.[20] Usually, changes in parameters affect the overall structure of the system. A parameter according to one system can be taken into account as variable in another system. Besides, the relation between a system and its parameters can be reciprocal, i.e., the system itself can affect the value of the parameter.[21] For example, according to standard model of a demand market, an equation of demand (i.e., structure description of the system) and the differential equation of it (i.e., process description of the system) compose a system description. The system includes quantity and price as directly relevant factors for the system, some of the other factors (tastes and fashions, incomes, price changes in complementary goods etc.) are included in parameters as indirectly relevant factors, and the rest (e.g. supplied quantity) is regarded as irrelevant. Since the equation of demand shows the relations between its components (i.e., the amount of the commodity which consumers wish to/able to buy at any value of price) it is the structure description part of the system description. The differential equation of it shows the effect of changes in prices on demanded quantity; it is the process description component of the system description.

As implied in the example above, by 'structure' in a system, I mean the causal relations between its variables, or elements. A description of states of the system would usually illustrate its structure. However, the notion of 'state description' of a system due to Simon might not be equal to description of a system's structure.

[19] van Gelder , p. 616.

[20] Ibid, p. 617.

[21] Ibid, p. 617.

Following examples of Simon for state descriptions, a blueprint of a system can be regarded as illustrating the structure of the system. It is because a blueprint of a machine, for example, usually shows the relations between its components, includes the knowledge about phases through which the machine is constructed (i.e., possible states of the system through its construction process). A picture is also a state description which shows the state of a system at a time. However, a picture usually does not illustrate the causal relations between the elements of the system which is pictured, and hence a picture does not show the structure of the system.[22] Thus, I claim that the description of possible states of a system which also includes clues about the change of the system is the description of the structure of that system – i.e., the description of the inter- and intra-relations between its variables and parameters.

By 'change' in a system, I mean the behavior of the system. A description of the behavior of a system would illustrate its processes. Thus, process descriptions which "… provide the means for producing or generating objects having the desired characteristics"[23] are descriptions of change in the system.

I will give an example from economics, again, to illustrate the distinction between the structure and change of a system. The classical equation system of supply, demand and equilibrium condition equations in a market illustrates the structure of that market according to different variables and parameters.[24] Similarly, the equation of supply illustrates the structure of the relation between supplied quantity (the amount of commodity) and different values of prices. For example the slope of the supply curve which is the graphical illustration of the equation of supply shows the sensitivity of the quantity to different values of prices: the steeper the slope of the

[22] Simon, p. 479.

[23] Ibid, p. 479.

[24] This illustration of a market is controversial. Since the discussion is irrelevant here, I will not mention the debate on idealization of markets. I think this standard version of a market description will suffice as an example of structure and process distinction.

curve the more the quantity is affected by prices. The differential equations which are used for a static or dynamic analysis of different equilibrium points in that market show the process in the market. In the case of static analysis the equilibrium price and quantity will take values which depend only on parameters of the system. In the case of dynamic analysis, on the other hand, the equilibrium value of price and quantity will be equations which still depend on parameters but also on time. Both equilibrium equations show the equilibrium values of quantity and price according to change of parameters of the system, and hence according to structure changes of the market. They show the nature of the difference between different equilibrium points of the system. Thus, they are process descriptions of the system.

Neither inclusion of time as independent variable in a description suffices for this description to be a process description, nor does lack of the time variable in a description mean that it is a structure description. A process description shows the difference between different states of the system according to change of the variables in the system, and/or it shows the nature of the change in that system.

a. Componential[25] and Integrated Systems

The structure of a system can be described in a componential or integrated fashion. It depends on the target phenomenon and its conception. Describing the distinction of componential *vs*. integrated systems (C/I) by using the notion of 'decomposability' (i.e., divisibility of the system into functional subsystems) would be useful.[26] The notion of decomposability enables us to decide whether a system is integrated or componential. In the C/I distinction, a system's integrity or componentiality depends on their degree of decomposability. The more a system is decomposable the less it is integrated and vice versa. The extreme form of a decomposable system is simply

[25] I refer to 'component systems' in Bechtel & Richardson by 'componential systems'.

[26] Bechtel & Richardson, pp. 23-27.

decomposable, or aggregative, system. Wimsatt lists the following properties of aggregative systems:

1. Intersubstitutability of parts;

2. Qualitative similarity with a change in the number of parts;

3. Stability under reaggregation of parts; and

4. Minimal interactions among parts.[27]

A system approaches aggregativity to the extent that it has these properties. "In the simplest departures from aggregativity, we may still maintain intersubstitutability; however, when this also fails, we have what we call *composite systems*."[28] According to one of the two types of composite systems, namely componential systems, parts of the system interact with each other. Though, parts or components of the system are taken into account as autonomous elements with intrinsic functions. Componential systems show higher degree of decomposability than that of integrated systems; they are nearly decomposable. The structure of social systems can be the most obvious example of such systems.[29] For example, if we regard a family as a system, or a

---

[27] Wimsatt, "Foms of Aggregativity", pp. 260-268, quoted in Bechtel & Richardson, p. 25. According to Wimsatt these are the properties of aggregate systems for which Bechtel and Richardson use the term 'simply decomposable systems'. I will use the terms 'simply decomposable' and 'aggregate' interchangeably to indicate extremely decomposable systems. There is also a discussion of 'localization' on page 24. Localization is the identification of components in the system with their functions. There can be simple/direct, or complex/indirect localizations of functions in the systems. In the simplest form we identify a function to one component only. This can be done in simply decomposable systems. However, by the decrease in decomposability we start to attribute functions not to one component only; rather, to several components and their interactions. None of the theories under consideration in this thesis (namely, CTM, PDP and DST) proposes a simply decomposable system. Prima facie CTM might seem to attribute a function to a component (for example language acquisition to language faculty), the functions attributed by CTM performed by interaction between different components (for example, interactions with different faculties via interfaces). Even if the interaction is minimal, the function is not performed by the independent performance of a single component. Thus, I will not discuss localization problem further.

[28] Bechtel & Richardson, p. 26.

[29] Simon, p. 469.

subsystem in a social system, a family is a componential system: We cannot talk about a family which is composed of three and a half people, all members in a family (e.g., mother, father and children) have functions which can clearly identifiable etc.

According to the other type of composite systems, namely integrated systems, components of the system are interdependent and the organization of the system can determine the functions of the components: "There may be, for example, mutual correction among subsystems, or feedback relations that are integral to constituent functioning."[30] Elements of integrated systems are simple in the sense that they do not have intrinsic functions and autonomous mechanisms, and are interdependent such that there are "… organizational properties that fix the interaction of the parts and determine their significance for system behavior."[31] Cellular structure of multicellular organisms is a good example for such a system: Although they have components (viz., cells) functions of the components are determined by the organization of them, i.e., "… we cannot … understand how they function if we neglect their incorporation in, and integration into, the complex activities of the [organism]."[32]

b. Discrete and Continuous Change

Discreteness *vs.* continuity (D/C) is about the process description of the system. In the case of process descriptions we investigate *how* the system changes when its variables change, or *how* its elements are combined to derive the structure of the system, or what the *difference* between different states of the system is. In the case of structure descriptions, on the other hand, we describe the relations between the

[30] Bechtel & Richardson, p. 26.

[31] Ibid, p. 25.

[32] Ibid, p. 26. The original example from Bechtel and Richardson is mitochondria and its relation to cells. I derived and changed the example depending on the quotation from Levins, R. "Complex Systems" on the same page and the quotation from Grobstein, C. "Hierarchical Order and Neogenesis" on page 31.

elements of the system. By structure description of a system, we can see *what* value the dependent variables can take when independent variables or even parameters change. However, we cannot see *how* the change of the overall system takes place, or what the *difference* between different states of the system is according to different values of its variables; these are illustrated in process description of the system. Processes description of a system can be discrete or continuous. If the relevant change of the variables of the system, i.e., their detectable and/or relevant contribution to the system, is incalculably minute, the process description of the system is continuous. Even if the change of the overall system is discrete, the continuous change of its variables results in continuous process descriptions. If the change of the variables of the system takes places discretely in time, or intermediate states are irrelevant to the explanation of the process, the process of the system is described discretely. Typically, in mathematics, we describe continuous change by differential equations and discrete change by difference equations. However, mathematical equations are not the only way to describe change in the system.[33]

Componential systems have autonomous elements which have a specific, identifiable contribution to the system. Usually, their relevant change, or their contribution to the system is not incalculably minute and hence, the change in a componential system is not described continuously. As in the case of a family, we do not see the change in a family composition in a continuous manner. For example, in a family composed of mother, father and two children when a child gets married, or one of the parents dies the difference between two states, in other words the change in the system, will be a discrete change. Since components of a family cannot affect the composition of the family in an incalculably minute way, we cannot show the change in the composition of the family due to incalculably minute changes in the components of the system. However, when we try to understand their relations concerning their emotions, the

[33] ‘‘Merge’ which will be introduced later is also a discrete description of the change in the target systems.

system would show continuous change, though it is still componential. The system would stay componential because the effects from its components to the system would be still identifiable and specific to that component. Though, components of the system, i.e., members of the family, can show continuous change in their emotions toward each other which would affect the overall system. For example, loss of respect to father in the family can be detected in a continuous manner and it can affect the relations in the family continuously.[34]

Although an integrated system is usually conceived as continuous, change in an integrated system can be discrete when important changes in the variables of the system occur discretely in time. Discreteness is about the change in a system, or process description of the system – ignoring the intermediate states in time or between states. Besides, the nature of the phenomenon under consideration might urge theorist to describe processes discretely. For example, profit maximization process of a firm which sells automobiles will be described discretely while that of a firm which sells water will be described continuously.[35]

Another way of describing an integrated system with discrete change is the choice of time scale. The processes of an integrated structure can be formulated discretely or continuously depending on the conception of time. However, the choice of the time scale depends on the phenomenon under consideration and is not an arbitrary choice. For example, long-term development processes take place in a larger time scale but short-term ones take place in shorter time scales.[36] The relevant change in the system can take place discretely in time, regardless of the degree of the system's

[34] John Bolender, personal communication.

[35] For details of the example, see next section.

[36] According to both time scales the development process can be described as continuous or discrete; for example, van Geert proposes a continuous process description, even for long term development. However, I assume that the example of different time scales for short- and long-term development processes is still a good example for the dependence of chosen time scale on the phenomenon under consideration.

decomposability. A system can be an integrated system, but the evolution of the system can take place discretely in time.

Table-1:[37] Illustration of the distinction of C/I-D/C.

Structure (C/I distinction)	Componential		Integrated	
Change (D/C distinction)	Discrete	Continuous	Discrete	Continuous

c. Two conceptions of cognition

I will propose two conceptions of cognition in terms of structure descriptions. I have chosen to distinguish conceptions of cognition with respect to structure descriptions because nature of the structure determines the nature of the process in a system. The reason and the nature of continuity or discreteness of the processes in a componential system would be different from those of an integrated system due to their structures. CTM proposes a componential system, and PDP and DST propose an integrated system. Their system descriptions and kinds of these descriptions will be explained in the fourth chapter.

[37] Ayhan Sol, personal communication. He suggested me to clarify the distinction of C/I-D/C by illustrating them in this table.

A componential system might show discrete change because the items upon which the system works are discrete. For example, discreteness of lexical items results in discrete processes in language faculty.[38] Another reason for discreteness of the processes in a componential system might be the idealization of time as sequences. In this case the system will be changed step by step and the process description of the system will not include intermediate states of the system. The most obvious example of this kind of process description in cognitive science is the instantaneous model of language acquisition due to Chomsky.[39] On the other hand, a componential system might show continuous change due to contribution of its one or more components to the change in the system continuously. It is possible for this kind of componential systems to show both discrete and continuous change due to operations of different components.

> On my interpretation, the mental processes underlying [Relational Models Theory] result from a facultative system, given that there are three distinct mental powers in play: The Social Pattern Generator (a kind of lexicon), a computational component consisting of Copy and Merge, and the preonic systems. The system is componential because of these three faculties. Some change in the system is discrete, given that the SPG and Merge operate in a discrete fashion, but some change can also be continuous insofar as the preonic system can behave continuously. It is perfectly possible that there are continuous changes in the preonic systems. For example, the emotional intensity of a relational model may gradually change over time. The fact that the entire system consists of components does not negate this fact.[40]

Process description of an integrated system can also be continuous or discrete. Chosen time scale might be one reason. Processes at larger time scales usually described discretely, even though it is not a necessity. Discreteness in long-term development processes and continuity in short-term development processes can be

---

[38] John Bolender, personal communication. (For details, see chapter 4.)

[39] Annette Hohenberger, personal communication.

[40] John Bolender, personal communication. (For details, see chapter 4)

given for this kind of reason. Although it is not necessary to describe long-term development discretely, explaining long-term development at a larger time scale is not an arbitrary choice. In explaining processes at larger time scales theorists usually ignore intermediate states and investigate the changes discretely, for example, year-by-year.

Another reason of discreteness in an integrated system might be the nature of the phenomenon under consideration. For example, in the cost-revenue analysis for a firm, the equations of cost and revenue exhibit integrated structure. Their components' (i.e., price and quantity) contribution to the system is determined by the organization in the system. Total revenue, for example, will be equal to the multiplication of the quantity (amount of the commodity) sold and the price of the commodity. Since in a competitive market individuals cannot affect the price, a firm can increase its revenue and/or maximize its profit only due to decisions of the quantity which will be supplied. Increase in the quantity will cause increase in the revenue. However, that will also cause increase in the cost, and hence decrease in the profit. Thus, firm should decide the optimum value of the quantity to supply. The tools for this optimization problem are difference or differential equations.[41] The decision of the mathematical tool depends on the nature of the change in the system which is dependent on the nature of the commodity. If the commodity supplied is not divisible, for example automobile, analyst would use difference equation which is an expression of discrete change. If the commodity supplied is divisible, as water[42], our mathematical tool for the analysis can be differential equations which presuppose continuous change.

[41] The analysis described here depends on the standard marginalist analysis in which the quantity that maximizes the profit of the firm is determined by equating the marginal amounts of revenue and cost. 'Marginal amount' means the amount of the contribution of the last unit to the system. Although this is a controversial view in economics, I will not mention the debates about the issue.

[42] We can regard water as a commodity as if it is not divisible. For example we can choose one kg. as the indivisible unit for our analysis and use difference equations as our tools. However, this would be an arbitrary choice because we can sell, for example, 1,25 kg. of water. By contrast, indivisibility of automobiles is not an arbitrary choice; we cannot sell one and a half automobile.

Components of a componential system can have integrated inner structure and an integrated system can be a component in a componential system at higher-levels. An integrated system can *create* a componential system when it *seems* componential at higher-levels. Although several levels will be jumped, the example family mentioned before can be an obvious illustration for this kind of part-whole relation. All members of the family are multicellular organisms which have integrated structures but the members are components in a higher-level componential system; namely, the family.

None of these conceptions seem to be enough to understand cognitive phenomena, without contribution of the other. We seem to need different conceptions proposed by different theories to work together. In the componential conception of cognition we need a different theory in order to understand the underlying continuous processes and integrated aspects of the cognition, or we can simply ignore those processes as *irrelevant*. On the other hand, the possibility of formulating discrete mechanisms of cognition in the integrated conception does not mean that we have an account of higher-level componential structures only due to discreteness of our theory. In my opinion, the full account of cognition includes both of these conceptions. Different theories which have one of these conceptions contributes to our understanding of cognition, just like the contribution of anatomical, genetic, physiological etc. explanations to our understanding of organisms.

CHAPTER 3:

IDEALIZATION

As explained in the previous chapter, models and theories may propose different system descriptions. These different descriptions assume different conceptions of the phenomenon under investigation. The proposed conception mostly depends on the idealization at work. As it is clear because of our limited cognitive capacities, we need to idealize the target phenomena while theorizing them. Although it need not be the justification of the idealization, too, the limited capacity of human minds is the reason of the need of idealization. We need to represent reality partially and sometimes with false assumptions to understand the phenomena under consideration and in some cases even to recognize them in the first place. We idealize the reality while theorizing which is crucial to bear in mind; not to be like these blind men:

> Suppose that the five blind men of the legend, perceiving different aspects of the elephant, nonetheless recognize their common referent. “Good for them!” you might say. But not so fast: given the tremendous difficulties of reconciling their views of it, they nonetheless decide to treat their views as if they were of *different* objects![43]

Idealization is an intentional distortion of the reality, thus, it necessitates justification. However, the justification of idealization need not be the same as the reason of it, i.e., the limited capacity of human minds. This kind of justification implies inclusion of more causal factors in a model by advance in technical and mathematical tools. A theorist might represent only the core causal factors of a phenomenon and this kind of representation might be the final goal of theorizing. Thus, we can choose different kinds of idealizations with different justifications. It depends on our purposes for constructing a theory and the nature of the phenomena. According to different kinds of idealization which will be explained in a while, the

[43] Wimsatt, p. 180.

distortion of the reality has different justifications. The kinds of idealizations, their justification, and relation to theorist's purposes will be explained in the first section. In the second and third sections the nature of complex phenomena and their relation to the kinds of idealizations will be investigated.

In this chapter, I will try to clarify my claim that cognitive science show a progress from the conception of cognition as levels of organization to conception of cognition as levels of explanation. According to former conception, cognition is a higher-level phenomenon and the lower-level phenomena which are the processes by which cognition emerges are irrelevant to our study of cognition. In this conception the lower-level processes are the processes of the physical substance which realizes the cognition, i.e., brain and the central nervous system. Since the cognitive phenomena are multiply realizable[44] by different physical substances, the way of understanding the nature of cognition should be investigating those multiply realizable properties, not the diverse physical processes. According to latter conception, on the other hand, both cognition and the underlying processes contribute to our understanding of cognition. In other words, cognition has two important aspects: one aspect is emergent properties of cognition and the other is emerging processes of it; both

[44] The 'multiple realizability thesis' is supposedly a supporting thesis for irreducibility and autonomy of higher-level sciences, especially psychology. The argument goes like this: There are higher-level, autonomous sciences which deal with functional relations. Those functional relations demand functional explanations which can be realized by very diverse physical substance. The attempt to identify terms of higher-level sciences to that of lower-level sciences would fail because of the one-to-many relation between those terms. Besides, the theory obtained after reduction of higher-level sciences would include irrelevant features and dismiss important functional features in the explanation of those higher-level phenomena. Thus, multiple realizability of higher-level explanations results in irreducibility and autonomy of higher-level sciences. Although I will not discuss the multiple realizability thesis and reduction debates further, just to state my opinion, I wish to say that Wimsatt is right in mentioning the same feature, even in physical sciences – e.g. one-to-many relation between terms of physical sciences which deal with different levels of organization. We cannot say that for example, physical and chemical explanations do not contribute each other, and hence chemical explanations are autonomous explanations. This fact weakens the part of the argument about the autonomy of higher-level sciences, even if it does not support the idea of reducing higher-level sciences to lower-level ones. The part of the discussion which is related to the claims of Wimsatt about ontological categories of complex phenomena will be clearer in the second and the third sections, though discussions about reduction will not take place, explicitly.

deserving different kinds of explanations. In this second conception the explanations of those processes which *produce* cognition are not exactly the explanations of the physical substance; rather, they are also multiply realizable functional explanations of emergence processes. Thus, different theories of aspects of cognition have partially overlapping different targets and hence, propose different conceptions of cognition. My last claim is about the possible future of cognitive science in which, I assume, *syntheses* will have special importance. This last conception will lead us to propose unified theory, rather than proposing different theories with partially overlapping targets for different aspects of cognition. In the process through which we will obtain a unified theoretical perspective, hybrid theories will have special importance in advancing technical, mathematical and theoretical tools. Although, recently proposed hybrid theories of cognition seem to support my last claim, it is only a speculative claim, for now, which necessitates further advances in mentioned tools.

My claim depends on Galilean idealization, in spirit, for the construction of theoretical perspective. However, I also claim that the other two kinds of idealizations, namely, minimalist idealization and multiple models idealization, have roles in this progress and they are complementary, rather than competitive. The former has role especially for recognizing and refining the phenomenon and the latter has role especially in the model construction process for complex phenomena, even from a unified perspective. For the sake of minimalism ignoring the contribution of advances in techniques and mathematical tools to our theoretical perspective, for the sake of completeness ignoring the improving function of simplicity, or for the sake of constructing compatible models ignoring the complexity do not seem plausible. I claim that we may need all of the three kinds of idealizations in different phases and areas of our inquiry in cognitive science because of the complexity of cognition. Before stating their relation in detail, I should explain the kinds of idealizations.

a. Kinds of idealization

Weisberg offers three kinds of idealization with several representational ideals. He defines 'representational ideals' as follows:

> Representational ideals are the goals governing the construction, analysis and evaluation of theoretical models. They regulate which factors are to be included in models, set up the standards theorists use to evaluate their models, and guide the direction of theoretical inquiry.[45]

For example, representational ideal of Galilean idealization is completeness. It is the goal of the theorist to represent reality as complete as possible. Representational ideals have two components: inclusion rules and fidelity rules. Inclusion rules enable the theorist to identify the relevant components in the theory and fidelity rules enable the theorist to evaluate the accuracy and precision of the theory.[46] According to the representational ideal which is mentioned above, i.e., completeness, the theory should include all factors which are related to the phenomenon under consideration (inclusion rule). The best model would be the model which satisfies the inclusion rule with a high degree of precision and accuracy (fidelity rule) – the model which represents all causes including the external ones. As mentioned above, every kind of idealization has a different justification for their distortion of the reality. I will try to explain several kinds of representational ideals while explaining kinds of idealizations and their justifications.

1. Galilean idealization

Justification of Galilean idealization is pragmatic. The goal of the theorist is to represent reality as realistic as possible and idealizes it mostly because of the lack of

---

[45] Weisberg (2007b), p. 648.

[46] Ibid, pp. 648-649.

proper mathematical and technical tools. Thus, as mathematical and technical tools advance, the theorist de-idealizes the target phenomenon in a more realistic fashion.[47]

The main representational ideal of Galilean idealization is 'completeness', i.e., complete representation of the phenomena with the most precision and accuracy possible.[48] In order to achieve this maximum precision and accuracy, theorist should include all causal factors which are related to a phenomenon. The theorist who accepts Galilean idealization wishes to construct a theory which is a complete representation of reality as complete as possible with maximum precision and accuracy. Thus, the theory should include all properties of the phenomena and external causes of those properties and "… the best model is one that represents every aspect of the target system and its exogenous causes …"[49] However, the final goal is not achieving the maximum accuracy and precision in the output of the model; rather, complete representation of the reality.[50]

2. Minimalist idealization

According to minimalist idealization a model should include only the relevant causal factors, "… the causal factors which give rise to a phenomenon."[51] It is not a phase in scientific inquiry to idealize reality minimally which can be accepted until some advanced techniques invented; it is the final goal to represent reality with only the relevant causal factors. The representational ideal of minimalist idealization urges the theorist to include only the core causal factors in explanation to be able to represent the causal structure of the phenomenon. As implicit in the previous sentence, only the primary causal factors which are responsible for the occurrence of the

---

[47] Weisberg (2007b), pp. 640-642.

[48] Ibid, p. 655.

[49] Ibid, p. 649.

[50] Ibid, pp. 652-653.

[51] Ibid, p. 642.

phenomena should be included in the model.[52] The claim is that the models generated by this kind of idealization represent the relevant causal structure of the phenomenon under consideration.

Because of the demand for one model for target phenomena in minimalist idealization and Galilean idealization, for them, 'generality' is also important in construction of models. According to this representational ideal, i.e., generality, a single model should capture as much target as possible.[53]

3. Multiple-models idealization

"Multiple-models idealization … is the practice of building multiple related but incompatible models, each of which makes distinct claims about the nature and causal structure giving rise to a phenomenon."[54]

Justification of multiple-models idealization is complexity of the phenomenon, limited nature of human cognition, and tradeoffs between different goals for representations like generality, simplicity, precision and accuracy.[55] Accordingly, representational ideals of multiple-models idealizations can be different depending on the goals of the modelers in their representations. According to this kind of idealization a model does not give a full representation of reality, even as a causal structure, in a single model; rather, the model represents different aspects of the phenomenon with different purposes. For example, a model can aim to give accurate predictions while another represents the core causal structure. Similarly, the underlying processes and the outcome of those processes might demand different, incompatible but complementary models.

---

[52] Ibid, p. 651.

[53] Ibid, p. 655.

[54] Ibid, p. 645.

[55] Ibid, p. 646.

I agree with the idea that Galilean idealization is not the only legitimate idealization kind and minimalist idealization can in some cases enable us to recognize an interesting target phenomenon.[56] It is really an efficient strategy to recognize, identify and refine the target phenomenon and in some cases, to explain it. For example, recognizing the spiral shape of galaxies might be possible only by some unrealistic assumptions. We should abstract out the performance restrictions to be able to obtain this explanandum. Despite the fact that a galaxy cannot be shrunk or dilated, or it cannot be rotated, literally, if it could extend infinitely, we would see that it would look the same. Ignoring these physical limitations reveals an explanandum, i.e., the dilational and rotational symmetry of the galaxy. "It then becomes a question for astrodynamics how to explain this symmetry."[57] However, in some other cases, explanation of that phenomenon might necessitate more causal factors to be taken into account than the factors we have included at the beginning, even if we still wish to idealize the phenomenon minimally. We need to take more causal factors into account in explaining the emergence of the spiral shape of the galaxy in our example – causal factors like gravitational force, or density of the matter at the galaxies. On the other hand, when it comes to explanation of emergence processes a mixture of minimalist idealization and Galilean idealization seems more proper, i.e., addition of new relevant causal factors – only the relevant ones without all details – in our explanation by progression of techniques and tools to be able to represent the core causal structure of the reality as realistic as possible.[58] While new relevant causal factors can be taken into account by new tools and they can improve our understanding, resisting to them for the sake of minimalism does not seem plausible.

---

[56] Weisberg (2007b) and Bolender (2010), pp. 41-42.

[57] Bolender (2010), p. 42.

[58] Although identifying the *relevant* causal factors is another problem, I will just assume that we are able to identify them and not discuss the problem.

For example, two of the most controversial points in cognitive science are the relevance of the environment and time. Taking time and environment into account might not be only due to endorsement of Galilean idealization as the only legitimate idealization type but also because of their necessity in understanding the phenomenon, even in a minimalist idealization. One important property of some complex systems is their emergence from very simple rules, especially when they are emerged from nonlinear interactions between its components.[59] Notions of self-organization and emergence seem to provide fairly minimalistic explanations of the emergence of cognitive processes without dismissing time constraints and environment. "Ironically, then, complex systems can sometimes be described by fewer variables than can relatively simple systems."[60]

On the other hand, I claim that different aspects of cognition might necessitate different conceptions of cognition which are idealizations, indeed. These different conceptions have different assumptions about the core causal factors which give rise to the phenomenon. For example, *emergent* properties of cognition can be idealized as a system which has componential structure with discrete change, while *emergence* of it might urge the theorist to suppose a completely different system which is integrated and has continuous or discrete processes. Furthermore, the only reason of the usefulness of the cooperation between different conceptions of cognition need not be the fact that emergence processes and emergent properties usually demand different conceptions in their explanations. Different sub-targets of a target phenomenon might demand assuming different conceptions of cognition, too. Thus, we might need different systems with different structure and process descriptions for different aspects of the same *thing*. Accordingly, I suppose that we need to construct different models from different perspectives for different phenomena to be able to understand cognition which is highly complex. But, what is complexity? I will

[59] Nowak, p. 185.

[60] Ibid, p. 185-186.

discuss complexity, the ontological categories of it and the theories related to those ontological categories in the next two sections.

b. Complexity

It is clear that cognition is a highly complex system. Explaining complex systems is a difficult task and demand different treatment than explaining relatively simple systems. Before attempting to explore possible strategies in explaining complex phenomena, describing complexity would be proper.

The main property of complex systems is their hierarchical architecture. They are composed of parts in a non-aggregative way that "… the whole is more than the sum of the parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole."[61] For example, the markets in economics are composed of individual consumers and firms, but not as the sum of its components. The whole markets show properties and processes specific to them due to the inter- and intra-relations between individual consumers and individual firms; to the extent that those specific properties and processes of the markets are subject matter of a discipline, namely, economics. The relation of a whole and its parts as explained by Simon is obvious in competitive markets in which any individual cannot affect the processes in the market; rather the properties and processes of the market can only be obtained through interactions between individuals.

Wimsatt offers two kinds of complexity: descriptive complexity and interactional complexity. According to him a system is complex not only because of the part-whole relation between its components. There are different ways to regard systems as complex. Different individuations or spatial decompositions, of the same system can have common parts. If these common parts show almost one-to-one mappings (for

[61] Simon, p. 468.

example, different decomposition of a piece of granite under chemical composition and density), the system is a descriptively simple system.[62] If different individuations of the system overlap in arbitrary ways and parts do not show one-to-one mappings (for example, different decompositions of an organism under anatomical, physiological and biochemical perspectives), the system is a descriptively complex system.[63] Interactional complexity shows the complexity of causal interactions of systems under different decompositions. Interactionally simple systems can be decomposed into sub-systems in which intra-systemic causal relations are stronger than extra-systemic ones. In interactionally complex systems, on the other hand, causal relations cross boundaries between different spatial decompositions.[64] Interactionally simple systems which are described from different perspectives have problems that belong to one perspective and those problems are solvable by the tools and information from that perspective. However, interactionally complex systems have trans-perspectival problems and "… require the use of information from more than one perspective for their solution."[65]

Cognition is both descriptively and interactionally complex system because of its trans-perspectival problems and those problems' trans-perspectival solutions. Of course, purpose of theorizing determines the problems of the theory and thus complexity.[66] Decomposition and interactions between sub-systems in a complex system can be proposed as if simple in a model depending on the purpose of the model. However, if a system which is described as if it is simple under different perspectives urges the theorists to consider trans-perspectival problems, we can conclude that the system is complex, independent of purposes.

[62] Wimsatt, p. 182.

[63] Ibid, p. 182, 229.

[64] Ibid, p. 184.

[65] Ibid, pp. 228 (emphasis deleted).

[66] Ibid, p. 228.

Wimsatt derives his ontology from the practice of scientific inquiry. He defines the "concept of an object" as "... a concept of something that is knowable robustly."[67] For example, according to his definition of concept of an object a sphere made up of glass is a robust object, knowable through at least two means of our sensory access, i.e., we can touch it and look at it. The properties (e.g., roundness, transparency, hardness) of the glass sphere are detectable through our different means of sensory access. These different means of sensory access illustrates different perspectives. The glass sphere will seem different under different perspectives. Besides, it will have different properties and different problems related to those properties according to the way we access the object. In addition to his definition of an object, he proposes higher-level ontological categories which show the degree of complexity of the *object*. He claims that 'levels of organization', '*perspectives*'[68] and 'causal thickets' are "... higher-level ontological features, *Organizational Baupläne*, related to things that people usually talk about under the topic of ontology (things like objects, properties, events, capacities, and propensities) as paragraphs are to words and phonemes or morphemes."[69]

The explanations at different levels are explanations from different perspectives. A system can be decomposed as sub-systems with respect to interactions between its parts under different perspectives. Interactional complexity shows the degree of complexity of the system which is decomposed such a way. "Simple systems as well as complex systems can be described from a variety of perspectives, but will differ in

[67] Ibid, p. 197. There is detailed discussion on 'robustness' available in the fourth chapter, namely, "Robustness, Reliability, and Overdetermination" (Ibid, pp. 43-74), and throughout the book. However, the definition of the notion, viz., 'robustly knowable', suffices for the discussion which takes place here. The definition is this: An *object* is robustly knowable, or a robust object, if it is knowable through different means of access.

[68] To avoid confusion between 'perspectives' as an ontological category and 'perspective' in the ordinary usage of the word, I will use italicized version of the word to indicate *perspectives* as one of these ontological categories.

[69] Wimsatt, p. 194.

the degree to which they have problems that are trans-perspectival …"[70] To the extent that the system is simple, we will need less perspectives to be able to explain that system. In the simplest cases one perspective will be enough to our explanations. Interactionally complex systems urge the theorists to consider trans-perspectival problems. Trans-perspectival problems also vary according to their complexity depending on to the complexity of decompositions under relevant perspectives. Different decompositions according to different perspectives can be spatially coincident, can show part-whole relation hierarchically, or overlap in arbitrary ways.[71] The last two cases illustrate descriptively complex systems.

The property of part-whole relation of parts of a system is the property of 'levels of organization' which is one of the ontological categories proposed by Wimsatt.[72] There are different organizational properties at different levels of organizations and there are inter- and intra-level interactions among parts of such a complex system.[73] Part of a system in a higher-level organization can be regarded as whole in a lower-level organization which in turn has parts. Since intra-relations between parts at one level would be relatively stronger than inter-relations between levels,[74] there is a one-

---

[70] Ibid, p. 228.

[71] Ibid, pp. 228-229.

[72] Ibid, p. 201. However, there is a sense of lower- and higher-level distinction other than that of Wimsatt. Sub-symbolic and symbolic levels are regarded as lower- and higher-levels respectively (Abrahamsen & Bechtel, p. 163). Levels according to levels of explanation are not the levels in levels of organization which show part-whole relation. Firstly, according to levels of explanation, levels are not ontological categories, although they have ontological implications. Besides, even if they show part-whole relation, levels of explanation would not be the synonym of levels of organization, but theories of levels which deal with the phenomena which belong to levels of organization category. However, levels of explanation do not show part-whole relation. Levels of explanation fit to perspectival theories which imply *perspectives* category as ontological category of its phenomena, rather than theories of levels. 'Perspectival theories', 'theories of levels' and 'syntheses' will be clearer in a while.

[73] Wimsatt, pp. 182-186.

[74] Although intra-relations between parts at one level will be stronger than inter-relations between levels, there is still inter-relation between levels, at least as part-whole relation. One level is not exhaustively enough to understand the phenomenon which belongs to levels of organization category because of the trans-perspectival problems related to the phenomenon.

to-many mapping between parts of higher- and lower-levels, thus higher-levels are multiply realizable. These mappings can both be between parts of the system under different spatial decompositions and/or between parts of the system as sub-systems under different interactional decompositions. Since Wimsatt constructs his theory of levels of organization for mainly physical systems, he claims that multiple realizability is not a property which is specific for psychological processes.[75]

According to Wimsatt systems which belong to *perspectives* category ontologically do not hold part-whole relation between their parts under different decompositions at different levels, or the relation should be reciprocal, i.e., "… information from each perspective is relevant to the solution of at least some problems in the other."[76] The properties of *perspectives* are as follows:

> *Perspectives* … (a) [span] more than one level, and thus [cannot] be ordered as higher and lower or more primary and secondary than one another; (b) [give] criteria for decomposing systems into parts using the properties and tools appropriate to that perspective; (c) [are] manifestly incomplete descriptions of their objects; (d) [are] such that different perspectives (for complex systems) [can] cut up systems in quite different ways that [are] not easily comparable to one another; (e) [have] a class of problems that they could solve in isolation; and (f) (for complex systems) [have] other problems that could not be solved without bringing in the resources of another perspective or perspectives.[77]

Following the same example, we would find different, but sometimes overlapping, qualities of our glass sphere, via different means of access. When we look at the object we see that it is round and transparent; and when we touch it we sense that it is round, again, and hard. Thus we can conclude that it is a round object and continue refining its properties via different means of access. According to different senses it would seem different partially. For the sake of simplicity, suppose that these two

---

[75] Ibid, pp. 201-202.

[76] Ibid, pp. 233-234.

[77] Ibid, pp. 230-231. (My emphasis)

means of access are two different *theories*. These two theories would deal with the glass sphere which belongs to *perspectives* category. The object has different properties detectable through two different perspectives. According to these perspectives one quality, viz. roundness, is common. However two of them, viz., transparency and hardness, are different. Although different perspectives' proposed qualities partially overlap and they together constitute the conception of the object, we can still distinguish these perspectives' contribution to the conception of the object and method of the access to the object. On the other hand, its material, namely, glass, would belong to levels of organization category ontologically: roughly, molecules constitute the glass, atoms constitute the molecules of it, so on.

As the complexity of the system increases problems that occur in the system becomes multi-perspectival. Then it becomes difficult to identify which part of the problem belongs to which perspective. This breakdown of boundaries of perspectives leads *perspectives* to causal thickets where still different perspectives save their problems and methodologies partially. However, in the case of causal thickets, disputes about boundaries, methodologies and key terms appear. The solution should "… await the development of conceptual structures, methodologies, and new explanations of mechanisms in terms of them."[78] For example if the glass sphere mentioned as an example above, is a sphere of a wizard, it would belong to causal thickets category. What makes it a *magical* tool can be detected through history, theology, cultural studies, even may be sociology. However, our regarding it as a *magical sphere* depends on a conception which includes all these perspectives, though probably in an indistinguishable manner or distinguishing them would be controversial.

[78] Ibid, pp. 238-239.

c. Levels of organization, perspectives, and causal thickets

As I have implied in the example of glass sphere, the means of access to *objects* are proposed by our theories from different perspectives. Besides, the conceptions of *objects* are provided by these different perspectives. According to this classification of ontological categories our theories would imply one of these categories. Physics, for example, implies levels of organization for the ontological category of its *objects*. I wish to call these kinds of theories, theories of levels to prevent confusion between theories and the ontology implied by them. I will call perspectival theories for theories which deal with different aspects of an object and its properties which belongs the *perspectives* category ontologically. Finally, the theories which deal with the phenomena which belong to the most controversial category, causal thickets, demand a different name, I think, to avoid confusion. I will call them *syntheses*[79]. In these kinds of theories, there are contributions from different perspectives in an undistinguishable manner. For example, a theory which deals with the history of wizards' glass spheres would include contributions from different perspectives such as theology, cultural studies, sociology etc. It might be the case that occurrence of *syntheses* is a phase which is directed to a unified theory or the only way to deal with highly complex phenomena might be dealing with them by *syntheses*.

Prima facie, CTM, PDP and DST are theories about different levels of organization: CTM for higher-levels and the other two for lower-levels. Indeed, until recently, it is accepted as if these theories are theories of levels. According to this conception of cognition, there is cognition and there is brain which is its realizer. Cognition is a higher-level phenomenon and truly cognitive processes are explained functionally. The other theories deal with lower-level processes, processes belong to the physical realizer of cognition, and they are irrelevant for truly cognitive processes. Thus,

[79] For the same reason to use italicized 'perspectives', viz. to avoid confusion between 'syntheses' as the name of theories which imply causal thickets as ontological category and 'syntheses' in the ordinary usage of the word, I will use italicized version of the word to indicate *syntheses* as the theories which imply causal thickets.

lower-level theories are not theories about cognitive phenomena, but theories about their implementation; because cognitive processes are higher-level processes which are implemented in humans by brain and the central nervous system and can be implemented by various kinds of physical substance.[80] However, PDP and DST are not about the micro-structure of brain. They are both about the cognitive phenomena with mathematical models which are inspired by micro-processes – they are not *about* micro-processes, or explanation of them. For example PDP also propose functional explanations, like CTM, which are multiply realizable. It tries to "… understand how various computation problems can be solved …"[81] The difference between the models of PDP and CTM comes from the system which inspires them and the aspect of the cognition which they deal with. However, both of them have cognition as target phenomenon. One can claim that PDP attempts to explain higher-level processes by explaining lower-level processes and hence although its target is higher-level processes it is about lower-level processes and not multiply realizable. However, the studies of various areas which use neural networks as mathematical tools show that PDP uses a multiply realizable system.[82] Similarly, DST offers models which are clearly about higher-level processes.[83] This property of DST is more obvious than that of PDP because it is inspired from dynamical systems, but not from neurobiology directly like PDP. Indeed, inspiration from the supposed physical substance which is responsible for cognition, namely, neural processes, might be really confusing. I claim that these three theories are not different theories related to different levels of organization; they imply *perspectives* for the ontological category of their subject matter.

[80] Fodor & Pylyshyn.

[81] Smolensky et al. p. xi.

[82] Ibid, pp. 1-2.

[83] Townsend & Busemeyer, Editors' Introduction, p. 101.

If we accept that PDP and DST are about the emergence of cognition and CTM is about the emergent properties of cognition (especially for language acquisition), they seem to be perspectival theories. "… [I]nsofar as a theory deals with only a subset of the causally relevant properties of an object, it has a perspectival character, but if the properties it deals with are sufficiently robust and fruitful, it may be easy to forget this fact."[84] Since these theories use minimalist idealization and deal with different sub-targets of a target cognitive phenomenon, they deal with only the relevant causal factors according to their perspectives, thus they have perspectival characters. It is reasonable, then, that they regard different causal factors as relevant for their explanations. On the other hand, they seem to need each other to be able to give a complete picture of the cognition.[85] For example, the phenomenon of language acquisition includes language, pre-learning device and the process of learning in interaction with environment as sub-targets to be explained. We can see CTM, PDP and DST are different perspectives which explain each sub-target respectively with their overlapping domains. Although the part of my argument in which I matched sub-targets with theories might be controversial, I assume that the uncontroversial part of my argument is this: All three sub-targets must be explained to be able to understand language acquisition. I wish to clarify my argument in the next chapter.

On the other hand, models generated by an approach which is hybrid of these three approaches[86] offer another picture about their relation: The conception of cognition proposed by hybrid theories fall under causal thickets category and hybrid theories are *syntheses* in which identifying the boundaries of different perspectives is difficult.

[84] Wimsatt, p. 236.

[85] Abrahamsen & Bechtel, p. 167.

[86] For example, the model offered in Hohenberger&Peltzer-Karpf.

According to Wimsatt causal thickets might not be a kind of *waste basket category*.[87] However, I assume that it is almost equally possible that cognition as causal thickets might persist or it might turn to a unified conception. According to my observation, cognitive science is in a phase in which theories become *syntheses* from perspectival theories. However, I cannot say the same for multiple models idealization; we might need more than one model to understand cognitive phenomena and possibly, it is not a phase in the way to construct the most realistic theory with a single model. Thus, I assume that in the long-term, even if we will be able to construct a unified theoretical perspective for cognition, we might still need multiple, incompatible but related models.

The relations of these three approaches according to the discussions took place in this chapter and in the previous one will be clearer in the fourth chapter, Applications in Cognition, where I am planning to discuss these issues over the examples from language acquisition and social cognition.

[87] Wimsatt, p. 240.

CHAPTER 4:

CONTROVERSIES BETWEEN CTM, PDP AND DST

The relation between three main theories of cognition is discussed in Abrahamsen and Bechtel in detail over types of mechanisms required for providing explanation of different phenomena.[88] They defend a kind of pluralistic view and claim that the main theories of cognitive science and their models are complementary, rather than competing, in our understanding of cognition.

> We have … emphasized that cognitive science, despite its many disputes, has progressed by continually combining and recombining a variety of influences. The use of equations both in characterizing and explaining phenomena are among these. When combined with other influences and commitments, the outcomes … have ranged from information processing models with quantified operations to connectionist networks to both global and mechanistic dynamical accounts…. Cognitive science takes multiple shapes at a given time, and is protean across time.[89]

The discussion which will take place here is about several controversies between these approaches which are the relevance of time and environment, and innateness. I will also defend a pluralistic view similar to the view described above. However, I will focus on the controversies firstly over language acquisition models and secondly over social cognition models and try to explore the conception of cognition needed to explain the phenomena under consideration. I claim that both of the target phenomena (i.e., language acquisition and social cognition) have sub-targets to be explained and explanations of those sub-targets might demand different conceptions of cognition.

---

[88] Abrahamsen & Bechtel, p. 182.

[89] Ibid, p. 182.

According to innateness assumption there is a device which is responsible for the acquisition and knowledge of several cognitive capacities specific for humans. The most known capacity among other capacities which depends on an innate device is language, regarding both its acquisition and knowledge. Since among the organisms only humans are able to acquire and use a language, there must be *something special* for this capacity *in* human mind/brain. Although none of these theories ignores that there is something innate which enables language acquisition, CTM, PDP and DST attribute different roles and mechanisms for this pre-learning device. Thus the role and the working principles of this device show variety among these approaches to the extent that they seem to talk about three different *things*. I claim that they investigate different aspects of the same device; namely, its content, mechanism and relation to the environment.

According to CTM language acquisition is only maturation of this device, namely, language faculty. The theory will be explained in the next chapter. However, to mention roughly, according to CTM the conception of language acquisition is a switch in the language faculty from Universal Grammar (UG) to a particular language. UG represents the core formal properties of all particular languages. In other words, it is *raw material* of all possible languages which is genetically endowed and present at birth. It switches to a particular language by the linguistic data obtained in early childhood. Thus, in the process, time is considered as a step and environment is considered as a trigger. Critics of CTM mostly insist on the importance of time and environment and criticize CTM for ignoring them. For example, a criticism from the DST perspective for the conception of language acquisition according to Chomsky's "Instantaneous model of language acquisition" is as follows:

> The "Instantaneous model of language acquisition" (Chomsky 1975, 1986) deliberately abstracts away from any temporal aspects of the real acquisition process. If one thinks of the initial state S_0 of the child as being a pure reflection of Universal Grammar (UG) and the final steady state S_s

> as representing full-fledged knowledge of language, then the instantaneous model of language acquisition holds that S_0 is mapped onto S_s in a single moment, "as if it were instantaneous" (Chomsky 1986, 52). This mapping simply applies the principles of UG to the entirety of the primary linguistic data, disregarding any temporal characteristics, e.g., order of presentation of the data or exposure time. As these temporal characteristics are considered irrelevant to the ultimate outcome of the acquisition process, namely knowledge of language, the idealization is legitimate, in Chomsky's view. He stresses, however, that this is an "empirical hypothesis" (ibid.), hence debatable on the grounds of empirical evidence. Indeed, this idealization has engendered a lot of criticism from various theoretical camps. Dynamic systems proponents are particularly displeased in this respect since they consider time as the most important aspect of cognitive processes (van Gelder 1998).[90]

Prima facie, the debate is about the 'competence' and 'performance' distinction. 'Competence' is the knowledge of language and 'performance' is the usage of it.[91] According to this distinction the core causal mechanism which is responsible both for competence and performance is explained by the mechanism of competence. It is a minimalist idealization in which only the relevant factors are included. Time constraint and relation of organism to environment are not relevant. However, given the fact that both DST and PDP intend to provide universal principles of acquisition process, I do not think that the difference between these theories is only due to competence/performance distinction. In my opinion, just like the idealization of UG from individual languages, PDP and DST idealize mechanisms from individual performances, indeed, in a fairly minimalistic way. For example, in a model of language acquisition constructed by neural networks, or in the model of learning coordinated behavior provided by Kelso, hormonal processes or anatomical differences which might also affect the performance indirectly are not taken into

[90] Anette Hohenberger, personal communication (Chomsky, N., *Syntactic structures*, The Hague: Mouton, 1975 and Chomsky, N., *Knowledge of language: its nature, origin, and use*, New York: Praeger, 1986).

[91] Chomsky (1995), p. 14.

account.[92] Besides, these models do not imply that those factors which affect language acquisition indirectly will be taken into account by improvement in technical and theoretical tools. Thus, I do not think that their difference comes from the kind of idealization at work, rather from the difference of their targets. Sub-targets in understanding language acquisition phenomenon necessitate different causal factors to be taken into account, and the description of different systems in which those causal factors interact.

According to Elman et al. there are different levels of interaction which have their specific environment. The relation between those levels and their environment has different outcomes. The relation between different levels of interaction, their relevant environment and the outcome of these interactions are illustrated in a table as follows:

Table-2:[93] Illustration of the outcomes of interaction between the organism and relevant environment at different levels. 'Innateness', according to Elman et al., is one of those outcomes.

Level of interaction	**Environment**	**Outcome**
molecular	**internal environment**	**INNATE**
cellular	**internal environment**	
organism-external	**species-typical environment**	**PRIMAL**
organism-external	**individual environment**	**LEARNING**

[92] Although individual performances are taken into account in the model provided by Kelso, it is applicable to all individuals in the same way. Indeed, in my opinion, it is a model which intends to explore universal principles of learning process as interaction between individual predispositions and environment. In other words, it might be claimed that it is a theory of learning processes of individual competences.

[93] Table-2 is quoted from Elman et al., p. 23.

According to them the term ‘innate’ refers to the supposed brain structure and the outcomes of the interactions which take place internal to the organism. The interaction between species typical-environment (e.g. gravity) and the organism has primal as its outcome which might be indistinguishable from innate.[94] The interaction between the organism and species-typical environment provides as reliable and universal outcomes as the interaction between molecular/cellular levels and internal environment provides. For example, we expect that a normal child who is born in a language speaking society acquires a language. It is not only because children are born with an innate brain structure proper for language acquisition, but also because there *is* linguistic data from the environment, there is need for communication etc. However, PDP models generally intend to explain the innate part by neural networks. Although neural networks are not directly explanations of the interaction between biological stuff, they respect biological plausibility. According to this view, biologically plausible mechanism which produces the innate content is provided by neural networks. In neural networks, qualitatively different outcomes are obtained by a single mechanism, rather than supposedly distinct faculties/organs for each outcome.

Indeed, the idea that acquiring a language is not only due to genetic predispositions and learning is compatible with the view proposed by Chomsky for different factors which enter into the growth of language in the individual. He adds “[p]rinciples not specific to the faculty of language”[95] to the factors which are genetic factors and experience. According to him genetic factors enable the child to interpret the

[94] Elman et al., pp. 22-23.

[95] Chomsky (2007), p. 15. Though, according to Chomsky UG is provided genetically but ‘principles not specific to the faculty of language’ are not. If we do not define innateness as Elman et al. do, those principles are not innate. However, I follow their definition and thus accept that those principles not specific to language faculty are innate, in the sense that they are not learned and restrict the outcome of development.

environment as linguistic experience and experience permits variation in a very narrow range.[96] The constraints which come from the inner processes of organism and the species-typical environment support the idea that experience permits variation in a narrow range. Thus, if we distinguish the content of the outcomes of innate procedures and those procedures, the definition of innateness provided by Elman et al. is compatible with the view of Chomsky.

> An alternative definition is to reserve the term innate to refer to developmental outcomes that are more or less inevitable for a given species. That is, given the normal environment for a species, outcomes which are more or less invariant between individuals are innate.[97]

I would conclude that the contents of innate outcomes are not only genetically endowed, but also emergent through an innate mechanism explanation of which is provided by PDP.

Last sub-target in the explanation of language acquisition is the relation between predispositions of individual and its external environment which would correspond to the last line in the Table-2. I claim that DST provides the most proper account for this sub-target. 'Intrinsic dynamics' of Kelso provides the account of predispositions of individual and the coupled system in which those predispositions and the environment of the individual interact provides the explanation of process of learning. Although the mechanism which produces those predispositions is not explained in the theory, they are determined and the evolution of intrinsic dynamics by the process of learning is explained clearly. These processes and notions will be clearer in the fourth chapter.

I think the pluralistic view proposed here is also compatible with the theory provided by Bolender for social cognition. He provides an explanation of the social cognition

---

[96] Chomsky (2007), p. 15.

[97] Elman et al., p. 22.

from the perspective of CTM and suggests a research program for the emergence process of the basic relational models of social cognition from the perspective of dynamical approaches.[98] According to his theory, social interactions are formed by a grammar in which four basic relational models and a null relation are combined according to rules of that grammar. The four basic models are communal sharing (CS), authority ranking (AR), equality matching (EM) and market pricing (MP). The relations in which members of the group are regarded as same are constructed by CS, the relations in which there is a set of hierarchical duties and rights attributed to members are constructed by AR, the relations in which members of the group have equal rights and duties are constructed by EM, and finally, the relations in which decisions are made through analyses which respect ratios or rates are constructed by MP. These basic models show descending degrees of symmetries when a transformation in a relation is made – i.e., the difference in the relationship by several transformations increases from CS to MP. These models, then, correspond to a descending chain of subgroups of symmetries. “Descending chains of subgroup types are a phenomenon widely observed in nature; their presence in social cognition is consistent with there being a relevant neural network, the activity of which can undergo symmetry breakings.”[99] The theory and the terms will be clearer in the next chapter but the quotation suffices for the dual nature of his theory, for now.

I claim that we gradually came to a conception of cognition in which contributions of different perspectives are important and distinguishable, yet. Recent theories imply *perspectives* as ontological category for cognition. For example, explanations of sub-targets in explaining language acquisition are provided by different approaches. These approaches describe the system under consideration by using different system

---

[98] In Relational Models Theory Fiske (1993) presents four basic relational models, their definitions and their recursive character. Besides, he provides four measurement scales’ correspondences to those basic models and argues for the innateness of those models. The theory proposed by Bolender is a refined version of RMT. The notions mentioned such as ‘measurement scales’ and their relation to RMT will be clearer in the next chapter.

[99] Bolender (2008), p. 1.

conceptions. CTM describes its target system as componential, and PDP and DST describe their target systems as integrated. Since these theories decompose the system differently and all of them seem to claim that they explain the target phenomenon, they seem to be competing theories. However, sub-targets in explaining language acquisition might demand different theories and different conceptions of cognition. Besides, understanding of language acquisition might demand cooperation of these different conceptions; just like understanding of an organism demands contributions from anatomy, genetics, physiology, etc.

The version of RMT which is proposed by Bolender implies two system conceptions together as responsible for processes at different stages. In the first stage the system is integrated and performs dynamical processes. Through those processes discrete basic items of the lexicon are obtained for the use of computational component. Computational component which works upon those discrete items exhibits discrete processes. The overall system in RMT has componential structure and one of the sub-systems/components of the system, i.e., preonic systems, has integrated structure and continuous processes. Thus, the system can exhibit both discrete and continuous processes at the second stage. RMT will be described in the next chapter, however from this sketch of the theory suffices we can observe that different conceptions of a system can work together without contradiction. Furthermore, these different conceptions in the system are complementary.

CHAPTER 5:

APPLICATIONS IN COGNITIVE SCIENCE

In this chapter, I will try to explain my thesis about cognitive science, its current status and its possible future according to the distinctions described in chapter one (C/I-D/C) and issues related to idealization described in chapter two. I am planning to defend a kind of pluralism for modeling cognitive phenomena – i.e., multiple models idealization. According to multiple models idealization the phenomenon under consideration can only be understood via multiple, sometimes incompatible, but complementary models. Besides, multiple conceptions of cognition might be necessary to understand different aspects of it, as well as multiple models. Since, all of the theories under consideration (namely, CTM, PDP and DST) have special role in explaining cognition, and/or different aspects of it, they contribute to our understanding of cognition. Because of the increase in the number of the hybrid theories, I see a tendency to a unified conception of cognition which can be understood via multiple models idealization. Prima facie, my position seems contradictory. However, there seems to be coevolution of different perspectives in which the contributions of them, and system conceptions which are proposed by them, become more integrated and the boundaries of different perspectives become blurred. This tendency seems to be directed to a coherent conception of cognition.[100] I hope my claims will be clearer after explaining them over several examples from cognitive science. Before explaining my thesis over those examples, reminding the concepts mentioned before seems useful now.

There are two basic components in the description of a system: structure description in which systems can be described as componential or integrated systems (C/I) and

[100] However, I should admit that I can be wrong about the *future* of cognitive science. Wimsatt might be right about the nature of complex phenomena which can only be understood by contributions from partially autonomous, partially interdependent perspectives (p. 240).

process description in which the processes of systems can be described as discrete or continuous (D/C). Componential systems and integrated systems are both decomposable systems but they differ in the degree of decomposability. Componential systems are nearly decomposable systems which mean that components of the system have intrinsic, autonomous functions whose contribution to the system can be identified. Integrated systems, on the other hand, are minimally decomposable systems. These kinds of systems also have components, yet their components are simple in the sense that the specific contribution of an individual component to the system cannot be identified or identifying it is useless, and interdependent in the sense that organization of the system is important for the function of the individual components. Process description of these systems can be discrete or continuous; it depends on several factors such as chosen time scale, the nature of the components of the system, etc.

Cognitive capacities, according to CTM, are performed via different faculties which are specific devices devoted to related capacities. Faculty of language is one of them.[101] According to this facultative/modular account, cognition is conceptualized as a componential system in a hierarchical manner whose components are autonomous subsystems between which there are minimal interactions. On the other hand, every faculty also has a structure description and process description. I will mostly try to describe the system of the faculty under consideration (for example, language faculty in the section for language acquisition), rather than the full account of cognition according to CTM.[102]

---

[101] Fodor.

[102] "Strictly speaking, CTM is not committed to faculty psychology. One could view mind as computational without dividing it into distinct power. It happens, however, that most CTM people also adhere to faculty psychology." (John Bolender, personal communication) I will deal with faculty psychology in describing CTM. One reason for this is the fact which is stated by him: Mostly, models in CTM view cognition as a componential/modular system. Another reason is that the model provided by CTM for the chosen example, i.e., language acquisition, describes the system as a modular system.

I will attempt to describe the systems assumed by CTM, PDP and DST for language acquisition phenomenon in the following section.

a. Language Acquisition[103]

The phenomenon, language acquisition, implies a change in the system, thus can be seen as a process description in nature; even if this process is as short as a *switch due to a trigger from the environment* in the structure.[104] However, every process description is tied to a structure description which will be changed. Indeed, process description can be explained shortly as the description of the processes which *produce* difference between the states of a structure.

Trying to explore the nature of language acquisition and hence the proper structure and process descriptions in a theory of language acquisition are the main concerns in this section. I will try to find a common ground between CTM, PDP and DST in order to be able to compare them. Thus, I will attempt to redescribe language acquisition according to CTM, PDP and DST as system descriptions (namely, the combination of structure description and process description); apologies for my mistakes in translation of them into system descriptions.

The structure of the system which will be changed by language acquisition can be regarded as pre-learning structure.[105] Only in CTM pre-learning structure which is a

---

[103] For the sake of neutrality, I prefer 'acquisition' rather than 'learning' to indicate the *access* to using and understanding a natural language, in other words, competence in using a natural language (See Lindner & Hohenberger for an explanation of the relationship between the notions learning, growth and acquisition).

[104] Pinker, pp. 111-112.

[105] Since every theory has a different conception of already existing structure (for example, innateness in CTM, intrinsic dynamics of Kelso, kind of *tabula rasa* of some early PDP models etc.) before language acquisition which is directly relevant to acquisition, I will use *pre-learning structure/device* to indicate those structures/mechanisms which will be changed by the acquisition process, again, for the sake of neutrality. Pre-learning structure can be specific for the acquisition of language (viz., a device devoted to language acquisition), or only specific to humans which show predisposition to language acquisition but not as a devoted device for language acquisition. These accounts will be clearer throughout the section.

componential system is devoted to language. In PDP and DST the supposed structures to be changed are integrated systems and are not devoted to language.[106] These three theories also have different accounts for the process of language acquisition, and hence different process descriptions. Roughly, the process of acquisition in CTM is a shift in the structure by the primary input from environment, but in PDP and DST it is mainly a learning process that takes place in time and embedded in environment. Both approaches (i.e., language acquisition as a maturation process and as a learning process) seem to capture different aspects of the language acquisition phenomenon and have valuable consequences in the explanation of sub-targets in understanding language acquisition phenomenon.

1. CTM

As mentioned above, according to CTM, language acquisition is not a learning process, rather maturation of a biological organ in human brain, i.e., language organ. CTM proposes language faculty as responsible for language acquisition. Before trying to explain language acquisition according to CTM, I need to explain several concepts.

Almost all resources which explain language acquisition according to CTM attribute special importance to the distinction between competence and performance. Despite the fact that we generally use language with mistakes, we know that our language include lexical items and a grammar. According to Chomsky, this knowledge is our competence in language and our usage of this knowledge is our performance.[107] Language faculty is about competence, not performance. It is our organ which enables us to learn and use our languages. The competence of a native speaker due to language acquisition is I-language which enables the native speaker to generate and

[106] However, this does not mean that PDP and DST do not have an account for innateness of language, i.e., its specificity for humans. Their described structures also have accounts for the uniqueness of cognition but not specific to language acquisition.

[107] Chomsky (1995), p. 14.

understand sentences of his/her language. "… [A] grammar of a language is a theory of the I-language …"[108] Universal grammar (UG) which is a grammar of all possible languages is generalized from grammars of these I-languages.

In general, grammar (or language as part of cognition) consists of lexicon, syntax, logical form component (LF), and phonetic form component (PF). Lexicon and syntax form the syntactic structure (f_{syntax} (lexicon)) in a grammar. LF is function of lexicon and semantic representations (SRs), and PF is function of lexicon and phonetic representation (PRs).[109] They are interface subsystems which link language faculty to other faculties such as thought systems and speech systems, respectively. Grammar, by combining these components in a specific way mediates between thought systems, speech systems and language faculty.[110] I will attempt to symbolize them as follows:

Grammar = f (f_{syntax} (lexicon), LF, PF)

LF = g (SRs, f_{syntax} (lexicon))

PF = h (PRs, f_{syntax} (lexicon))

LF links the syntactic structure to SRs and hence to thought systems. PF links the syntactic structure to PRs and hence to speech systems. The relation between components of a grammar might be described as follows:

The relation of components regarding thought systems:

f_{syntax} (lexicon) → LF → SR

---

[108] Radford, p. 7.

[109] I use 'function' here to indicate the proposed relation between components. For example, grammar is a function of syntactic structure, LF and PF in the sense that specific relations of these components provide us grammar. If they are inappropriate, I apologize for the symbolic illustrations.

[110] Radford, pp. 9-10 and Chomsky (1995), p. 34.

SR → LF → f_{syntax} (lexicon)

The relation of components regarding speech systems:

f_{syntax} (lexicon) → PF → PR

PR → PF → f_{syntax} (lexicon)

This representation, I assume, illustrates the structure description of a grammar. Every language has its own lexicon, own syntactic rules, own thought systems and own speech systems which constitute I-languages. I-languages are used to generate symbolic objects such as phrases; sentences etc. which are structural descriptions. I-language is a *machine* which produces infinitely many symbolic objects from finite procedures and lexicon. These objects include items from all subsystems of a grammar.[111]

Beyond the rules of particular languages, "… UG principles which are innately endowed are wired into the language faculty …"[112] Thus, IS of language faculty overlaps UG. For the sake of brevity, I will not try to list and/or explain all properties of merge, the rules of UG, or the parameters. I assume that the explanation provided here is enough to understand the componentiality and discreteness of a grammar. Although it does not show all of the properties of simply decomposable systems,[113] it has minimal interaction among its subsystems (i.e., syntactic structure, LF and PF) and those subsystems are partially autonomous. Furthermore, those subsystems have also componential structures.[114]

---

[111] Chomsky (1995), p. 14-15, p. 35.

[112] Radford, p. 16.

[113] Indeed, since componential systems are composite systems, to be a componential system, grammar should not show all properties of simply decomposable systems. (See Chapter 1)

[114] Chomsky (1995), p. 34.

How those possible I-languages (UG) settle down as a particular I-language, then? In other words, how does a person acquire his/her language? The answer of CTM for this question is this: Language acquisition is the maturation of language faculty by the primary linguistic data (PLD) which provide the values of parameters which will be settled in UG.[115] It is a determined path (biologically) which will be followed due to genetically endowed properties of our minds/brains. We only need a trigger. There is an initial state (IS) of language faculty before the acquisition process (at around birth) which is set to a final state (FS – grammar of particular language(s)) by the input from environment as PLD. Hinzen describes the process as such:

> Exposure to *primary linguistic data* (PLD) thus leads to switching from one possible language to another, and we may think of language learning as an IS converting PLDs into a particular FS: IS (PLD) = FS, via a number of intermediate stages.[116]

The final state of language faculty is competence of a native speaker, namely, I-language. The language acquisition process according to CTM can be described as follows:

Experience (input/ PLD) → UG (IS of language faculty) → grammar (output/FS of language faculty)

Input from environment sets the value of parameters and the grammar of the particular language is acquired. The change in the system is just a one-step process which is clearly discrete. The structure description of this system – the structure to be changed by acquisition process – is language faculty which is a componential system, as explained above.

[115] Illustration of the procedure of producing phrases and of one of the parameters, i.e., order parameter, is provided by the description of "X-bar architecture". (For details, see appendix).

[116] Hinzen, p. 119.

Since according to PDP and DST the language acquisition phenomenon is conceived as a learning process, I will not specifically try to explain any language acquisition model of these theories; rather I will try to describe their conception of learning process. Shortly, according to their conception, the pre-learning device which is not specific to language acquisition evolves by experience.

2. PDP

The explanation of language acquisition in PDP is provided by models which use neural networks as mathematical tools. The models in PDP, simply, work on brainlike networks which are "… collection[s] of interconnected elements of units."[117] The components of the system (i.e., units) are not components like the components of a componential system, rather, they are simple parts of the system with no specific and autonomous function which can be attributed to one unit and they are interconnected. Although the amount of the effect of one unit on the other connected unit can be computed via weights of connections one unit's *specific role* in the system is indistinguishable (or distinguishing it is useless). The cognitive structure which PDP models propose is composed of simple units with their activation functions and connections between these units. Thus, they are integrated systems. Although they are decomposable too, the degree of their decomposability is much less than that of componential systems. For example, we can identify units but as mentioned above we cannot, or need not, identify their specific contribution. Besides, these units contribute to the system by their one-to-many connections to other units. Thus, the organizational properties of the system determine the contribution of a particular unit. The activation value of one unit can be accepted as the effect of the unit on another unit. The effect is transferred to the latter unit via a

[117] Smolensky et al., p. 1.

function. The amount of the effect of one unit on the other is specified by a real number, i.e., "the weight or strength of the connection".[118]

> [T]hese systems are viewed as being plastic in the sense that the pattern of interconnections is not fixed for all time; rather, the weights can undergo modification as a function of experience. In this way the system can evolve.[119]

The learning processes, including language acquisition as a learning process, proposed by PDP depend on this kind of evolution of the system. I will turn to the conception of language acquisition phenomenon in PDP. Firstly I will try to distinguish structure descriptions and process descriptions of PDP models. The description of the relations between units of the network is structure description of system. The structure description of the system is as explained above, indicates an integrated system. In neural networks there are two main types of connectivity which determine the process description of the system: feedforward connectivity and feedback connectivity. If the process description shows only the change of input to output, the connectivity of the network is feedforward. Thus, the process description of the system is a static function of input in which every input changes into an output without the effect of the output. If the relation between input and output is iterative and hence, output of the system enters system as input in the next stage the connectivity of the network is feedback. In the case of feedback connectivity, the process description shows input-output relation of the system as a function of time, rather than output as function of input. The difference between steps is process description of the system. These kinds of process descriptions can be discrete or continuous depending on the conception of time – i.e., taking time into consideration as discrete or continuous.[120] I will not discuss feedforward connectivity and feedback

---

[118] Rumelhart et al., p. 46.

[119] Ibid, p. 46.

[120] Smolensky et al., pp. 3-4.

connectivity further; rather I will try to explain the conception of learning phenomenon in PDP models.

PDP models conceptualize learning as the evolution of the network due to inputs from environment. The description of this evolution is provided by describing the change in weights of connections between units (namely, weights, for short). Every time an input enters system it affects the weights which in turn affect the output. The process of introducing the system inputs repeatedly is called training. Although there are differences between different learning models of PDP mainly the system works as this: firstly inputs and the target are presented to the system and then the system computes the difference between the output and the target (namely, error). Weights are adjusted in order to decrease error gradually through the training process. The mechanism holds for both feedforward connectivity in which inputs and outputs of every training step are partially independent, and feedback connectivity in which output of the system enters system as input in the next step. By '*partial independence*' of the training steps in feedforward connectivity, I mean that although every training step affects the next step due to changed weights, this effect is not as direct as in a feedback connectivity. In feedback connectivity, output of one step directly affects the next step, in addition to the effect of changed weights.

The system description of PDP models of language acquisition is mainly a process description in which an integrated structure evolves as roughly described above. The discreteness or continuity of the process depends on the type of the connectivity in the system and the conception of time. Different PDP models of learning propose slightly different models and different conceptions of the phenomenon. For example, Rumelhart and Zipser propose four different paradigms of learning according to PDP: auto associator, pattern associator, classification paradigm, and regularity detector. (1) According to auto associator paradigm network stores presented patterns through training and when a similar pattern presented to the system, the original pattern is retrieved. (2) In pattern associator paradigm the task is not retrieval of the original pattern, but association of pairs of patterns. The system is supposed to produce one of the pairs when the other is presented after learning the task. (3) The system proposed by models constructed according to classification paradigm associates the stimulus patterns and a fixed set of categories which are presented to the system. After the training process system learns to classify a particular stimulus or a different but similar version of it. (4) In the last paradigm, viz. regularity detector, task of the system is similar to the preceding paradigm, viz. classification paradigm. However, now, system is supposed to construct its own categories by presentation of stimulus patterns with some probability. Categories constructed by the system captures the important features of the stimulus patterns.[121] The models of these four paradigms are constructed by multiple models idealization. They are constructed according to different paradigms of same theoretical perspective and are related, complementary but incompatible models because of having different claims about the nature of the phenomenon.[122]

3. DST

I will mostly depend on Kelso and van Geert in explaining conception of learning phenomenon of DST. Although Kelso does not propose his theory exactly for language acquisition, I assume that his theory about coordinated behavior also provides the conception of learning process in DST. van Geert proposes a similar approach to the approach of Kelso while he mostly depends on the example of lexical learning. According to him "*[d]namical systems theory* is an approach to the description and explanation of *change*"[123]

Both of them accept learning as the change in already existing structure. It is a similar approach to CTM in this respect, i.e., DST also attributes special importance

[121] Rumelhart & Zipser, p. 161.

[122] Weisberg (2007b), p. 645.

[123] van Geert, p. 242.

on pre-learning structure. However, the pre-learning structure of DST, unlike that of CTM, is not a static structure which is indistinguishable among individuals. Indeed, DST proposes a mechanism for learning in which individual differences are important. According to proponents of DST a theory of learning should account for, for example, why one individual learns in a relatively shorter time than the time in which other individual learns. Thus, intrinsic dynamics of Kelso, for example, respects individual's already existing capacities and knowledge.[124] In other words, DST has an account of performance. It idealizes a kind of universal *law* from individual differences according to which we can also determine those differences. Another important difference between DST and CTM is their positions regarding environment and time. The former approach, unlike the latter, suggests that individual and the environment constitute a coupled system which evolves in time through interaction of individual and his/her environment. For example 'intrinsic dynamics' and 'specific parametric influences' in Kelso, and 'endo- and exo-systems' in van Geert constitute such coupled systems – former terms imply the individuals' pre-learning devices and the latter terms indicate the environment. On the other hand, CTM idealizes acquisition process in which time is a step and environment is trigger which is the cause of the switch in the system.

I will quote their mathematical models and try to explain their accounts over these models. Firstly, Kelso provide a formulation of his account in which evolution of the system depends on intrinsic dynamics (already existing structure) and specific parametric influences (input from environment, or task in learning process) as follows:

$$\phi \quad = f_{intr}(\phi) + f_{inf}(\phi)$$

Left hand component is intrinsic dynamics component which is the already existing structure of the system. Intrinsic dynamics illustrates the pre-learning structure in

[124] Kelso, p. 163.

which there are individual's experiences and capacities prior to learning. Right hand component is specific parametric influence which comes from environment specific and relevant to the pattern. "We say that a behavioral pattern is learned to the extent that the intrinsic dynamics are modified in the direction of the to-be-learned pattern."[125] System is applicable to both learning process of a child, or infant, and learning process of an adult as a theoretical framework. Intrinsic dynamics is illustrated as an attractor layout according to Kelso which changes wholly through learning process. Individuals behave according to already existing attractor layout and by learning the attractors of the system changes, thus the behavior of the individual changes, too.

Similarly, van Geert proposes an equation for lexical learning as follows:

$L_t = K / (1 + (K / L_0 - 1) e^{-rtK})$[126]

L: growth level or lexicon

r: growth rate

K: carrying capacity

t: time[127]

I will not deeply investigate these equations from Kelso and from van Geert. However, the second one is important in showing similarity by taking into account almost same variables, i.e., already existing knowledge and capacity of the individual. Both approaches assume integrated structures for explanation of the change occurred among learning process. Their components are highly integrated.

[125] Kelso, p. 163.

[126] van Geert, p. 247.

[127] Hohenberger & Peltzer-Karpf, p. 495.

Change in the system can be formulated as discrete or continuous depending on the data.

Two idealization types are important in DST models: They aim to be complete theories. They try to include all relevant factors in the model, thus they are constructed by Galilean idealizations according to their long-term aims, or final goals. However, they also try to be minimalistic, i.e., they include only the relevant factors which give rise to phenomenon, namely, learning process, and the models are simple in the sense that they include as few components as possible.

In my opinion language acquisition phenomenon as a target of these three approaches can be divided into three sub-targets: language, pre-learning device and the process of learning in interaction with environment. One of my claims in chapter two might be controversial which this is: According to their language acquisition models CTM, PDP and DST explain those sub-targets respectively. Firstly, it is controversial especially because of the claim that PDP explains pre-learning device. It is generally accepted that the theory which attributes special importance to pre-learning device is CTM. Besides, PDP is generally accused of including an assumption of *tabula rasa*. However, proponents of PDP object the accusation by distinguishing the content of pre-learning device from the mechanism of it. According to them, for example, language is the content of developmental outcomes.[128] They define 'innateness' as a term which refers to "… putative aspects of brain structure, cognition or behavior that are the product of interactions internal to the organism."[129] Accordingly, the explanation of the pre-learning device should be the explanation of the processes through which the innate content emerges. Indeed, the claim can also be restated as this: The tools PDP provides are *better* than that of CTM for the explanation of the evolution and the mechanism of pre-learning

[128] Elman et al., pp. 21-22.

[129] Elman et al., p. 23.

device; but CTM gives the most sophisticated explanation of the innate content before language acquisition process, its final structure after a language is acquired and their relation to other capacities.

The other controversial part of my argument is the claim that DST explains the process of learning in interaction with environment. Of course, it is not a controversial claim on its own; DST really intends to explain that sub-target. However, there is an implicit claim which might be controversial that DST explains that sub-target *better* than PDP does. That implicit claim is controversial because of the importance of environment in PDP models. They also claim to give an account for the relation between the existing structure and the environment. In my opinion, their models provide the working principles and the evolution of the already existing structure, rather than its relation to the environment. Of course, they explain the evolution of that structure through its relation to environment. However, the target of the models is not that relation, rather explaining the effect of that relation on the mechanism. I agree with the following claim:

> Learning, viewed as the mere strengthening of synaptic connections, tacitly ignores the presence of any meaningful relation between the things being learned and the intrinsic organization of the system doing the learning. …. In particular the cooperative or competitive interaction between specific learning requirements and intrinsic organizational tendencies … has important consequences for how learning is to be understood.[130]

Thus, I claim that CTM explains the structure of the innate content and what happens to the innate content of language by acquiring a language, PDP explains through which mechanisms that content emerges and changes (mechanism and evolution of pre-learning device), and DST explains how learning process takes place between the predispositions of organism and environment. Thus, their differences come from the sub-targets which they deal with, even if they intend to explain language acquisition

[130] Kelso, p. 170.

wholly. The final conception of language acquisition to which all these three conceptions contribute would belong to the *perspectives* category ontologically and these theories are perspectival theories. Although their conceptions of the phenomenon overlap partially, we can still distinguish their contribution to our final conception and they all provide explanations for different sub-targets/aspects of language acquisition; none of these explanations is exhaustive when considered alone.

b. Social Cognition

I will mostly depend on John Bolender's approach in explaining social cognition. He makes a distinction between the traditional social cognition theories and Relational Models Theory (RMT).[131] According to this distinction, RMT is not directly a theory of interpersonal perception;[132] rather it is a theory of "… formal properties of the mental models used in structuring social relationships."[133] However, RMT redefines interpersonal perception as application of the basic relational models.[134]

John Bolender's approach has two main stages. In the first stage the lexicon is produced by a dynamical process for the use of the computational component. In the second stage computational component works upon the lexical items in a discrete manner. There is a universal moral grammar (UMG) in which formal properties (mods) and semantic properties (preos) of basic relational models of RMT have role. Firstly, lexical items are produced by social pattern generator (SPG) and then, they are copied from lexicon which includes SPG for the use of UMG. In UMG, lexical items which are copied from lexicon are merged by following a set of syntactic and

[131] Fiske, A. P. (1991), *Structures of Social Life: The Four Elementary Forms of Human Relations*, NY: The Free Press; and Haslam, N., (Ed.) (2004), *Relational Models Theory A Contemporary Overview*, Mahwah, New Jersey and London: Lawrence Erlbaum, 2004, cited in Bolender (2010) p. 11-15.

[132] Bolender (2010), p. 15.

[133] Ibid, p. 7.

[134] Ibid, p. 15.

semantic rules. Infinitely many relational models can be produced by this system and those models are used in preonic interfaces for the preonic interpretation of the individual set. The second stage of the process clearly illustrates a system from the perspective of CTM.

I will firstly try to explain the first stage in which basic relational models are produced by a dynamical process. There are four basic relational models which include formal properties and preonic properties, and a null relation. Basic relational models are communal sharing, authority ranking, equality matching, and market pricing.[135]

Communal sharing (CS): In the relationship structured according to this relational model, all members of the group are regarded as same in their relationship. One member is identified with the group. “Nationalism, romantic love, racism, and indiscriminate killing of anyone outside of the group in retaliation for an attack upon the group are forms of CS.”[136]

Authority ranking (AR): This model shows a relation among linear hierarchy. “Subordinates are expected to respect and obey, and superiors enjoy greater prestige whilst also having duties of protection and care for their inferiors.”[137] The typical example for this group can be monarchies.

Equality matching (ER): The relations according to this model are constructed among people which are regarded as equal, distinct but interchangeable. Rules for taking turns can be an example of ER.

---

[135] Ibid, pp. 11-14.

[136] Ibid, p. 12.

[137] Ibid, p. 12.

Market pricing (MP): "Interactions are conceived in terms of rates or ratios such as prices, wages, interest, rents, tithes, or cost-benefit analyses."[138] A competitive market in which firms and consumers make decisions depending on cost-revenue/benefit analyses, judgments made by *calculating* right proportion of reward or punishment are forms of MP.

Null relation: This shows a specific relation in which none of the above models is applied. A person who constructs his/her relationships according to null relation (for example, psychopaths) might treat other people as if they are not living entities, or humans, but objects.

As proposed by Bolender, formal properties of these basic relational models have a dynamical emergence process. He claims that the innateness of UMG is not only genetically coded, but it also dynamically develops through a symmetry breaking process.

> What is actually encoded genetically may be an extremely fundamental feature of CS, the rest of social cognition being the result of physics and interactions with other mental faculties. Perhaps the only element of CS represented in the genome is sympathy or love, essentially the view anticipated by Darwin in his chapter on the moral sense in *Descent of Man*.[139]

The symmetry is detected in relational models by observing the difference which can be obtained by some transformations. If a transformation in the mental representation of the system (viz., social faculty) makes a social difference then the representation is asymmetric. On the other hand, "[t]ransformations which make no difference to the output of the social faculty are symmetries."[140] According to this account, genetically coded sympathy, as the core element, develops into the formal parts of four relational

---

[138] Ibid, p. 12.

[139] Bolender (2007), p. 264.

[140] Bolender (2008), p. 12.

models by self-organization. Finally, the complex relational models with their semantic parts are yielded by interaction with other mental faculties and experience.[141]

> Each relational model takes the form of one of the four measurement scales (Fiske, 1991), and symmetries in social cognition correspond to symmetries found in measurement scales. Communal Sharing (CS) takes the form of a nominal scale, AR an ordinal scale, EM an interval scale, and Market Pricing (MP) a ratio scale.[142]

These scales and hence relational models correspond to a descending sequence of subgroups which show descending symmetries. According to this, the group for ratio scale will be a subgroup of the group for interval scale which is a subgroup of the group for ordinal scale and so on. The group for nominal scale will contain all subgroups and any member of the group for ratio scale will also be a member of all other groups. To explain the relation between the symmetries found in relational models and that of scales, suppose that a father punishing his child by depending on a relation constructed by AR. The type of punishment will be decided by MP and equality of all children according to obeying a rule of the family is constructed by EM. In our example the rule is this: 'Do not break the vase. If you break, your pocket money will be decreased.' To be able to trace all models, let us suppose that this family share their meals according to CS. The last type of relation would show the highest symmetry. For example at a meal of this family anyone can eat any amount of the food; it is not shared among members of the family according to any ratio or hierarchy. In this case any transformation in their relation would not cause a difference in their social relations. In other words, if we change the role of the father and a child, this change will not cause a change in their relations according to sharing the food; still they can eat as much as they want. Changing roles in a relation constructed by a less symmetrical relational model, AR, would cause difference in

[141] Bolender (2007), p. 264.

[142] Bolender (2008), p. 12.

the social relationship among members. For example none of the children can punish his father when he breaks the vase. In the relation constructed by EM which is less symmetrical, in our example fault-punishment relation, if we change the role of the child who breaks the vase with an *innocent* one, we would change the relation. Doing this would cause an unfair punishment. Obviously, AR and EM are less symmetrical than CS. AR is more symmetrical than EM; though, it is less obvious. The order of the hierarchy among members of this family is not restricted specifically but the equality of the children is restricted. I mean, the right of the father to punish his children is indifferent according to fault. For example, among different pairs of fault-punishment such as breaking the vase-decrease in pocket money, or going out without permission-staying at his/her room for two days etc. the role of the father does not change; he judges the punishment. On the other hand, according to EM changing these different pairs cause difference in the social relationship. For example if father decides to decrease pocket money when one of his children breaks the vase but decides not to permit going out of his/her room for two days when other child breaks the vase, these difference in his decision would cause difference in the relationship between his children; i.e., it will no more be a relation constructed by EM. Thus, EM is less symmetrical than AR. Similarly, MP is the least symmetrical relational model. The amount and the type of the punishment are decided in proportion to the fault which implies a more restricted relational model. For example decision for the punishment staying at room for three days, rather than two days, would yield a difference in the social relation; it might be unfair to decide that punishment.[143]

The correspondence of relational models to measurement scales and their relations according to symmetries implied by them, show a symmetry breaking process. They are produced by social pattern generator (SPG). SPG is part of a component, i.e.,

[143] Bolender (2008), pp. 12-15.

lexicon. Symmetries of the relations constructed by relational models decrease through relational models, namely, CS, AR, EM, MP, respectively.

> Fiske notes that the less symmetrical the model the later it first manifests itself in childhood. Communal sharing is the first to manifest itself, appearing in infancy. Authority ranking appears in judgment by the age of three. Equality matching first evidences itself soon after the fourth birthday, and market pricing after the eighth (Fiske, 1991: 48-9, 1992: 696).[144]

It is crucial to bear in mind that the construction of these primes through self organization with broken symmetries is only about the formal structure of the system. The rich semantic features are gained over this core structure by interactions between other mental faculties and experience.[145]

There is a distinction between formal properties of RMT which has five primes and their grammar, and preonic systems which can be seen as the semantic component of UMG. Every model has a formal part (mods) and a semantic part (preos). Formal parts of these models (namely, CS, AR, ER, MP and null relation) constitute lexical items of a componential formal system/grammar. Preos are, roughly, determiner of the context in which relational models apply.

> These are the four fundamental, innate, human relational proclivities. To signify that they are cognitively modular but modifiable modes of interacting, I call them "mods." However, these open-ended generative potentials are insufficient in themselves to determine action or evaluation, or permit coordination. In order to use these mods to act or to interpret others' action, people need socially transmitted prototypes, precedents, and principles that complete the mods, specifying how when and with respect to whom the mods apply. I use the term "preo" to signify the class of paradigms, parameters, precepts, prescriptions, propositions, and proscriptions that can be conjoined with mods. A mod must be conjoined

[144] Bolender (2007), p. 259.

[145] Bolender (2008), p. 20.

with a preo that complements it to generate a specific cultural coordination device.[146]

These mods are relational structures that can be implemented in innumerable ways. Hence a mod requires cultural complements that complete it, a set of *preos*. Preos are cultural prototypes, precedents, precepts, or principles that orient the mod by indicating with whom, what, when, where, and how it is implemented in any specific instance. Preos can be folk tales, proverbs, holy books, sermons, children's stories, movies or television shows, admonishments or punishments experienced or observed, explanations, or experiences of interaction that the child is able to use to orient her mods. The preos are what is culturally transmitted that the child must detect and learn.[147]

The formal parts of the models are basic lexical items (mods/primes) as follows:

c = CS, a = AR, e = EM, m = MP, n = Null[148]

Lexicon, computational component and preonic systems are the components of the system. Lexicon provides basic lexical items for operations of computational component. Computational component provides infinitely many combinations of those lexical items (complex mods) for the use of preonic systems. The primes can be applied in a recursive manner. There are two operations according to process description of the system: copy and merge. Copy operation, as its name implies, copies items or single tokens from the lexicon. Those tokens can enter merge operation. By merge, tokens of mods are combined.

Lexicon $\xrightarrow{copy}$ Computational component $\xrightarrow{merge}$ Relational models → Preonic interface → Preonic systems

Syntactic rules and semantic rules describe the relations between primes. According to syntactic rules, components can enter derivation once a time, derivations should

[146] Fiske, 2004, pp. 3-4, quoted in Bolender (2010), p. 22.

[147] Fiske, forthcoming, quoted in Bolender (2010), p. 22.

[148] Bolender (2010), p.51.

have an end and the last step of the derivation can only be a merged prime.[149] Semantic rules are also about formal properties of these mods. They do not describe preos which determine the context in which these formal structures apply. According to semantic rules, "[o]nly a set can be interpreted as a social relatum" and "[a] dominant prime determines the type of relation." For example, suppose that in a family father decides to punish one of his children for his/her fault, say, breaking the vase and the punishment is a decrease in his/her pocket money. In this case there will be a relation as such: {a, {m}}; the relation is MP, dominated by AR. Father has right to punish his child according to the linear hierarchy of their relation which is one of the characteristics of AR. On the other hand, father decides the punishment in a proportional fashion. Further, he might decide the amount of the decrease in his child's pocket money which can compensate the price of the vase. These features of the decision of punishment in our example illustrate MP. If we think that this relation would hold for all children of the family, we can reconstruct the relation as {a, {e, {m}}}; a relation of MP dominated by EM dominated by AR. AR is still the dominating mod. The children will be regarded as distinct but MP will be applied to them indifferently; thus father would punish any of his children in the same way due to same fault which is a characteristic of EM.

The structure and process descriptions of the system are similar to those of language faculty without its interface systems. There is also an interface system in UMG; i.e., preonic systems which links the formal structure to social relations specific to cultures. UMG mediates between different cognitive faculties. However, this time we have a much more complex relation between different faculties than the relation of faculties in language. UMG provide "… cognitive underpinnings of moral judgment."[150] Before trying to explain the relation between relational models and moral judgments, I will try to explain preonic interface of UMG. Preos determine

[149] Ibid, p. 51.

[150] Ibid, p. 57.

where and how those mods apply. For example, the order of occurrence in a meeting of a group structured by AR, or one person-one vote principle which is related to EM are semantic parts of models. They are interpretations of the mods.[151] Even though, whether all parts of preos learned is controversial, we can accept that it is mostly (even if not wholly) learned.[152] It might be the case that formal rules of semantics can be innate and the other parts related to interpretations of formal structure are learned. In this respect, i.e., their determination on the usage of formal structure, learning preos seems as a process similar to parameter setting. However, rules of the grammar of mods are not as explicit as rules of the grammar of language, yet; thus it might or might not be due to the same process. Since the study of RMT is a relatively new area, I assume that further investigations might provide an explicit account on how preos and hence relational models are set through their acquisition process.

It is clear that it is a kind of *universal grammar* but why universal *moral* grammar? The reason is its role in moral judgments. UMG does not propose a moral code, but provide the mental operations used in construction of moral codes.[153] All relational models are applied in moral principles, though they are not constituted by moral principles. For example, "[e]thical principles of equal treatment and compensation, uniform contributions, and even distribution are structured by EM."[154] Bolender summarizes the relation between UMG and moral principles as follows:

> Universal Moral Grammar (UMG), in the sense of the term that is of interest to cognitive scientists, consists of properties true of all non-pathological human minds. Rather than being itself a moral code, UMG consists of the cognitive mechanisms that mostly directly enter into the production of moral codes. Different basic relational mods are linked

[151] Ibid, p. 23 and p. 13.

[152] Ibid, p. 23.

[153] Ibid, p. 57.

[154] Ibid, p. 58.

> preonically to different moral expectations, evidently due to some neurological necessity.[155]

In understanding social cognition we have two sub-targets, at least for now. One is the explanation of formal properties of relational models and their relation to preos (i.e., UMG), and the other sub-target is the explanation of emergence of those formal properties. According to dual approach of Bolender we need different conceptions of these sub-targets and different approaches to be able to explain them. However, it is still a research program yet. If the project proposed by Bolender is pursued, I assume that we will reach a position similar to language acquisition phenomenon: different perspectives for the explanation of different sub-targets in understanding the phenomenon. Besides, we might be able to have an account for the acquisition of social cognition given by perspectival theories.

155 Ibid, p. 62.

CHAPTER 6:

CONCLUSIONS

I have tried to classify three main theories of cognitive science, i.e., CTM, PDP and DST, according to their system descriptions and the conception of cognition which they propose in their models. These theories of cognitive use one or the other of two conceptions of cognition (i.e., cognition as a componential system and cognition as an integrated system) while constructing their theories. CTM proposes a componential system in explaining cognitive phenomena. DST and PDP, on the other hand, propose integrated systems in their models of cognition. I have used several other classifications as tools for understanding the relations between different conceptions of cognition. I have observed that cognition and the phenomena related to it necessitate cooperation between two conceptions of cognition to be able to have a full account of cognition. The cooperation has leaded to a phase of cognitive science which I call perspectival theories from theories of levels. I also observe that next stage of cognitive science will be *syntheses*. As techniques and mathematical tools advance this stage can turn cognitive science to a unified theory in which cognition will have a unified theoretical perspective, or it might be the case that the stage of synthesis persists.

I have used the ontological categories which are derived from the practice of scientific inquiry by Wimsatt. These ontological categories are levels of organization, *perspectives*, and causal thickets. I have attempted to distinguish those ontological categories from the theories which have implications about ontological categories; theories of levels, perspectival theories, and *syntheses*. Objects and/or causal relations which are proposed by a theory belong to one of those ontological categories. These two related classifications seem to be applicable to cognitive science temporally, rather than as classifications among theories. I observe that cognitive science has a changing conception of its *object*, i.e., cognition, by the progress in theories. According to early conception of cognition, it was a higher-level *object* of levels of organization category. Lower-level explanations are regarded as explanations of micro-structure of the physical stuff, i.e., brain. The real cognitive theories were the theories about the higher-level phenomena of levels of organization. At this stage theories of cognitive science were theories of levels.

In the second phase which I assume that the current phase, theories of cognitive science belong to the perspectival theories category. The conception of cognition provided by them belongs to the *perspectives* category. In the current conception of cognition every theory has its own contribution in a distinguishable manner. Although they mostly intend to explain the same phenomenon, they focus on different sub-targets in the explanation of the phenomenon. Their focus is implicit because mostly their claims regard the whole phenomenon. For example, in language acquisition models of CTM, PDP and DST the target is understanding language acquisition, explicitly; not its sub-targets (i.e., language, pre-learning device and learning process by interaction with the environment). Only after refining the sub-targets and analyzing the theories closely we can see that they focus on different aspects of the phenomenon, i.e., sub-targets in understanding language acquisition. Proposing different conceptions for the phenomenon according to these different theories becomes reasonable, then. Explaining language acquisition as the change in the innate content is provided by the conception of cognition as componential system. Explaining language acquisition as the change in the structure of pre-learning device is provided by the conception of cognition as an integrated system. Finally, explaining language acquisition as the change in the attractor layout in which the pre-learning dispositions of individual and the relevant environment constitute a coupled system is provided by conception of cognition as mainly a process which implies an integrated structure. By refining the sub-targets and analyzing the contributions of the theories in explanation of those sub-targets, we can recognize that the phenomenon itself includes trans-perspectival problems. For example, explanation of innateness demands explanations of both the content of pre-learning device and the mechanism of pre-learning device itself. These explanations are provided from different perspectives; the content of innate device and the mechanism of it are explained through different decompositions of the system. This nature of the explanation of innateness makes it a trans-perspectival problem.

Our other example, social cognition, supports the view that the current conception of cognition belongs to *perspectives* category. Indeed, unlike the evolution of the conception of language acquisition, the theory about social cognition mentioned here presupposes the cooperation of different perspectives; which, I think, strongly supports the current position of cognitive science. In the theory, there are several trans-perspectival problems such as explanation of the emergence of basic relational models and usage of them. They are explained through systems which are decomposed according to different perspectives. Besides, the target phenomenon itself (i.e., construction of social relations in human minds by relational models) is posed as trans-perspectival and explanation of it demands cooperation of different system conceptions. From the beginning, Bolender proposes two research programs for different sub-targets in understanding social cognition which are the explanation of the emergence processes of basic relational models at the first stage and the construction of social relations by relational models at the second stage. The system in the theory is a componential system which works upon discrete items. However, the basic lexical items are produced by a dynamical process and they are used by the computational component by discrete operations. On the other hand one of the sub-systems of the system, i.e. preonic systems, is an integrated system and is able to affect the overall processes of the system by its continuous change. Using both of the conceptions of systems in the theory does not result in a theory which makes contradictory assumptions; rather, these two conceptions are complementary in the theory.

Finally, I assume that the evolution of our conception of cognition will lead us to conceptualize cognition as causal thickets and the evolution of the theories which explain cognition will result in theories as *syntheses*. Recently proposed hybrid theories seem to support my claim. The possible future of the cognitive science might be a unified theoretical perspective or it might be the case that cognitive science will be shaped by *syntheses*. I assume that advances in our theoretical perspectives, technical and mathematical tools will show.

BIBLIOGRAPHY

Abrahamsen, A. and Bechtel, W. (2006), "Phenomena and Mechanisms: Putting the Symbolic, Connectionist, and Dynamical Systems Debate in Broader Perspective", in Stainton, R. (ed.), *Contemporary Debates in Cognitive Science*, Malden, Mass.: Blackwell Pub., pp. 159-185.

Bechtel, W. and Richardson, R. C. (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton, N.J.: Princeton University Press.

Bolender, J. (2007), "Self-Organization in The Development Of Social Cognition: Symmetry Breaking and the Relational-Models Framework", *Psychologia*, Vol. 5, pp. 255-272.

Bolender, J. (2008), "Hints of Beauty in Social Cognition: Broken Symmetries in Mental Dynamics", *New Ideas in Psychology,* Vol. 26, pp. 1-22.

Bolender, J. (2010), *Digital Social Mind*, in press.

Braddon-Mitchell, D. and Jackson, F. (1996), *The Philosophy of Mind and Cognition*, Oxford, Mass.: Blackwell Pub.

Cartwright, N. (1989), *Nature's Capacities and their Measurement*, Oxford: Clarendon Press.; New York: Oxford University Press.

Chomsky, N. (2000), *The Architecture of Language*, Mukherji, N., Patnaik, B. N., Agnihotri, R. K. (eds.), New Delhi; Oxford: Oxford University Press.

Chomsky, N. (2001), *The Minimalist Program*, Cambridge, Mass.: The MIT Press, 1995 (fourth printing).

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. (1998) (Eds.), *Rethinking Innateness: A Connectionist Perspective on Development*, Cambridge, Mass.: MIT Press.

Fiske, A. P. (1993), *Structures of Social Life*, New York: Free Press; Toronto: Collier Macmillan Canada; New York: Maxwell Macmillan International.

Fodor, J. A. (1983), *Modularity of Mind*, Cambridge, Mass.: MIT Press.

Fodor, J. A. and Pylyshyn, Z. W. (1988), "Connectionism and Cognitive Architecture: A Critical Analysis", *Cognition,* Vol. 28, pp. 3-71.

van Geert, P. (2009), "Nonlinear Complex Dynamical Systems in Developmental Psychology", in Gaustello, S. J., Koopmans, M. and Pincus, D. (eds.), *Chaos and Complexity in Psychology: The Theory of Nonlinear Systems*, Cambridge, NY, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo, Delhi: Cambridge University Press, pp. 242-281.

van Gelder, T. (1998), "The Dynamical Hypothesis in Cognitive Science", *Behavioral and Brain Sciences,* Vol. 21, No. 5, pp. 615-665.

Hohenberger, A. and Peltzer-Karpf (2009), A., Language Learning from the Perspective of Nonlinear Dynamic Systems. Linguistics, doi: 10.1515/LING.2009.017

Hinzen, W. (2006), *Mind Design and Minimal Syntax*, Oxford; New York: Oxford University Press.

Kelso, J. A. S. (1995), *Dynamic Patterns: The Self-Organization of Brain and Behavior,* Cambridge, Mass.: MIT Press.

Lindner, K. and Hohenberger, A. (2009), Introduction: Concepts of Development, Learning, and Acquisition. Linguistics, doi: 10.1515/LING.2009.009

Pinker, S. (1995), *The Language Instinct*, NY: Harper Perennial.

Radford, A. (2004), *Minimalist Syntax*, Cambridge, UK; New York: Cambridge University Press.

Rumelhart, D. E., Hinton, G. E., and McClelland, J. L. (1989), "A General Framework for Parallel Distributed Processing", in Rumelhart, McClelland & the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, Cambridge, Mass., London: MIT Press, (ninth print), pp. 45-76.

Rumelhart, D. E and McClelland, J. L. (1989), "PDP Models and General Issues in Cognitive Science", in Rumelhart, McClelland & the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the*

Microstructure of Cognition. Volume 1: Foundations, Cambridge, Mass., London: MIT Press, (ninth print), pp. 110-146.

Rumelhart, D. E. and Zipser, D. (1989), "Feature Discovery by Competitive Learning", in Rumelhart, McClelland & the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, Cambridge, Mass., London: MIT Press, (ninth print), pp. 151-193.

Searle, J. (1992), *The Rediscovery of the Mind*, Cambridge, Mass.: MIT Press.

Smolensky, P. (1989), "Information Processing in Dynamical Systems: Foundations of Harmony Theory", in Rumelhart, McClelland & the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, Cambridge, Mass., London: MIT Press, (ninth print), pp. 194-281.

Smolensky, P., Mozer, M. C. and Rumelhart, D. E. (1996) (eds.), *Mathematical Perspectives on Neural Networks*, Mahwah, New Jersey: Lawrence Erlbaum Associates Pub.

Townsend, J. T. and Busemeyer, J. (1995), "Dynamic Representation of Decision-Making", in van Gelder & Port (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge: MIT Press, pp. 101-120.

Turing, A. M. (1950), "Computing Machinery and Intelligence", *Mind*, New Series, Vol. 59, No. 236, pp. 433-460.

Weisberg, M. (2007a). Who is a Modeler?. British Journal for the Philosophy of Science. doi:10.1093/bjps/axm011.

Weisberg, M. (2007b), "Three Kinds of Idealization", *The Journal of Philosophy*, Vol. CIV, No. 12, pp. 639-659.

Wimsatt, W. C. (2007), *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*, Cambridge, Mass.: Harvard University Press.

APPENDIX:

X-BAR ARCHITECTURE

The components of symbolic system which I have tried to symbolize as 'f_{syntax} (lexicon)' are primes (atomic elements) and symbolic objects which are generated from primes by finite procedures. These symbolic objects are phrases (XP). X is the head of a phrase, XP. A phrase is about a situation, object, action, pre- or postposition, etc. and head of it governs the other components of the phrase. A phrase can contain another phrase in it. For example the determiner phrase, *a picture of a flower*, has *a* as its head followed by a noun phrase, *picture of a flower*, which has *picture* as its head. The noun phrase is about *picture*, not about *flower*. Thus, *picture* governs *of a flower* in the phrase, i.e., it is the head of the phrase. Accordingly, head of a verb phrase is the verb in it and the head of a pre- or postposition is the pre- or postposition in it, respectively.[156] Since being a picture implies being a picture *of something*, *of a flower* plays a role in the phrase; in other words, it is the role-player of the phrase. In the noun phrase, for example, *picture of a flower in a frame*, *in a frame* is the modifier. Modifiers give additional information about the head which are not as essential as role-players. Subjects of phrases, the causal agents, are optional and called 'specifier' (SPEC). When the phrase denotes an action, usually it takes specifier.[157] For example *the girl likes pictures of flowers in a frame* is a verb phrase which is followed by a noun phrase. In it, *the girl* is the subject who *likes pictures of flowers*. The quotation from Pinker illustrates the X-bar structure which is the structure of a phrase as follows:

$$XP \rightarrow (SPEC)\ \bar{X}\ YP^*$$

[156] Chomsky (1995), pp. 34-35.

[157] Pinker, pp. 99-110.

> "A phrase consists of an optional subject, followed by an X-bar, followed by any number of modifiers."
>
> $\bar{X} \rightarrow X\ ZP^*$
>
> "An X-bar consists of a head word, followed by any number of role-players."[158]

UG holds these relations for all possible I-languages by inclusion of 'parameters'. "The piece of information that makes one language different from another is called a parameter."[159] For example in the X-bar architecture the universal law will be as follows:

> $\bar{X} \rightarrow \{ZP^*, X\}$
>
> "An X-bar is composed of a head X and any number of role-players, in either order."[160]

The parameter in this super-rule (a rule of UG) is the order of the components. English, for example, is a head-first language and Turkish is a head-last language. The head of the phrase *picture of flowers* is *picture* in English. The head of the phrase, *çiçeklerin resmi*, is *resim* in Turkish which follows its role-player, viz. *çiçeklerin*.[161]

Phrases can contain other phrases recursively as in the example *the girl likes pictures of flowers* which is a verb phrase with a noun phrase embedded in it. These objects constitute sentences and sentences themselves can enter other sentences recursively.

---

[158] Pinker, p. 110.

[159] Pinker, p. 111.

[160] Pinker, p. 111.

[161] *Resim* means picture, *çiçekler* means flowers and *-in* is the affix which means *of* in the phrase, viz., *of flowers*.

This procedure is called merge – "… a technical term meaning 'combining' …" [162] Selection, on the other hand, provides lexical items for the use of merge. The selection process of a lexical item must take LF and PF, namely, interface subsystems, into account. A chosen word from lexicon corresponds to an item from LF and an item from PF, and these two items should be compatible; in other words, "… it is not the case that any sound can mean anything."[163] Merge and selection are the process descriptions of a grammar which are used to construct phrases and sentences. They are the means of producing infinitely many structural descriptions from finite input.

[162] Radford, p. 66.

[163] Chomsky (1995), p. 225.