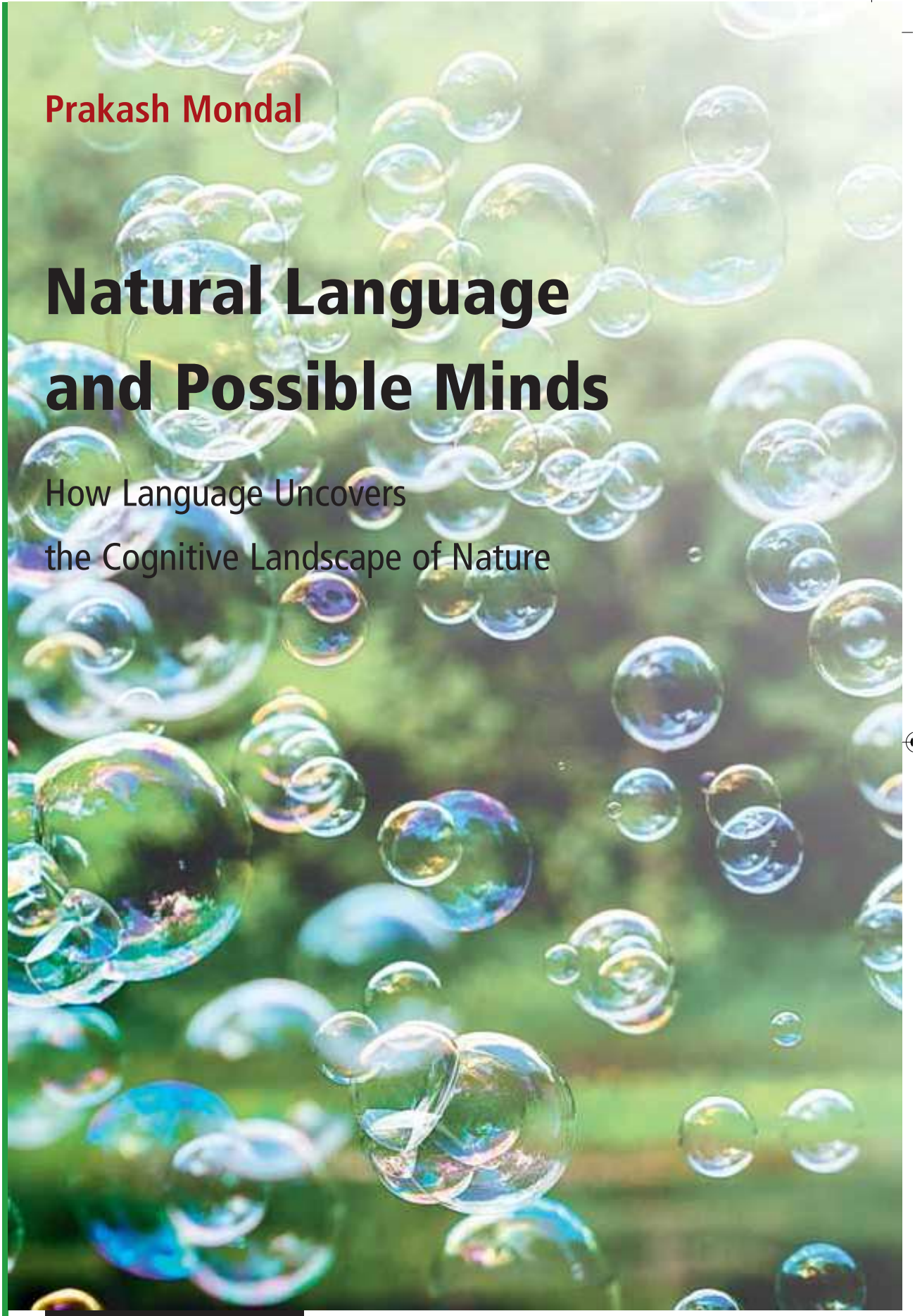
Prakash Mondal

Natural Language and Possible Minds

How Language Uncovers the Cognitive Landscape of Nature



VALUE INQUIRY BOOK SERIES

BRILL | RODOPI

Natural Language and Possible Minds

Value Inquiry Book Series

Founding Editor

Robert Ginsberg

Executive Editor

Leonidas Donskis

Managing Editor

J.D. Mininger

VOLUME 303

Cognitive Science

Edited by

Francesc Forn i Argimon

The titles published in this series are listed at *brill.com/cosc*

Natural Language and Possible Minds

How Language Uncovers the Cognitive Landscape of Nature

By

Prakash Mondal

BRILL
RODOPI

LEIDEN | BOSTON

Cover illustration: The cover illustration is used with permission from Pixabay.

The Library of Congress Cataloging-in-Publication Data is available online at http://catalog.loc.gov
LC record available at http://lccn.loc.gov/

Typeface for the Latin, Greek, and Cyrillic scripts: "Brill". See and download: brill.com/brill-typeface.

ISSN 0929-8436
ISBN 978-90-04-31665-2 (paperback)
ISBN 978-90-04-34420-4 (e-book)

This book is printed on acid-free paper and produced in a sustainable manner.

To

Mahamaya, my divine mother

∴

Contents

Preface

This book is part of an attempt to understand the nature and form of mentality in a vaster range of beings than is currently acknowledged in mainstream thinking in cognitive science. While the nature of human mentality is not under suspicion, when it comes to the exploration into the nature of other kinds or types of mentality people often raise eyebrows. Set against this backdrop, this book tries to trace the very foundations of minds and, in doing so, seeks to project a vista within which a tapestry of different types of minds can be delineated. Perhaps the only way the whole project here differs from other avenues of thinking in certain quarters of cognitive biology and cognitive semiotics is that this book attempts to project a picture of distinct types of mentality across organisms by drawing up a formalism extracted from natural language within which the abstract components of mentality in biological substance-independent terms can be described. This formalism is thus descriptive rather than explanatory. The formalism tries to grapple with the problem of describing the structural forms of possible types of mentality across the organismic spectrum. This is necessitated by consideration of the point that predicting or determining what other non-human organisms (can) do on the basis of what they have in their inner realms is still beyond our grasp. This is faintly understood because we do not yet get a handle on the question of whether other non-human species have anything approximating to mentality. Hence a more fruitful way of approaching the forms of possible types of mentality across the spectrum of various organisms or species is to first settle the descriptive problem of expressing what there exists that can be individuated inside the inner realms of non-human beings. With this goal set, the book undertakes to understand possible minds from natural language. Rather than injecting an anthropomorphic bias into an otherwise exploratory account, natural language turns out to carry certain advantages that have never been harnessed. Additionally, it emerges that this has surprising consequences for a description of the type of mentality we can attribute to machines. The nature of computation vis-a-vis natural language is also explored in this connection. Overall, this lends credence to the idea that natural language has the limitless potential for the task of unraveling many as yet unresolved riddles and puzzles surrounding language, minds and computation.

Nevertheless, I also think there is a danger inherent in any attempt at uncovering the mental world of the other—an issue which deserves careful consideration as well as caution. But as the discussion proceeds, I've tried to show that some, if not all, of the concerns can be adequately addressed if the sources

of the confusion and vagueness are identified. The problems are not simply theoretical; they are empirical too. In fact, one of the underlying assumptions adopted in the current work is that the cognitive landscape of nature is far more variegated and kaleidoscopic than we may ever conceive in the safest corners of our chambers of intellectual inquiry. No matter how the details of the proposal presented here unfold, one thing that seems clear is that even a description of distinct types of possible mentalities grounded in the natural world will require a stupendous amalgamation and aggregation of facts, ideas and insights drawn from various modes, fields and aspects of intellectual inquiry. This book is a mere fragment of this rather daunting venture.

Finally, I invite readers of all stripes to explore the book and decide for themselves what they need to seriously understand about minds in nature and the relation of natural language to minds.

05 January, 2017 P.M.
Hyderabad

CHAPTER 1

Introduction

Natural language has an intimate connection to the nature of minds. Ideas, thoughts, feelings and other nebulous entities attributed to minds are often expressed and represented in natural language. One also comes closer to understanding what the other person entertains in the mind by using natural language. It may now become clearer that natural language is here taken to be human language. But the link between natural language and the nature of minds still possesses an inchoate character because the precise nature of the relationship between natural language and minds is still obscure, and thus has to be explicitly articulated. To be clearer, by saying that natural language has an intimate connection to the nature of minds we do not mean to simply assert that the use of natural language in actual circumstances helps discern and decipher the intricate patterns of human thoughts, intentions, plans, mental strategies, emotions etc. The present context in fact demands much more than this, although it is tempting to stick to this kind of behaviorist exploitation of natural language for explorations into the mental territory. The complex system of linguistic structures manifest in any natural language unpacks a host of assemblies of mental organization that reliably correspond to patterns of linguistic structures. In other words, a range of systematic patterns of linguistic structures reveals an ensemble of 'mental structures' which can be taken to be the representational properties of the underlying mental organization. For instance, in English the sentence 'He pushed the bottle towards her' expresses a mental assembly of conceptual relations which is to be distinguished from that expressed by the sentence 'He crushed the bottle for her'. Thus, a variety of linguistic structures express a variety of mental structures that can be traced to the internal representational and/or encoding machinery of the mind. In this sense, variations in the representational properties of the underlying mental organization can be correlated with variations in the forms of minds. At this juncture, it appears that the notion of forms of minds must be such that it helps track variations in representations of thoughts in across diverse sections of humans. This is indeed the case, insofar as the focus is restricted to the members of *Homo Sapiens*. But the present book raises the question of whether possible forms of minds across a range of non-human organisms and/or systems can be tracked by marking patterns of variations in a vaster gamut of mental structures extracted and projected from natural language. As far as the whole enterprise of the inquiry undertaken by the present book goes, the answer provided here is resoundingly yes.

As the current level of inquiry into the nature of the mind, as it stands, penetrates into the human cognitive machinery, it is striking that the structure of the human mind is often explored by looking into what natural language reveals about our cognition. This appears to induce an anthropomorphic bias when we go about figuring out what other non-human types of mentalities look like. The reason this is so is that the mental structures revealed by examining natural language(s) are considered to be intrinsically human, given that the structures of human language, as opposed to the expressions of formal languages, are the sources of significant insights into the nature of our minds. This is the cornerstone of this line of inquiry. The current work proposes to turn around the idea behind this inquiry by showing that the humanness of natural language does not introduce an anthropomorphic bias. Far from it, linguistic constructions and phenomena can actually tell us a lot about the structure and form of other possible types of mentality in other non-human creatures and plausibly even in some plants. One reasonable way of approaching this is to incorporate the idea that the mental structures that can be tapped through natural language can yield the mental structures that cannot possibly be tapped through natural language once the former is deducted from the collection of all possible mental structures which is independently postulated on the basis of cogent generalizations. Crucially, this book aims to demonstrate that there is nothing in the existing hypotheses and theories on the relationship between natural language and cognition that prevents mental structures from being realized in other non-human organisms or creatures and plausibly in some plants. Before we proceed further to see what the book has to offer in ways of understanding the character of mentalities through natural language, some of the basic concepts employed throughout the book need to be clarified.

1.1 On Minds and Mental Structures

So far the discussion has conveyed the impression that the notion of minds deployed in the present context is special. But one may now wonder in what substantial way it is actually special. We know for sure that minds have a special status in the realm of biological entities. But a notion of minds that is flexible enough to be tailored to diverse biological contexts of various organisms and species is, of course, desirable as mental structures are postulated as the cognitive ingredients of mentality in a general sense. For one thing, all that matters when one speaks about minds is whether it allows for invisible capacities that are regarded as the central facets of any mind, whereas what we require here is

a handy conception of minds that does not simply attach to any unique form of the biological substrate but rather links to many realized forms of substance found in the biological space. It may be noted that not much over and above mental capacities and/or processes needs to be considered when speaking of minds just in the same way as not much over and above biological processes needs to be taken into account when talking about life. In this sense, minds are ways of speaking of certain entities and processes just like life is a way of speaking of biological processes (see Mayr 1982). Plainly, minds are not physical objects made up of certain types of substance. Therefore, from a certain perspective, the contention that the boundaries of life are exactly the boundaries of minds may seem justified. On this view, one has to look no further than the domain of living entities to delimit the possibilities of minds. But it carries with it the presupposition that the living world reliably coincides with the world of properties we usually or naturally ascribe to minds. The problem is that such a view, sensible though it may seem, inherits a weak form of anthropomorphism, for the properties we usually ascribe to minds are to be discerned from our own case that we usually or naturally observe and theorize about. It is also worthwhile to note that simply stating that the boundaries of life are identical to the boundaries of minds trivializes the very notion of what minds really are. This is so because minds are then plainly and somewhat grossly equated with emergent forms of life.

This issue deserves a bit of elucidation here. On the one hand, we cease to get a handle on how to demarcate minds in such way that minds are recognizably specified with respect to multiple forms of biological substance found in the realm of life, for if all life forms have minds, all life forms have an equally consequential stake in the business of mentality. On the other hand, this notion is not fine-grained enough to tell birds apart from plants, for example, or for that matter, humans apart from other primates. This seems to indicate that there is something over and above the mere demarcation of minds that goes beyond an identification of the boundaries of minds. Rather, this suggests that demarcation is not sufficient for the characterization of what kinds of things minds are. But note that this cannot be taken to imply that the idea that the boundaries of life are the boundaries of minds is inherently misguided. As the arguments in this book unfold, we shall see that this idea forms the background scaffolding of what is to be developed in the succeeding chapters of the book. That is, this idea can constitute the background assumption so that this can be fine-tuned, suitably modified and sharpened further to yield the desired concept of what minds are.

For all we know about minds, it is apparent that they cannot be simply identified with the body or any part of the body including the nervous architecture

or with anything akin to the nervous architecture having similar functions. This means that it would be too simplistic to reduce minds to bodies and brains or any such forms of biological substance because this invites the problem of Cartesian dualism that consists in positing two distinct ontological domains for the mental stuff and the biological substrate. A view of dualism may carry with it the danger that one may view the matter with a reasonable amount of skepticism in thinking that an anthropomorphic bias is imperceptibly being passed on to the characterization of what minds are. The reason this may appear to be so is that only those cognitive abilities and capacities that are usually championed or regarded as great achievements in the mental world of humans seem to be disengaged from the biological substance. Reasoning, thinking or cognizing and so on are the exalted candidates under this category, whereas eating, smelling, seeing, feeling etc. are usually not thought to fall under this category. This is largely rooted in the Cartesian bifurcation between (human) minds and non-minds either in inert objects or in non-humans. In any case, what seems clear is that the case for substance-independence of minds has to appeal not to a trivial notion of independence from any substance whatsoever, but rather to *potential* independence from different forms or types of substance which minds can be linked to. In other words, it would be wrong to simply say that minds are substance-independent to the extent that it inherits the dualistic segregation of minds and biological substance, thereby nullifying the case for distinct types of mentality in non-human organisms and creatures simply because mental types cannot, then, be said to vary depending appropriately on the type of biological substance chosen. Hence minds can be said to be substance-independent only insofar as the postulated independence is taken to be a kind of *non-unique dependence*, but not a general across-the-board type of independence. That is to say that the required notion of independence is cashed out positively in terms of a kind of dependence which is actually non-unique in nature. This leads to the view that minds cannot be linked to any *single* substance; rather, they can be linked to multiple forms of substance. Thus, mind's substance-independence is a kind of *logical* or potential independence but not a kind of ontological independence that requires non-trivial metaphysical commitments.

Now at this stage, it may seem that this is just another functionalist argument from *multiple realizability* which consists in the claim that a higher level (cognitive) function can be realized in multiple forms of substance. But this is mistaken for various reasons. First, multiple realizability in its essence derives from the computationalist view that a certain computable function can be realized in a number of hardware systems, thereby supporting a one-to-many mapping from computable functions to tokens of hardware (see Fodor 1975;

Pylyshyn 1984; but see Putnam 1988, Polger and Shapiro 2016 for a sustained critique). In simpler terms, this springs from the idea that a software system can be realized in many different types of hardware. A view of non-unique dependence, albeit apparently compatible with multiple realizability, differs from multiple realizability, in that multiple realizability does not really 'care' about the type of hardware chosen, while non-unique dependence is a type of dependence but it does not depend only on one single substrate. To put it in another way, a given type of substance matters for non-unique dependence precisely because it is a case of dependence in each particular instance or condition of something depending on a given substance, whereas multiple realizability is *not* a case in which individual instances of realization in various types (or even tokens) of hardware are instances of dependence since the relevant computable function need not be realized anyway. That is, mind's non-unique dependence on diverse types of substance demands that minds become anchored to *at least* two distinct types of substance because only then can we state that minds are not linked to any single substance. But the same cannot hold for multiple realizability, crucially because it is not even necessary that computable functions are realized in any single substance. Second, it may be noted that mind's substance-independence is a kind of logical or potential independence, as stated above, whereas multiple realizability cannot be simply a case of logical or potential independence—it is more than that. Since multiple realizability may obtain even when no function is actually realized in a substance, multiple realizability may be a case of ontological independence plus logical or potential independence. This distinguishes multiple realizability from non-unique dependence in a striking fashion as multiple realizability is a super-order concept a part of which is shared by non-unique dependence. This is because everything that is a case of logical or potential independence may not automatically be a case of multiple realizability (a non-computable function which is logically disengaged from any realizing physical system cannot be said to multiply realizable, and so is true of something like the largest prime number). Third, the underlying basis of multiple realizability is the very concept of realizability which can hold in many forms of substance, but the essence of non-unique dependence lies in the notion of dependence which warrants a relation logically distinct from that of realization. If A depends on B and B depends on C, we can always infer that A depends on C, but if A is realized in B and then B is realized in C, it does not follow that A is realized in C. Thus, for example, if a time calculation algorithm is realized in a digital wall clock which is in turn realized (in the sense of being embedded) in the concrete structure of a wall, it would be absurd to say that the time calculation algorithm is also realized in the wall.

Having clarified the notion of substance-independence we have targeted in the present context of the discussion on the relation between mind and its substrate, we may feel that the characterization of what kinds of entities minds are is achieved. As a matter of fact, this is not fully right. Specifying the relation between mind and its substrate does not simply amount to specifying what minds are any more than specifying the relation between a given equation and its realization in the mind or even on a piece of paper amounts to specifying what that equation really is in its fundamental nature. In fact, we require something more than this. But we may wonder whether minds are really the entities that can be characterized the way numbers, for example, are characterized. After all, it is quite plausible that minds conceived in a more general sense which is couched in broader terms do not fall within the phenomenal limits of the organismic envelope. On the one hand, the existence of minds within a generally broad organismic envelope is neither entailed nor made viable by considerations of intentionality, as expanded on in Section 1.3, or of consciousness and various kinds of cognitive processing. In particular, this is not guaranteed by consciousness in that consciousness is not a simple unitary phenomenon whose character can be reliably and appropriately fractionalized such that the separable parts can be in distinct combinations mapped on to distinct types of possible minds. Moreover, a general characterization of consciousness which is demarcated independently of specific biological substances and yet applicable to many organisms across the organismic spectrum is hard to come by. Thus, for example, sensory consciousness which is characterized by different degrees of integration of sensory features and qualities of the world through increasing hierarchical layers of abstraction can be candidate for a general level of consciousness (see Feinberg and Mallatt 2016). But the problem here is that it unjustifiably excludes living entities (such as plants) that may not have any kind of sensory integration through hierarchically organized neural structures. Thus, an appeal to consciousness to help build demarcations within the space of possible types of mentality would end up being too restrictive. Besides, the existence of minds is not equally made necessary by the individuation of the instantiation relation between mind and its substrate, for no instantiation relation between minds and the kinds of substance they are instantiated in can unequivocally determine the boundaries of (possible) minds. This is the case by virtue of the fact that the required instantiation relation between minds and various kinds of substance may be many-to-many. We cannot *prima facie* stipulate that the human biological substance can instantiate only its own type of mentality reserved only for *Homo Sapiens*—it is plausible that many other types of mentality are themselves embedded within what we may recognize as the human type of mentality, however characterized.

But then one may believe that minds conceived in a general sense—the sense which is warranted in the present context—can be approached by stripping the human type of mentality, as we understand it, of all its intellectualist attributes and then appealing to those aspects of interactions with the environment that are bottom-level or low-level processes to be postulated as characteristic of minds. The best candidate that comes closer to fulfill this goal is the basic perceptual process which seems to capture much of the low-level territory of mind's operations and processes. Since all organisms and plants interact with the outer environment and act upon certain properties, features and resources of the environment they are in constant touch with, it is the perceptual configuration in various creatures with respect to the perceptible or perceived world that appears to project a general version of mentality which can range over an ensemble of organisms simple or complex. That the perceptual mind seems to be the basic format of minds can also be traced to the view that non-human organisms live in a world of immediate perception beyond which their world ceases to exist (see Dummett 1993b, p. 123, for example). Now regardless of whether or not perceptual abilities in various organisms are attuned to the immediate world, it seems clear that perceptual abilities do not immediately lend themselves to being molded into the basic texture of minds. The reason is that what kinds of systems minds are cannot be entailed by the detection of perceptual abilities. Consider, for instance, the interaction of two magnets whose opposite sides attract but whose same sides repel each other. On the surface of it, a minimal form of account that does not presuppose the understanding of physics or behavior of magnets as physical objects may bestow perceptual capacities on magnets. But this is, of course, nonsensical. The problem emanates from the behavioristic criteria attaching to the way perceptual capacities are recognized. That is, it is only by looking at the behavior of a certain creature or even an inert object, one may surmise that the creature or the object concerned has perceptual capacities. But as soon as we recognize that behavior does not automatically underwrite the mechanisms or structures lying within, it is hard to see how one can drive home a general conception of minds by taking perceptual capacities to be the basic format.

One may attempt to resist this conclusion by insisting that we already know for sure that a creature is a living entity, while an inert object is a non-living entity. This argument misses the point altogether, for it is not the question of whether one has the required familiarity with the entity in question—rather, it is the question of whether one can really read back from mere observation of perceptual abilities to the recognition of a mind-like system. In fact, in many cases we may not even possess a rudimentary form of knowledge of or familiarity with even many living entities, whether they are plants or microbial

organisms, and then we may treat them as inert objects rather than living entities by reading much into apparent observations. Our familiarity with the taxonomy of living and non-living entities cannot dictate what kinds of things minds really are. The point raised here does not, however, imply that we cannot understand anything at all about the structure of minds by working back from perceptual abilities to the system within that generates the behaviors which can be predicated on the perceptual abilities concerned. As a matter of fact, we can capture insightful glimpses into the structure of minds by appreciating the significance of non-perceptual contents of minds as they interface with perceptual capacities across organisms. This brings us to a point raised by Bermúdez (2003) who thinks that an otherwise justified restriction to perceptual capacities leaves out of consideration many non-perceptual processes and their contents which may well exist in many creatures, and possibly in other smaller organisms. The caching behavior of scrub-jays, the courtship behavior of European starlings, passerine birds' behavior of bringing food to the eggs in the nest in anticipation, the nest building behavior of many birds in anticipation of eggs, tool uses in different primates such as chimpanzees, the hiding behavior in cats etc. and also many relevant behaviors of various creatures in unfamiliar situations illustrate cases where the actions in question are not simply attached to the immediate sensory-perceptual environment. Although it is reasonable to believe that non-perceptual capacities entail perceptual capacities, but not vice versa, and thus it is safer to postulate perceptual capacities for minds because they constitute the broader category of mental capacities and abilities, this is problematic for two reasons.

First, just because perceptual capacities constitute a broader category than non-perceptual capacities, this *minimalist* orientation does not *by fiat* gain a purchase on the character of minds. The minimalist orientation turns out to be nugatory, on the grounds that it is not fine-grained enough to motivate subtler distinctions among different types of mentality as it grossly brings together all organisms or species under its ambit. This suffers from the same defect that was pointed out above for the view that the boundaries of minds are the boundaries of life despite its significance on certain other grounds. The same consideration applies to any appeal made to Morgan's Canon that bans ascriptions of higher mental capacities to a behavior that can simply be interpreted as the outcome of some lower mental capacity (Morgan 1894).

Second, when it is the case that perceptual capacities are entailed by non-perceptual capacities as the former forms a broader kind, it is also the case that non-perceptual capacities are not entailed by perceptual capacities. Now this means that we cannot *necessarily* infer the existence of non-perceptual capacities from the presence of perceptual capacities. That is to say that perceptual

capacities may *imply* the presence of non-perceptual capacities. And if this is so, what justifies the restriction to only perceptual capacities for the conceptualization of minds in a general sense? After all, many possible minds can have non-perceptual capacities along with perceptual capacities. So, for example, even if one hundred organisms out of a thousand do not possess any non-perceptual capacities but have perceptual capacities, we need to have an account of those nine hundred organisms whose mentalities are constituted by both non-perceptual capacities and perceptual capacities. Surely this cannot be a matter of quantitative weights that can be read off from the statement that the existence of non-perceptual capacities cannot necessarily be inferred from the presence of perceptual capacities. Hence the logical relation between non-perceptual capacities and perceptual capacities cannot be cashed out, at least in a straightforward way, in terms of biologically significant relations among the mental capacities across species.

It is noteworthy that Bermúdez (2003) has gone on to offer an account of the non-perceptual contents in non-human creatures by proposing a non-linguistic (or simply non-propositional) version of semantics called *success semantics* aimed at recasting beliefs and desires inherent in goals in terms of certain external conditions. According to him, beliefs have as their contents utility conditions—conditions, or rather states of affairs that make a belief true, and desires have satisfaction conditions which are construed as states of affairs that match desires to actions, thereby terminating the desire. While this may well wedge non-perceptual and yet non-propositional contents into the mental machinery conceived in a non-human way, this transfers the burden of non-perceptual contents to the external world and borders on a behavioristic way of individuating contents. The present proposal aims to approach this in a quite different way. Given the vagaries in articulating what minds really are, it is far more appropriate to pose the question in a way that is tailored to meet the requirements of the present inquiry without running into the sort of problems delineated above. That is, instead of asking what minds really are, we may now ask what constitutes what we recognize as minds conceived in the customary general sense of the term. Or simply, what constitutes mentality? The answer the present proposal aims to advance is that it is mental structures that constitute the basic texture of what we may usually discern as mentality. Mental structures are the ingredients of mental types, but not of mental tokens. In fact, mental structures can suitably replace that which we ascribe to organisms as possessing as part of the resources that allow them to perceive, represent, interact with, or act upon the world.

Now it may appear that this formulation is circular because minds or mentality is characterized or defined on the basis of some structures which are true

or characteristic of minds. This apparent circularity is dissolved once we realize that the formulation above is not so much a definition or a reduction as a *substantive* (re-)description that is indicative of the fabric of mentality. That is to say that even if we substitute some other term, say, *m-structures* or even *x-structures* for the phrase 'mental structures', there is no loss of the substantive sense attaching to the phrase 'mental structures'. Thus, the formulation above does not ride on a linguistic reduction or characterization of terms like 'minds' or 'mentality'. More will be said on this later on, and this issue will be further explicated in Chapter 3. Suffice it to say for now that what is important about mental structures is that they have two different dimensions or modes. One is that they underlie linguistic expressions, and the other is that they are not *just* representations or reified structures that abstract away from the biological substrate. They may be grounded as internal states either in a nervous architecture or within the bodily system of an organism as a whole.[1] In this respect, mental structures are not to be aligned with basic perceptual processes if conceptualized in line with what Burge (2010) appears to think. He takes perception to be a quasi-algorithmic process that does not have a representational character, and hence it can be said to be implemented in non-humans and also infants who may lack representational or higher-order cognitive capacities. However, the problem, as one may see with his views on perception, is that he considers it to be largely modular, which does not comport well with the present view of mental structures which may be embedded within the biological constitution of a species, or simply, within the internal states of the body as a whole.

For similar yet slightly distinct reasons, Gauker's (2011) view that many non-human creatures can have imagistic representations is also not ripe for the development of species-general structures individuating types of mentality. Imagistic representations, just like perceptual representations, do not allow for unique or partial decompositions the way mental structures can by virtue of the fact that they will have the logical structure of *relations* in its mathematical sense (as will be formalized in Chapter 3). The image of a tree, for example, mentally organized cannot be logically linked to the part of the image for the trunk—only whole images rather than parts of them count. This prevents imagistic representations from being subject to partial *virtual* manipulations for reassembly, re-integration, and creation of new combinations from parts. Quite aside from that, imagistic representations cannot be postulated for organisms such as plants that do not have any perceptual apparatus configured in terms of sensory-motor organs. This does not, however, impose any ban on

1 But see Chapter 4 for a slightly distinct way of implementing mental structures in connection with the relationship between mental structures and machine cognition.

imagistic representations being linked to or fed into mental structures, or vice versa, especially for organisms that possess the perceptual apparatus. It is reasonable to think that perceptual or imagistic representations can often index and shape mental structures when certain perceptual or imagistic representations are evoked more than once for the recognition of objects and features via re-identification and re-extraction. Thus, for example, when the features of certain food items in an organism's environment are perceived and gradually evoked over and over again for the *reification* of a sign-like form linked to the set of food items naturally found, the relevant perceptual or imagistic representations can give rise to and thereby shape mental structures that are finedgrained to determine if some arbitrary item is the same as some member of the set of items found, or is of the same category. But this cannot be taken to imply that mental structures are *in themselves* perceptual or imagistic representations, for mental structures do not directly interface with the actual world. In this connection, it is noteworthy that perceptual or imagistic representations regarded as signs that (may) map themselves onto goals or actions in the form of responses these signs stand for enter into a causal or semiotic relation that directly transforms something extracted from outside into responses/actions which in turn conserve such causal or semiotic relations over many instances of events. Once perceptual or imagistic representations assume sign-like forms, it is mental structures that determine which sign relations are to be deployed in each given situation, thereby paving the way for the emergence of new and novel sign relations in unfamiliar settings and situations. This is made viable by the open-ended form of mental structures which, by virtue of not projecting onto the world directly, can link to multiple sign relations between inner needs and actions within and across organisms. Thus, there is nothing that can stop a certain sign relation that obtains in food gathering, for example, from being employed for hunting or even playing. This has ramifications that will be developed as we proceed to formulate mental structures for kinds of organisms in Chapter 3.

From another perspective, the exact relationship between mental structures as specified above and linguistic expressions deserves elucidation. When we state that mental structures are interpreted structures that underlie linguistic expressions, the underlying idea is that mental structures can be revealed by examining linguistic structures. That is, linguistic structures serve to disclose mental structures that are structures having no meaning in themselves and are pre-interpreted within the contextual constitution of the exercise of various capacities and actions of organisms. To give an example, the mental structure that can be uncovered from the sentence 'He danced with her but never sang for her', for instance, can be said to be pre-interpreted in the sense that it is an

abstract structure which is internally accommodated and assimilated by the encoding mechanisms of neural networks and bodily processes that engage in the relevant actions associated with the mental structure at hand (in this case, dancing and singing). In other words, the relevant mental structure must have to be assimilated and integrated into system of neural and other bodily processes (responsible for motor, proprioceptive, kinesthetic interactions) for later evocation, deployment and iterative reuse. Note that mental structures taken in this sense can be encoded representations or embodied structures or both at the same time. This means that the two dimensions or modes of mental structures correspond to two different scales—the link to linguistic expressions forms the abstract higher-order scale (which *we* figure out by applying our meta-cognitive abilities) while mental structures as internal states become part of the scale of physiological configurations.

Beyond that, it is crucial to understand that mental structures are not determined by linguistic structures, and in being so, they do not stand in a relation that can be cashed out in terms of an enabling or mirroring relation. Rather, linguistic structures bear certain logical relations which evince mental structures. This can be taken to mean that linguistic structures do not enable mental structures, in that enabling is ultimately a weakened or diluted causal relation and hence it inherits relations of a causal chain which does not harmonize with the character of mental structures. Mental structures cannot be either caused or enabled by linguistic structures because they can stand alone independently of linguistic structures. Likewise, mental structures cannot be mirrored by linguistic structures because mirroring demands a kind of isomorphism between the object that is mirrored and the image itself which has to be preserved over all transformations of the mirroring object, that is, over transformations linguistic structures may undergo.[2] This does not hold true for mental structures since a transformation of linguistic structures may alter the mapping to the mental structures concerned. Moreover, logically speaking, mental structures may be or may not be compositional relations, and hence they cannot be simply predicative. While all predicative relations (e.g. F(x) when x is true of F) are compositional as the composition of predicates gives

2 This point makes reference to the way mirroring physically works. In mirroring, no matter what the size of the image on the mirror is (that is, magnified or reduced), there must be an isomorphism between any points in the mirror image and the corresponding points in the actual object mirrored. Now if the mirror is tilted or placed in another angle with respect to the object mirrored, the isomorphism has to be preserved. This obtains even for the equivalence between the distance from the mirror surface to the object mirrored and that between the mirror image (which *virtually* stands behind the mirror surface) and the mirror itself.

rise to a relation which can be traced to the way the given predicates have been syntactically combined, mental structures may *not always* be characterized this way and be non-compositional as well because mental structures may contain elements which cannot be syntactically combined. This distinguishes mental structures from linguistic structures which are generally compositional as far as syntax goes. This will be elaborated on in Chapter 3 when a formalization of mental structures is presented. Also, more will be said on the linguistic relation to mental structures in Section 1.3. With this in place, we are now geared up to offer some remarks on the legitimacy of, and justification for, the methodology of the current work.

1.2 A Note on the Methodology

It is vital to understand that the present study is a modest attempt to unravel the intricately knitted complex that minds are by having them decomposed into their basic structural forms. The proposal to be advanced is that (possible) mentalities are, *at least in part*, structurally constituted by mental structures which can be uncovered from linguistic structures. Note that this does not in itself ban any investigations and explorations into the psychological procedures and mechanisms that may be postulated as part of the machinery of minds as well. Insofar as this is so, any experimental and/or comparative ethological studies on various non-human organisms and creatures may serve to complement the development of the formalism of mental structures in the present context. That is to say that any experimental or ethological studies on other non-human organisms and creatures may discover more about the variations in psychological procedures and mechanisms realized in non-human species, and can in turn feed these insights into the present framework for their assimilation into the formalism of mental structures. In this respect, one important advantage of the present study is that it employs a descriptive apparatus which is to be framed by way of the articulation of a logical formalism of mental structures that can easily accommodate experimental and ethological studies. This is so because the formalism to be developed is supposed to be neutral with respect to its extrapolation to experimental and ethological findings. Thus, the formalism of mental structures in later chapters, especially in Chapter 3, will be deployed to see how a range of mental structures fits into experimental and ethological findings on cognitive capacities across species. Plus the descriptive formalism of mental structures will also be employed to figure out what can be said about the character of a type of mentality that can be attributed to computing machines (in Chapter 4). Surely this part of the exploration into

non-human types of mentality when considering machines in particular cannot simply be a matter of experimental investigations because no amount of study of machines can decide either in favor of or against the contention that machines have minds. This point will be taken up in Section 1.3. In this respect, the descriptive formalism of mental structures will have an edge over other competing proposals, in that it makes no claim as to whether experimental investigations into machine computations can reveal the nature of machines' type of mentality. This will be further clarified in Chapter 4.

In a nutshell, the present study will apply a top-down approach in tackling the problem of finding out other types of mentality in non-humans. That is, it will first attempt to solve the problem of *description* of other possible minds in non-humans and then get down to understanding how this can be squared with experimental and ethological explorations into non-human organisms' cognitive abilities and capacities. We are aware that experimental and ethological studies have been the conventional type of studies in understanding non-human creatures' behaviors and cognitive abilities. Significant as this approach is, this cannot adequately address the question of how to describe various types of minds other than the human kind. The problem is much more severe than is commonly recognized since no amount of experimental and ethological studies can unequivocally demonstrate that other non-humans have distinct mentalities. The central goal of the present approach is to get a handle on the problem of description of non-human types of minds first and then to see what we can learn about the cognitive mechanisms that can act upon mental structures to realize cognitive behaviors. With this goal as part of the methodology, this book sets out to examine the unique connection between natural language and naturally possible minds. While the conceptual apparatus required for later discussions has now been refined, we have not yet addressed the question of why natural language is so special. Section 1.3 will assess the merit of this question by contextualizing it in the wider domain of the investigation into the very nature of intelligence whether biological or otherwise. In this connection, various other proposals that approximate to the exploration of non-human mentalities will also be evaluated so as to check how the case for mental structures can be reinforced once the weaknesses of these approaches in getting to grips with the non-human type(s) of mentalities are shown.

1.3 Why Natural Language?

Investigations into the nature of the mind have proceeded with the supposition that an understanding of the structure of the mind can offer insights not

only into the properties of mentality but also into the very possibilities of having a mind. One of the ways of examining the structure of the mind is to study the mental structures that human language as a cognitive organization gives rise to. On the other hand, a way of understanding the mind itself is to understand the nature of intelligence which seems to encapsulate everything we tend to associate with a cognitive system that evinces aspects of mentality. The former naturally lends itself to being made into a linguistic inquiry, insofar as it relates to an aim of understanding the mental structures behind the linguistic structures and representations. But the latter extends to a vaster intellectual territory within which the nature and form of intelligence of humans, different types of machines and other creatures in substance-independent terms is examined from computational, philosophical, biological and perhaps anthropological perspectives. Even though these two threads of natural inquiry in its general sense have different natures, goals and methodologies, they have a lot in common. It is not quite hard to see that an inquiry that projects a window onto the hidden texture of cognitive structures, insofar as it is revealed by an inspection of linguistic structures, can also reveal something about the form of intelligence. This is so because the cognitive structures underlying linguistic structures connect and shape what cognitive systems operate on, manipulate and exploit in any activity that counts as intelligent in some demarcated manner. In fact, the latter inquiry is often linked to what is usually done in artificial intelligence (AI). It may be noted that an inquiry into the nature and form of intelligence of humans, different types of machines and other creatures in substance-independent terms subsumes, rather than forms a part of, the study of AI per se. It is not unreasonable to argue that the underlying raison d'être behind the study of AI is the quest for other *possible* forms of mind. And it is this aspect that informs the inquiry that delves into the nature and form of intelligence of humans, machines, other creatures and even plants.

At the same time, it is also vitally important to recognize that the quest for other possible forms of mind can make sense only if the necessary and sufficient properties of minds are adequately understood. Clearly the marks of what it is to be mental have some substantive connection to the language capacity, on the grounds that the language capacity makes viable certain cognitive structures, especially certain kinds of thoughts that we as humans entertain. However, we have reason to believe that this cannot be the whole story in itself, for creatures other than humans such as dogs, cats, parrots, crows, pigeons, dolphins or other primates do not possess the *kind* of language capacity humans are endowed with. Be that as it may, there seems to be something irreducibly linguistic in any conception of what it is for something to be mental, and by virtue of this, a conception of intelligence that borrows something

from this conceptualization of cognition is bound to incorporate aspects and properties of the linguistic organization of intelligence. Most significantly, when Alan Turing, the father of modern computer science, came up with the concept of a test that would count as the operational diagnostic for the inspection of the marks of intelligence in digital computers, the test which is known as the Turing Test (Turing 1950) was described essentially as a linguistic test. The test involves a computer and a human both hidden behind a screen or veil on the other side of which sits a human who as the judge scrutinizes the linguistically framed responses from both the computer and the human in reply to questions posed by him/her. Both the computer and the human are certainly indistinguishable to the judge, since the judge does not know which response comes from whom. All that the judge will have to do is check the verbal responses in response to his/her queries in order to tell the machine apart from the human. Note that the entire test has been designed in a fashion that involves natural language conversations which humans have to verify with a view to determining whether the responses come from a machine or from a human. Regardless of whatever demerits the test in itself has (see for a relevant discussion, Proudfoot 2011), the test has a lot to say about the connection between natural language and (natural) intelligence.

The foremost question is: why did Turing think of natural language conversations when designing a test that could decide the case for machine's intelligence? After all, there is no *logical* reason why the test as such could not have involved a task such as generating visual images or analyzing sounds or moving things around or even drawing a picture. The question that bothered Turing is 'Can machines think?' Since no definition of thinking that will be appropriate enough to conform to well-demarcated specifications applicable in diverse scientific contexts or to the normal use of the word 'thinking' can be formulated, Turing replaced that question by another relatively unambiguous and precisely framed question which asks whether a machine can play what he called an 'imitation game'. Clearly, when thinking of natural language conversations, Turing had something in his mind, as he says

> The new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man. … We do not wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane (pp. 434–435).

It is clear from the passage above that Turing differentiated cognitive capacities and processes from mere physical capacities of humans, given that physical

capacities are insignificant and irrelevant when the goal is to test machines on the capacity for thinking. Furthermore, while considering potential objections to his proposed test, Turing also thought it appropriate to take into account possible disadvantages that machines could face during the performance of the test. For example, humans can pretend to be machines and this action may weigh heavily against the machine involved in the test, for humans are not good at many tasks computers are good at (such as mathematical calculations) and hence computers can be easily caught. For the simple reason that this could put machines at a disadvantage, he also considers the following objection.

> May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection (p. 435).

It should be noted that Turing acknowledges that the objection to engaging in natural language conversations in the test is a strong objection indeed, although he does not say anything concrete so as to weaken or eliminate any possible problems the objection in itself may carry. Let's consider the objection because the objection in question constitutes the crux of the matter this book will be concerned about. Suppose a task other than engaging in natural language conversations is fixed as the task which both the machine and the human involved will have to perform. Thus, for example, machines can carry out the task of analyzing images pixel-by-pixel which humans never do. But then this will disadvantage the human engaged in the test, since humans cannot, without any external aid, execute this task anyway. Plus it is not clear how the task of analyzing images pixel-by-pixel can be equated with thinking. Let's then consider the other possibility. What if we pick up a task which is different from what humans do *and also* from what computers do, along with the condition that the task chosen should be taken to be thinking in some sense? This question seems meaningless, on the grounds that some activity or task that can be taken to be thinking in some sense cannot plausibly be disjoint from what both computers and humans do. If a task or activity can be reckoned to be thinking in some sense, why cannot it be performed by humans, regardless of whether it can be performed by machines or not? Without doubt, there can be some potential tasks or activities that both computers and humans do not or cannot do. For instance, traveling backward and forward in time, gazing at all stars in the universe at once, playing with gigantic physically located buildings, running with the speed of light and so on are not the kind of tasks that

machines or humans do. Even if any of these tasks is benchmarked for the test of machine intelligence, it is not clear whether these tasks have the marks of mentality. Nor do we know whether these activities can be identified with the process of thinking per se, although they may require and also involve planning and appropriate processes of reasoning on the part of the agents that *may* engage in such activities. Faced with such a difficulty, we may attempt to fix the definition of thinking so as to let it apply to a certain range of cases reasonably constrained. However, it is a pointless task that Turing also recognized.

Whatever way we may try to settle the question of what constitutes thinking per se, we undertake to reflect upon a different question, that is, the question of how the language capacity connects to the way we characterize something as mental so that we can figure out what possible minds may look like. We may now explore two possibilities that may help us get a handle on the complexity of the issue at hand, depending on whether we take into account natural language or not. So let's first suppose that the nature and form of possible minds can be investigated by not postulating natural language as an intrinsic element of cognitive systems. If we adopt this possibility, we can explore the structure of possible minds that do not possess the cognitive capacity language affords. At this juncture, it appears that we will have to take into account the mind-like properties of languageless creatures such as ants, cats, snails, dogs, squirrels, pigs and so on. While the uniqueness of natural language may dispose many people to include other animals for the exploration of the question of how to shed light on the nature of mentality minus natural language, one may, with a reasonable degree of certainty, impugn the statement that the word 'natural' in natural language should be reserved only for humans, unless one has stipulated the demarcation of the meaning of the word only for humans (see for a related discussion, Lyons 1991). Assuming that the denotation of the word 'language' can be made to incorporate the systems of signs—however developed or impoverished—that are rudimentary enough to be used by different animals, we may see how we can make certain conjectures about the nature of possible forms of mentality. If we follow Luuk (2013) in this regard, a whole hierarchy of referential properties of symbolic systems emerges, provided that we accept that the word 'symbol' can be reliably interpreted to have developed from the broader-level category specified by signs in the Peircian system, which has a tripartite organization structured around icons, indexes and symbols.[3] The hierarchy can be represented as follows:

3 Icons bear an imagistic (physical) resemblance to the object which an icon is an icon of (for example, pictures). Indexes have some natural or causal or sensory connections to the object an index is an index of; for instance, smoke is an index of fire. Finally, symbols are signs

Signs → Denotation → Paradigmatic Connotation → Syntagmatic Connotation → Definition

Monadic signs are simplex signs that depend on various stimulus–response relations. Alarm calls of vervet monkeys, crows can be of this kind. Denotation, in Luuk's formulation, is a relation between a sign and its conceptual content, that is, the mental image of the entity referred to. Paradigmatic connotation depends on the logical-conceptual relations between signs; for example, part-whole relations, type-token relations, inclusion/exclusion relations etc. fall under this category. The same holds true for the predicate-argument relations; thus 'cars' predicated of 'red' instantiates a relation between an argument 'cars' and the predicate 'red'. Syntagmatic connotation consists in the combinatorial relations that obtain among different signs. That prepositions, for example, precede the nouns in English, as in the prepositional phrase 'in the garden', is a matter of syntagmatic connotation. Finally, definition is a higher-order relation which is by its very nature parasitic upon other relations. Thus, for example, if we define heat as the motion of molecules of matter, the notion of molecular motion depends on certain other relations involving molecules and motion, and so on. Luuk claims that the human symbolic capacity is distributed among the interpretative correlates of all these five elements such that in any instance of interpretation of a sign any subset of the five symbolic interpretative potentialities can be utilized. Most importantly, Luuk argues that the hierarchy corresponds to the evolutionary trajectory of the symbolic capacity in the biological world, and that animals such as vervet monkeys, gray parrots, bottlenose dolphins, bonobos etc, may possess the first two symbolic capacities, namely sign-making and denotation in the hierarchy shown in (1).

This raises some important questions for us. If we grant that sign-making and denotation are (also) *part* of the symbolic capacity characterizing the human linguistic capacity, it is not immediately clear how we can make sense of the question of throwing light on the nature of mentality minus natural language when we turn to other animals, assuming of course that we are referring to human language when talking about natural language. This is so because what is in essence a part of the human linguistic capacity (that is, sign-making and denotation) does not plainly detach itself from the extension or conception of what human language actually is. That is to say that sign-making and denotation cannot be independently characterized once for non-humans and then

that are arbitrary and stimulus-free, that is, are used in the absence of the object denoted, and they do not bear any physical resemblance to the object (words in natural language, for instance).

for humans. And if this is so, we cannot 'frame' the notion of mentality minus natural language because the traces of human language linger on even when we attempt to formulate the notion of mentality minus natural language, especially for other animals. The framing itself, in virtue of involving sign-making and denotation, carries over properties and conceptions of human language. On the other hand, if we are ready to accede to the proposal that other animals (may) have different systems of signs, regardless of whether or not they have certain overlaps with, or are subsumed by, the entire repertoire of human symbolic capacities, some glimpse into the structure of possible minds can be thrown. Note that this proposal appears to obviate the anthropomorphic bias, so long as we are inclined to think that animals have their species-specific independent systems of signs and that such systems of signs may look impoverished when compared to the human system of signs just as humans' auditory or olfactory capacity is impoverished when compared to that of dogs or lions. In such a case, we understand a lot about the structure of possible minds minus human language. We come to observe that dependencies involving stimulus–response relations and mental imagery in denotational reference making portray possible minds with such capacities as systems that have a rich finely tuned low-level visual faculty along with a minimally structured memory attuned to the empirically perceived world out there. Such minds can make stimulus-bound responses, possibly categorize different *types* of stimuli, and conceptually differentiate between different tokens of stimuli. Additionally, having mental imagery requires episodic memory, semantic memory and perhaps a type of intermediate-term[4] memory which can organize the experiences and interactions with the outside world, thereby also facilitating learning construed in its generic sense. Importantly, such kinds of minds may also be able to have a minimal theory of mind in the sense formulated by Butterfill and Apperly (2013). That is, such minds may have capacities of goal-directed action, encountering and *plausibly* a form of mental registration. The first two do not require any form of mental representation, in that a goal-directed action is cashed out in terms of a *function* which specifies an outcome or a goal to achieve which an agent must engage in certain activities, and encountering is simply a non-representational relation between an object and an agent, while the third does require representations because mental registration requires not merely a triadic relation between an object, an agent and a location, but

4 The intermediate-term memory is sandwiched between the working memory and the full-fledged long-term memory remaining active over an extended period but does not thereby harden into long-term associations (see Donald 2001). The memory formed during symbolic communications can be of such kind.

also the mental representations of each. Pigeons, dogs, cats, snakes, chimpanzees, crocodiles may have such forms of minds.

However, this is again problematic on the grounds that possible minds are thus interpreted to have some sort of language-like capacities manifest in the systems of signs such minds possess. Any systems of signs are ultimately bound to be language-like, irrespective of whatever such systems of signs are rendered. If this is what the whole thing turns out to be, understanding the nature and structure of possible minds minus language broadly conceived can be more daunting than can be naturally supposed. Plus we do not certainly seek to understand possible minds in theoretical terms only in the animal world—we aim to understand possible minds in a broader sense in computers, machines of different kinds, and other artifacts designed by humans. Perhaps a better approach towards this problem can be taken by considering a broadly construed ontology of representational levels, as is articulated in Bickhard (1998), which specifies a series, or rather a hierarchy of differentiated levels of representations that emerge through rich matrixes of *interactions* obtaining between agents and the world out there. Many of these levels in the hierarchy do not presuppose the existence of representational capacity in the internal states of the systems concerned. For the purpose at hand, we may imagine that these systems could be systems implicit in machines, animals or maybe even in plants. Thus we cast our net wide enough in order to capture a gamut of possible minds which is as broad as possible. It needs to be stressed that the ontological hierarchy specifies various kinds of representational ascriptions that may be grounded in different intentional ontologies. What this means is that our ascriptions of representational capacities to some system-internal properties in a machine, for example, can have various interpretations that depend on exactly where in the ensemble of different intentional ontologies we locate such representational capacities. Thus, for example, any intentional stance that takes a system to be *about* or *oriented toward* something can be cashed out in terms of a minimal ontology or even no ontology. That is, we may show no commitment to any restriction that determines whether it is machines or animals or whatever on which we assume an intentional stance. Under this construal, even a plant can have belief-like states when a plant such as a pitcher plant is oriented towards an insect, or when a machine such as an elevator is oriented towards what it elevates, namely humans and other objects. Plus various kinds of presuppositions and representational constraints may also emerge through interactions of systems with the world located outside of the systems themselves. For instance, an ordinary fan is built in such a way that the functional presupposition that there would be air around and that there would not be any blockage that may prevent its wheeling is not to be

found inside the machinery of the fan—it is simply presupposed by way of the interaction of fans with the world outside. Many representational constraints that determine how a system will operate and which final state it will end up in are not built into the system as such; rather, they are part of the interactive potentialities determined by the functional relationships obtaining between a system and the environment in which it operates. Possible minds in automata, many animals including primates and plausibly some plants may be described in these terms.

Note that this way of approaching the question of how to explore the nature of mentality minus natural language is more appropriate and suitable for the examination of the *formal structure* of possible minds across a wider range of entities—machines, animals and plants. For, even on this proposal language and consciousness are higher-order cognitive phenomena or forms of representational capacities which arise from a much more enriched and specialized ontology of interactive dynamics involving the agent's internal systems, the environment and the social world. Thus, understanding the question of how to explore the nature of mentality minus natural language boils down to understanding the form of the virtual mind-like emergence of interactive possibilities that give rise to representational constraints as well as to certain layers of functional presuppositions which both facilitate and constrain learning during such interactions but are not *explicitly* encoded or present anywhere in the organism/system concerned. Whatever merits or advantages this proposal may have, this does not advance our understanding of our question anyway. Although it needs to be made clear that the hierarchy of ontologies of different kinds of representation is not *exactly* intended to be deployed for the exploration of possible minds vis-à-vis natural language, it is doubtful that the form of the virtual mind-like emergence of interactive possibilities will (ever) gain a purchase on the nature of self which constitutes the core of mentality. This holds even if interactive possibilities afforded by a system within its environment may approach and thereby approximate to the contours of many properties of what Deacon (2012) calls 'ententional' phenomena, which are intrinsically incomplete by virtue of being related to, or constituted by something which is not intrinsic to those phenomena in question. Note that such ententional phenomena are phenomena having certain properties that are *other than* what their (physical) constitution entails, and include, for example, functions which require satisfaction conditions, thoughts which have contents, purposes which have goals or even subjective experiences that presuppose the existence of a subjective self. Given this characterization of ententional phenomena, the functional roles and organization of the systems in artifacts, tools or plants with respect to their respective environments can certainly give rise

to functions or goal-like states by way of the interactive potentialities manifest in the systems concerned. However, there are certain problems that we need to consider here. Functions and goal-like states notwithstanding, non-biological systems, at least in artifacts or tools, may not have well-developed self-like states which may go on to constitute subjective experiences. The plain reason is that any self-like states that may emerge in such systems cannot be fully autonomous and agentive, although the systems may, whether now or in the future, exhibit properties of self-repair, self-production or self-reconstitution which characterize self-organizing processes found in nature. It needs to be made clear that self-organizing processes found in different organisms are in essence marked by the physical and chemical processes in organisms that draw energy, materials or other resources from nature and create and sustain themselves.[5]

From this perspective, it is in a sense reasonable to hold that plants and other animals can have self-like states characterizing and constituting subjective experiences in virtue of being autonomous and agentive. The form of subjectivity in plants and animals may arise not merely from the generation of functions, representations and goal-like states, but also from the subjective constitution of a self that propagates its organization which may well be called 'teleodynamic', to borrow a term from Deacon. Teleodynamic processes are those processes that represent within themselves their own dynamical tendencies by having the whole produced from the parts and then having the parts produced from the whole, thereby generating a self that continually creates, renews, preserves and interprets itself. Teleodynamic processes—which emerge from and ride on simple self-organizing processes, that is, Deacon's morphodynamic processes—give rise to properties that qualify as sentience and the locus of subjectivity. Most importantly, teleodynamic processes are characterized by the dynamical *constraints* of their organization which encapsulate a *restricted* space of possible degrees of freedom that the internal systems of organisms causally generate when organisms take energy from the environment converting it into something necessary for growth, metabolism and reproduction. Now it may be emphasized that plants and other animals can have a form of subjectivity that allowably fits a mind-like entity if we attribute our certainty to the assumption that various self-like properties of the mind derive from the dynamical organization of different absences. And if

5 Importantly, Deacon calls such self-organizing processes 'morphodynamic' processes which are such that organisms having morphodynamic processes (such as bacteria) take from nature what they need and then create order by incessantly producing new structures inside or outside the boundaries of their physical organization.

this is the case, we can justifiably say that the mind-like entities of plants and other animals can possess non-representational and/or perhaps minimally representational capacities in perceptual, motor and sign-making activities. Such minds can detect objects in the vicinity, recognize the relevant predators and the members of the same species, make certain reliable categorizations of kinds of prey, feel pain, sense and also transmit certain signals necessary for the survival. Even though we can perhaps figure out what these kinds of minds look like or really are, there is perhaps a lot that is whisked off from the ground. As Dennett (1996a) point outs, the commonalities between kinds of minds are easier to discern and possibly discover as they come under the same recognizable larger envelope, while it becomes harder and harder to determine the finer details of differences that can help track the characteristic cognitive differences among kinds of minds in substantive terms. Teleodynamic processes of minds identify general characteristics of a larger envelope under which various kinds of minds of plants and other animals can be brought together, but it is not clear to what extent and how this can in itself reveal substantive cognitive differences among different kinds of *possible* minds. Teleodynamic processes are just that—expecting something more than that is not what the form of such processes warrants.

Given these problems in characterizing in formally explicit terms the substantive cognitive differences among different kinds of *possible* minds, we can perhaps do much better by looking into the properties of natural language phenomena within and across languages. Natural language is important for many reasons. Our sensory systems organize experiences in terms of the sensory qualities that are combined or simply collapse by forming manifolds of percepts, and then language imposes its own organization on these percepts or organized forms of sensory experiences. Language is thus a second-order cognitive system that formats forms of sensory experiences and moulds linguistic representations that build on those forms of sensory experiences and creates more and more complex abstract representations divorced from their sensory origins. In a sense, language is perhaps the only cognitive system that projects a window for us onto the interior space of our own minds, and it is not clear whether any other cognitive faculty (such as the faculty of memory or the motor system or even the attention system) has this capacity[6] (see for a discussion,

6 In many cases and in significantly relevant respects the faculty of memory or even the attention system is shaped by the linguistic system, inasmuch as linguistic labels help index and thereby track items and experiences stored in and recalled from memory, and in addition, it also helps enclose diverse ranges of experienced items within a constrained space of our attentional focus. But this is not to deny that the faculty of memory or the attention

Torey (2009)). Even the human emotive system, which is phylogenetically older than the linguistic system, cannot be exactly said to come closer to this, although emotions are ways of feeling one's self rooted in the body (Slaby 2008). Additionally, higher-order emotions (such as shame, embarrassment etc.) often rest on the intricate interplay between the linguistic system and the symbolically grounded cultural praxis, as language and emotion develop in close harmony shaping one another's cognitive representations (see Mondal 2013). Most importantly, a diverse variety of linguistic structures with all the complexities and idiosyncrasies can offer insights into the mental structures that correspond to those linguistic structures. Assuming that a number of such mental structures may be shared by members of other species, regardless of whether or not these species have the means of articulating them or encoding them in expressions that can match the complexity of syntactic structures in human language, we can propose to investigate the nature of possible minds by extrapolating from what widely diverse types of linguistic structures across natural languages reveal. This has some crucial advantages that cannot be overlooked.

First, the harder-to-determine aspects of mentality conceived in its general sense can be tapped if the form and structure of possible minds is explored by figuring out what a wide variety of linguistic structures reveals about an assortment of possible minds. The only caveat in this proposal is that the syntactic structure of natural language is not taken to be the pivotal point for the extrapolation from various types of linguistic structures to ranges of possible minds. Rather, it is the corresponding mental structures that can be *inferred* from various types of linguistic structures which will constitute the fulcrum of the proposed extrapolation. The talk of mental structures corresponding to various kinds of linguistic structures is, to a great extent, in tune with Jackendoff's (1983, 1990, 2007) *conceptual structures*, which within the theory of Conceptual Semantics characterize what humans conceptualize—the language-independent mental representations that are structured around linguistic constructions. Since mental structures corresponding to various kinds of linguistic structures can be even human language-independent, this does not also invite the anthropomorphic bias.

Second, the projection of the space of possible minds through other sensory-cognitive systems is bound to run into severely paralyzing problems. It may be noted that other sensory-cognitive systems in humans are not as

system also facilitates the functioning of the linguistic system, especially during language processing.

well-developed as they are in many other species on earth, while the faculty of language is developed in *Homo Sapiens* to an extent which is perhaps unparalleled in the entire animal kingdom. Given that this gives rise to a complementary distribution of the relative differences in cognitive capacities of humans with respect to other species or of other species with respect to humans, it would be in any event biased to look into the nature of possible minds through the lens of whatever cognitive system/faculty we pick up. In fact, this can tip the balance in favor of other sensory-cognitive systems since humans do not appear to distinctively excel in all cognitive capacities except in the memory capacity and the linguistic capacity (see Tulving (1985), especially for episodic memory; see Chomsky (1985) for the linguistic capacity). Many other cognitive capacities including the capacity for socio-cultural cognition that can be said to be uniquely present in humans are in some sense or the other co-dependent or co-developing capacities of the memory capacity and/or the linguistic capacity (see Holtgraves and Kashima (2008); Fitch, Huber and Bugnyar (2010); but see Cheney and Seyfarth (2007), who think the linguistic capacity has arisen from the capacity for social cognition—which is not exactly at odds with the co-dependent development of the capacity for social cognition and the linguistic capacity). Overall, all co-emerging cognitive capacities are in a sense unique in humans, and most sensory-cognitive systems other than those co-emerging cognitive capacities/systems are shared with other species and may well have had a common homologous origin. Moreover, language being the prime cognitive capacity that helps humans to look inside themselves and also to engage in various kinds of thoughts that can be entertained, it would be more reasonable and appropriate to make an attempt to understand the question of exploring possible minds through natural language. As a matter of fact, it is hard to imagine how approaching the question of exploring possible minds through the window of other sensory-cognitive systems minus the linguistic capacity can even make sense, for any cognitive capacity minus the linguistic capacity in any creature cannot have the theorizing itself get off the ground in the first place.

Third, the point made just above readily relates to the problem of intentionality vis-à-vis (natural) language. Intentionality is a property of mental states, objects or events which characterizes aboutness or directedness at objects or states or affairs (Searle 1983; Lycan 1999). In other words, intentional states are directed at the world in virtue of the specific kind of relationship that obtains between intentional states and things (either in the mind or in the world out there). More significantly, Brentano (1874) hypothesized that all mental states are intentional states. What this means is simply that all mental phenomena involve directedness or aboutness toward objects or entities

or states or affairs.[7] At this juncture, it appears that it would be worthwhile to look into the question of exploring possible minds by examining the nature of intentionality, primarily because even natural language—especially linguistic meaning—can be supposed to have derived from intentionality which was probably present in the earliest life forms, as Searle believes. Thus it seems reasonable to investigate the question of exploring possible minds by verifying whether something possesses intentionality or not. However, this way of formulating the question has some crippling disadvantages. Checking whether or not something, say, X, rather than Y, possesses the property of intentionality cannot be done by checking the internal parts of either X or Y, for intentional states cannot be directly seen within a system or a living entity. Intentional states are *inferred* from the outward behavior of an entity or from the outputs of cognitive processes. Plus the ascription of intentional states to non-living things is fraught with a number of deep conundrums. So it is not even clear whether we can include machines if we wish to include machines while we investigate the question of exploring possible minds.

In addition, the philosophical debates on the question of whether or not humans' intentionality is intrinsic or machines' intentionality is derived also vitiate, if not entirely eliminate, the prospect of applying intentionality as a good test for exploring possible minds. On the one hand, ascribing intrinsic intentionality to machines risks having a blithe disregard for the relevant facts, for, if machines had intrinsic intentionality, machines would have been able to perform all sorts of intentional acts, for example, intending, pretending, believing, making commitments, lying, guessing, wanting etc. etc. So far as we know, machines do not engage in all these. Moreover, if machines can generate on their own algorithms, or rather programs to repair themselves or even 'reproduce' in accordance with 'goals' and 'purposes' that machines set on their own, this can indeed be taken to be a good test for the possession of intrinsic intentionality in machines. However things come about, there is more to it than meets the eye. Now, on the other hand, if humans had derived intentionality rather than intrinsic intentionality, this would invite a problem of infinite regress. If the intentionality of humans is derived from something else, say, from evolution, as Dennett (1996b) believes, then what is the intentionality of evolution derived from? And so on *ad infinitum*. One cannot, on

7 While the Brentano thesis has received support from Crane (2001), the thesis has been criticized by Millkan (1984) and Nes (2008) on the grounds that the feature of intentionality is also true of many non-mental phenomena (for instance, the directedness of the stomach toward (digestion of) food). Additionally, the absence of directedness of pain experiences is adduced to counter the claim that all mental phenomena are intentional.

any metaphysical grounds, maintain that the nature of the intentionality of evolution does not need to be traced to anything else save itself.[8] Arguing that the intentionality of evolution is fundamental is a *non sequitur*, since we have already allowed machines to have derived intentionality, and it is not clear why we should not ascribe derived intentionality to evolution as well. Why should evolution have a privileged status over machines in this regard? After all, the process of evolution is also machine-like or algorithm-like, as Dennett himself claims. Therefore, the argument does not go through.

Regardless of whether or not humans' intentionality is intrinsic—and in fact the present discussion does not hinge on whether or not it is so,[9] the question of exploring possible minds can be approached in a more sensible way. The range of possible minds can be reasonably construed and so constrained to include machines' potential form of mentality, especially if we inspect the intricacies of mental structures hidden behind natural language constructions. It is because mental structures hidden behind natural language constructions cannot be *solely* possessed by human minds even if humans uniquely possess the linguistic capacity. It needs to be emphasized that it is not the mental structures behind natural language constructions per se that have meaning; rather, natural language constructions/expressions have meaning (also noted in Davis (2003)). And if so, mental structures of various sorts that constitute the contents of linguistic expressions can be conceived of in human mind-independent terms as mental structures behind natural language constructions do not in themselves possess meanings for humans, or for that matter, for other entities. Mental structures concealed beneath natural language constructions or expressions may thus be projected for possible minds of animals, machines and also plants when the range of possible minds is explored in terms of such mental structures. This idea is, however, different

8 Even though evolution is not supposed to have any goal or aim and, for that matter, a form of intentionality, appealing to evolutionary grounds for making claims about the fundamentality of the design of evolution is circular. The reason is that the act of appealing to evolutionary grounds for claiming that the design process of evolution is fundamental is done in order to argue that the human intentionality cannot be intrinsic, but at the same time, the argument that the human intentionality cannot be more intrinsic than the intentionality of cats, for example, is adduced in order to establish that the design process of evolution is fundamental, and that evolution cannot be said to have a form of intentionality.

9 The view that will be postulated later in Chapter 4 is that humans' intentionality is intrinsic, fundamentally primitive and may well be grounded in the human body, for it is even impossible to talk about intentionality in other entities in the absence of humans' intentionality which ascribes intentionality to other entities by means of inferences as the very act of ascribing intentional states to other entities is always inferential.

from the view of language adopted by Sperber and Wilson (1995), who consider language to be a medium for storing and processing information and thus thinks that other animals must have languages.[10] This view more than trivializes the notion of language, and hence nothing appears to prevent any other cognitive faculty—insofar as it stores and processes information—from being reckoned to be languages. In fact, it is pointless to tinker with the demarcation of what may be called language, and any hypothesis that turns on tinkering of such kind seems like a play on words without much substantive import. Note that mental structures concealed beneath natural language constructions/expressions are hence independent of and outside the boundaries of the object of the human semantic interpretation, given that the mental structures are not in themselves part of the human semantic interpretation which belongs in the domain of the human mind. This is, however, not to deny that mental structures can be modulated and shaped by natural language expressions. But at the same time, this cannot also prevent such mental structures, however structured and shaped by natural language expressions, from being *possibly* manifest or realized in minds other than those ascribed to humans. This is precisely because mental structures as may be shaped by linguistic expressions are not an *intrinsic* property of any mind. Consider, for example, the mental structure underlying the sentence in (1).

(1) John will travel across Australia no matter what it involves.

The sentence in (1) involves a mental structure that contains two contrasting thoughts. The matrix clause 'John will travel across Australia' introduces the thought that will definitely be the case or is bound to obtain, whereas the thought that presents a conflicting condition is provided by the subordinate clause 'no matter what it in volves'. On the one hand, it is evident that the mental structure of such constructions is *structured* by the structural organization of the expressions concerned, and hence we cannot get the same mental structure if we say, for instance, 'what it involves no matter John will travel across Australia', which is not a well-formed expression in English. But, on the other hand, there is nothing inherent in the mental structure in itself that can *logically* prevent it from being realized in, say, machines or even animals. If, for the sake of argument, mental structures are characterized as having meanings, nothing in principle can stop the meanings of mental structures from

10 They have considered a linguistic system to be a cognitive system, as opposed to a communicative system, instantiated in terms of information processing, and hence insofar as this is so, they deem that other animals and machines can have such systems.

having their own meanings and so on *ad infinitum*, thereby triggering an infinite regress of meanings all the way down the hierarchy of embeddings. Furthermore, the mental structure *can* be identified with a mental state which can otherwise be distinguished from a mental structure constituting the contents of a linguistic expression by the specific kind of abstraction of mental structures onto an inter-subjective level of knowledge.[11] Likewise, the mental structure of sentences such as the following cannot also be the exclusive property of the human mind.

(2) The kids swim as well as they dance.

The mental structure of (2) cannot in itself be constituted by the expression in (2); rather, it is constituted by the thought that the kids' swimming and their dancing are equally good. And hence there is nothing wrong in having the possibility that the mental structure in question can be projected for other possible minds. Thus the notion of the word 'mental' in mental structures has to be cashed out in terms of its substance-independent properties.

Overall, what is important is that natural language constructions or linguistic expressions, in virtue of being subject to the human mind's interpretative constraints, cannot be projected outside the domain of the human mind, inasmuch as the language capacity in humans is unique. Simply put, *at least some* mental structures behind natural language constructions/expressions can be *ontologically* located in many possible brains and systems, and they cannot be the exclusive property or part of the human mind because they are not (intended) to be mapped to further semantic structures in the first place. Mental structures behind natural language constructions/expressions are thus metaphysically autonomous entities in this sense, and this being so, they can be searched in many places and the hope is that they can be found too.

We have so far considered various ways of understanding the nature of minds with respect to the linguistic capacity or its structural system. It turns out, as this book will argue, that we have to look no further than the domain of linguistic structures themselves to assimilate the building blocks of possible types of mentality across the spectrum of diverse organisms and creatures. Linguistic structures are not in themselves components of minds. Rather, the mental structures that capture the organization of expressions within

11 See Mondal (2012), who has argued that linguistic knowledge recognized as the knowledge of linguistic expressions, mental structures plus their correspondence possibilities can be said to exist at an inter-subjective level of individual language speakers by way of abstraction from individual minds, even though the states of individual minds can also instantiate properties of such knowledge.

linguistic structures can be the potential candidates of mental types. This can be looked at from another perspective. Linguistic structures vary as a function of what mental structures they express that capture the semantic relations in a construction or across a range of constructions. As mental structures are conceived of as something linguistic structures express, it appears that the more variation in linguistic structures we find, the more possible mental structures we may find out. But the entire range of variations of mental structures riding on the variations of linguistic structures cannot simply be assumed to instantiate variations in mental types across different species. This would be a fallacious interpretation of what mental structures are supposed to capture as part of the building blocks of minds. The claim in the present context is *not* that the variations in mental structures corresponding to the variations in linguistic structures are identical to the variations in mental types across species. This would conflate all possible mental types within the *exclusive* envelope of mental structures as found in different linguistic structures (across languages) on the one hand, and distribute various mental structures *of human languages* among distinct categories of organisms and creatures on the other. Rather, the point of the whole exercise implicit in the line of reasoning employed above is to show that certain, if not the whole range of, mental structures can be extracted from natural language constructions themselves in order to test their viability for other organisms as well. Some such mental structures may turn out to be more general, while some others may prove to be more restrictive. The adequate testing ground here will, of course, be the ethological or cognitive-behavioral contexts of different species.

If this line of reasoning is on the right track, we can, of course, expect mental structures to be related to minds in a special way. As we have explored the nature of mental structures as they relate to minds, we have come to understand that the relationship between mental structures and minds is more subtle than may be supposed. In the present context, minds are ways of talking about mental structures, although minds can have domain-specific processes over above mental structures. Although the relation between natural language and minds has been touched upon in this section, the foundational assumptions and the theoretical contexts underlying that relation have not been so far examined. The next chapter aims to do exactly this. This is what we turn to in Chapter 2.

1.4 Summary

This book will thus examine the extent to which the relationship between natural language and the range of possible minds can be intimate. The way this relationship can be intimate will also be a part of the inquiry this book

will engage in. We shall observe that natural language and the extrapolation of a range of possible minds are inextricably intertwined. Researchers in AI and cognitive science have not given serious consideration to understanding this connection as deeply as possible. A lot of concentration has been on either developing theories of intelligence or debating the nature of human minds so as to say how humans differ from machines in various cognitive capacities. Significant as these issues are, I believe these issues bypass the fundamental question, that is, the question of whether we can unravel something about possible minds by examining the nature of mental structures revealed by different assortments of natural language constructions within and across languages. If the line of inquiry this book will undertake has anything to unlock, it must be such as to unlock the immense potential of natural language for cognitive science and possibly beyond. The better we understand this, the more we understand about the properties of mentality and the nature of intelligence in general. This can be appraised in view of the fact that current biological theories and philosophical hypotheses do not help much in understanding the structure and form of other possible types of mentalities. This is not because our biological and philosophical understanding of other possible mental types is limited by the internal inadequacies of the existing theories and hypotheses. Rather, it is because the available tools of biology and philosophy cannot reach into the realms of mental phenomena in other non-human organisms or systems and even plants by studying some intermediary object that can take us inside the domain of other possible minds. Language being the *sine qua non* of cognitive capacities equips us with the exact intermediary object which can offer glimpses not only into the realms of mental phenomena that are biologically instantiated but also into the structures and representations that can be brought forward to bear upon the question about the form of mentality in non-biologically grounded systems (such as computing machines). Overall, this book attempts to show how to integrate the biological understanding of animals and plants, the philosophical understanding of mentality, and the linguistic understanding of the nature of mental structures hidden beneath linguistic expressions into a whole that uncovers the nature of possible forms of mentality. Needless to say, the present book will be interdisciplinary in drawing upon insights from disciplines as diverse as linguistics, philosophy, anthropology, computer science, psychology, neuroscience and biology in general. Hence every attempt will be made to keep to a common discourse so that it reaches a larger audience in the widest possible spectrum of cognitive sciences. Having this perspective in mind, I strongly hope that lay people can also partake of the discussion the book will engage in; many of the issues to be pondered over and thus dealt with are everybody's concern, as far as one may reasonably believe.

The book is organized as follows. The book is divided into six chapters. Chapter 2 will focus on the linguistic foundations of minds as the affinity between language and the character of mind needs to be scrutinized and looked at from various perspectives in order to see which conception is handy enough for the present context. In Chapter 3 the descriptive formalism of mental structures will be formulated, and then the form of various kinds of possible minds in distinct species or organisms will be specified after the formalism is fleshed out with reference to a plethora of natural language phenomena. Chapter 4 will develop further connections to machine cognition, and Chapter 5 will explore plausible consequences for everything that can be reckoned to be cognitive and also the connection between the cognitive and possible types of mentality. Finally, relevant concluding remarks as they follow from the entire discussion in the book will be made in Chapter 6. Even if some chapters, especially Chapter 1 and Chapter 2, may be read on their own, there will be a certain degree of continuity from Chapter 3 to Chapter 6. Confident readers may thus move directly over to Chapter 3 and follow the threads of the narrative as it unfolds, while other curious readers may track the flow of the arguments involved right from the beginning of the book if they wish to do so. With this we may now turn to Chapter 2 to find out what it can tell us about the relation between natural language and the (natural) foundations of minds.