# 6 The Logical and Extrinsic Sources of Modularity

Thomas Bever
*University of Rochester*

## MODULARITY AND LANGUAGE

For a number of years, researchers on language behavior have believed that it involves the interaction of different kinds of partially autonomous systems of general and specific knowledge. That is, language is a modality, a natural kind of mental organization. The differentiation of such modalities as *language, vision, taste,* is pre-theoretically satisfying, but requires scientific explanation. How is it that they coalesce and emerge? How does the child know that aspects of his or her early experience are interrelated together and which motor patterns are related to them?

There are corresponding questions about the organization of information within a modality. For example, successful language behavior involves the appropriate interaction of systems of phonology, syntax, semantics, discourse, pragmatics, and world knowledge. Fodor (1983) sketched one proposal on the laws governing mental traffic between such systems. He crystallized a modern form of the old doctrine of "specific energy" of sensory systems, now coined, "modularity." Fodor's specific proposal is articulated and discussed elsewhere in this book. Certain intuitively appealing and widely believed aspects of this proposal are important for this discussion: Modules are architecturally segregated, that is, their internal processes cannot be mutually influenced; modules are neurologically distinct and reflect devoted innate neurological predispositions; modules utilize processes and forms of memory unique to each, that is, principles of "general cognition" either do not exist or exist

outside of cognitive modules in the system of general intelligence, the "central processor."

In this chapter, I sketch an alternative framework for a research program on the interaction of mental systems underlying language behaviors, and some current results that support that program. I argue that language behavior recruits a heterogeneous set of distinct capacities and neurological structures, each of which has intrinsic constraints on how it can interact with others. Furthermore, I raise the possibility that the differentiation of cognitive processes is general, cutting across types of behavior. These facts and constraints can result in modular-like properties of certain aspects of language without being unambiguous evidence for an innate and architecturally distinct module for language, nor for modules within the language modality.

A currently fashionable form of demonstration of the modularity of language is to show that computation of linguistic knowledge proceeds independently of other kinds of belief. I argue that such demonstrations may only reflect the necessary computational incompatibility of different kinds of information. Thus, the current experimental evidence for modularity actually follows from the fact that distinct levels of representation have distinct internal computational languages.

The preceding is in part a point of logic, not fact. Arguments do not require facts to be interesting, but they do require facts to be convincing. Accordingly, I also explore a kind of fact about some general cognitive bases for processes used in language, based on the isolation of biologically coded individual differences. I show that logical distinctions among kinds of cognitive processes used in language are reflected in different cognitive strategies used in biologically distinguishable populations. Such biological variation suggests that language processes do not rest on an isolated innate module, but rather, at least in part, are drawn from a set of generally available cognitive mechanisms.

## ARE WE MAKING MOUNTAINS OUT OF MODULES?

The background assumption underlying the modularity hypothesis is that mental life is computations—it involves the transmission and trans-mutation of symbols as inputs and outputs of operations. This assumption partially defines the necessity of some sort of modularity because specific transformational systems have correspondingly specific input and output schemata. The corresponding property of a module is that it resists all but a specified set of inputs and has a specified set of possible outputs. There is an intuitive ordering of such systems; the most plausible are those

clearly based on physiologically tuned input/output systems; less plausible are "modules" defined only in terms of function.

### Sensory Systems

The classic module is sensory, based on isolatable physiological constraints: The doctrine of the specific energy of sensory systems reflects the existence of specialized sense organs that normally are sensitive only to specific kinds of stimulation. The output of such a system is fed into a particular modality, regardless of the input: A classic and often cited example is that pressure on the eyeball is partially perceived as light. The isolation and analysis of such modules rests on clear phenomena, with clear neurophysiological explanations offered in current theories. Modules of this kind are not controversial, and set the guidelines as to what to look for as modules in more complex behavior.

### The Whole Iguana

Much of complex behavior, such as object recognition, appears to operate as though it were a sensory/motor system or instinct, even though it is implausible to argue that it is completely innate. Claims that there are modules at this level of organization are the most interesting because they are also the most controversial. A major argument in favor of their existence in perception is that there must be some boundary between what we are perceiving and what we expect to be perceiving. To paraphrase an argument from Fodor (1983), if I believe there are no lions in the room, and there are, it would be dysfunctional (and an evolutionary failure) to have a lion-perception system totally dominated by my beliefs. Hence, the processing of complex percepts (e.g., lions), must proceed with some autonomy from other sources of belief.

Language lies in this range of phenomena. On the one hand, it is an elaborate system of behavior and knowledge that interacts with many of our thoughts and precepts. On the other hand, like lion-perception, it must be potentially autonomous from beliefs, or we would never learn anything unexpected from linguistic information. Accordingly, the claim that language behavior is modular is often taken to be an empirical claim about the insensitivity to contextual knowledge of ongoing language processing: For example, it is generally claimed that there is an "architectural" barrier between meaning and syntax processing—the syntax processor must complete its work before semantic context can be involved. Semantic discourse effects can occur, but only after syntactic structure is assigned to sentences.

A brief consideration of how this proposal is expressed in psycho-

linguistic practice illuminates the difficulty of studying modularity in general. A typical example of such empirical controversy comes from the work of Marslen-Wilson and Tyler (1987), researchers who believe that language comprehension is not modular, in the sense that contextual information of all kinds plays a role at each point in processing. They attempt to show that an on-line task sensitive to local syntactic processing can be influenced by semantically based information. For example, if a subject hears a sentence fragment and then must quickly read aloud a word that appears on a screen, local syntactic agreement between the end of the sentence fragment and the word influences reading time. Thus, reading time for "is" is faster following the fragment in (1a) than that in (1b), whereas just the reverse is true for reading time for "are."

1a. Finding support ...
1b. Talking mothers ...

This effect is taken to show that the syntactic relations between a verb form and the following noun are computed quickly enough so that a verb agreement expectation is set up immediately. Marslen-Wilson and Tyler then asked the question, would context preceding an ambiguous sequence like (1c) determine the reading time effect? That is, would the sequence in (2a) lead to faster reading times for "is" than "are," and would the opposite result pattern occur following (2b)?

1c. Visiting relatives ...
2a. When it becomes a duty, visiting relatives ...
2b. When they stay too long, visiting relatives ...

Marslen-Wilson and Tyler showed that contexts did have such effects, and they concluded that the modularity hypothesis for syntax processing is false because clearly the contextual information was guiding the syntactic processing. There are a number of responses to this, but the empirical one is most revealing. Townsend and Bever (1982) noted that Marslen-Wilson and Tyler characteristically had only singular verbs and nouns in the contexts like (2a) and plural verbs and nouns in contexts like (2b) (also noted by Coward, 1982). Thus there might have been a word-to-word facilitation of "is" and "are," rather than a semantically mediated effect on the syntactic processing. Townsend and Bever tested for this and indeed found such a direct lexical associative effect. That is, regardless of whether the semantic context influenced the expectation of a singular gerund or a plural phrase, the word reading time for "is" was facilitated by the presence of a singular noun in the context, and time for "are" by a plural noun. The results also showed some (weak) semantic context

effects, but only for facilitating the reading time of "is" after contexts like (2a), not of "are." Townsend and Bever argued that the context effect is limited to the gerund because the gerund form (as in 1a) maintains the canonical English phrase order, in which a verb precedes its object, the sequence can be recoded immediately and given a semantic interpretation: This contrasts with the adjectival interpretation (like 1b), which is not a complete proposition. Townsend and Bever argued that the reason that semantic context can have an immediate effect on the singular gerund interpretation is that it has an immediate semantic analysis that can interact with the preceding context, whereas the adjectival analysis is not yet available in the form of a complete proposition at a semantic level. Thus, the modularity of processing between levels is a function of when each level of representation has a complete unit available. Semantic context can influence new semantic units as they are processed, but semantic context cannot influence the choice of a syntactic unit except when there is a complete semantic unit to which the syntactic unit has been linked.

This exemplifies an important distinction between two kinds of modularity. The classic claim about research like that of Townsend and Bever is that modularity is "architectural," that higher level semantic processes cannot influence lower levels. The cases studied by Townsend and Bever show that the mediator of such interaction is the form of information itself. That is, temporal discontinuities in the use of semantic information from outside a sentence occur because of discontinuities in the formation of semantic units inside the sentence. This brings us to a point of logic underlying the necessity of modularity when different kinds of representational systems are concerned: If the computational language of two systems differ, one cannot affect the internal operation of the other. This does not necessarily demonstrate an architectural boundary between them, because their mutual computational opacity would lead to such discontinuities of influence anyway.

Consider a cross-modal example as an extreme, say the matching of pictures to words. At first, this would seem to be an obvious example of two distinct modules at work, separated architecturally. Operationally speaking, to demonstrate their modular independence one would want to show that if the perception of a word was facilitated by a corresponding picture, it was only after the word was initially sensed. That is, a picture of an iguana can not directly facilitate perception of the isolated letter sequences, I, or IG, or IGU, or IGUA, but only of a representation of the word, IGUANA. Such facts, if true, would support the assumption that picture processing and word finding are architecturally distinct modules. But in fact, the results show something weaker. It is empirically reasonable that the computational language of object recognition is not

expressed in letter sequences: It follows logically that object recognition cannot inform word recognition. That is, the picture of the iguana cannot constrain the word-finding process to search for words beginning with I, or IG, or IGU, or even IGUANA: It can only constrain the word-finding process to find words with semantic structure related to that of iguanas. Of course, the word-finding process itself may quickly provide the information that the most important word semantically related to iguanas is "iguana" and they constrain its visual expectation for that word. But that constraint does not interact with letter recognition directly, only via conceptual and lexical levels of representation (see Schwartz & Schwartz, 1984).

This logical point makes it necessary to be cautious about any evidence for architectural modularity between different sources of information relevant to language behavior. Because there are empirical reasons to believe that the internal computational languages of nonlinguistic knowledge, semantics, syntax, and phonology all differ, we must expect on those grounds alone to find discontinuities in the apparent influence of information from one system on another.

## THE CENTRAL PROBLEM AND THESIS

With this background, consider more closely the nature of the claim that language is a module. One essential component other than architectural segregation of processing is that it utilizes behavioral principles that are unique to it. This uniqueness is presumably related to specific neurological bases, such as the left-hemisphere superiority for language. There appear to be appropriately distinct components of linguistic knowledge, such as that between the syntax and lexicon—the former represents computational knowledge, while the latter includes associative information between words and concepts. Finally, the child's discovery of grammar is viewed as depending on unique innate mechanisms. In brief, the claim that language is a module involves the following associated claims:

1. the unique neurological bases for language account for its locali-
   zation in the brain.
2. the different kinds of linguistic knowledge are uniquely repre-
   sented within the language capacity.
3. grammar is acquired via unique learning mechanisms.

ᴸ In this chapter, I outline the thesis that modular-like properties of language may have an initial source in prelinguistic subcortical mech-

anisms. I then argue that each of the three unique properties just outlined may be an expression of general properties of cognition. I argue further that the relation of those properties to biologically defined groups suggests that they are innate at a general cognitive level, not limited to language.

### Perceptual/Motor Instinct and the Early Segregation of Modalities

The apparent existence of general modalities is evidence for neurological prefiguring of knowledge domains. An increasing number of researchers are concluding that the young child has a specific set of innate knowledge domains. Their domains include "naive physics," "naive biology," "naive interpersonal psychology," "language," "person recognition" (see Baillergeon, Spelke, & Wasserman, 1985; Carey, 1987; Keil, 1989; Leslie, 1987; Wellman, 1990). Each of these domains appears early in childhood, partially segregated from the others: The set of domains partitions experience and knowledge in ways that are functional for an adult world, and that are the basis for module-like systems in the adult world.

At first blush, the child's ability to segregate his or her world into mentally relevant packages would seem to demonstrate an innate modular structure at the cortical level for each of those systems. Clearly, something must be innate to enable the infant to parse the world into experiences relevant to the acquisition of distinct mental systems. The question is: How can such constraints be causally relevant to the cortex, before the cortex is fully operational? The answer may lie in the existence of postnatal perceptual schemata and motor reflexes, which have autonomous characteristics. For example, the infant in the first 2–3 months of life has many highly tuned perceptual abilities (see Mehler & Dupoux, 1990, for a review). The infant can isolate critical aspects of his or her species—specific behavior (e.g., verbal input, faces); the infant also has perceptual and motor schemata responsive aspects of the physical world (Baillergeon et al., 1985). At the same time, the infant engages in apparently higher social functions such as facial mimicry, smiling, and reaching outward (Maratsos, 1982; Meltzoff & Moors, 1983; Murray & Trevarthen, 1986). These highly adaptive behaviors are the basis for the child's first lie, in the sense that they convey to adults the belief that the child has a functioning, physical, psychological, and social personality, but they occur based on subcortical and automatically functioning neural substrates. Such behaviors may depend on instinct because the learning structures of the cortex are not fully operational at birth, or simply because the newborn lacks experience from which to extract

physically and socially functional behaviors. On either interpretation, the functional role of such autonomous structures is clear; it provides the precortical infant with a repertoire of physical, perceptual, and socially functional behaviors. These structures in turn shape the cortical organization as it emerges. This idea must be true, at least in part, because the innate schemata themselves give specific organization to what the emerging cortex experiences: The child cannot help but learn to segregate kinds of experiences as they are grouped by the schemata the child is born with. In this way, an initially unorganized cortex can be trained to parse the world in accordance with the shape and grouping given to initial experiences by the set of early innate sensory/motor schemata. That is, the child's first successful lie, that it is actually a functioning mind, succeeds because the instincts themselves have been selected to survive in the physical world and fool the adult world. But in surviving and fooling the world, they also package experiences for the emerging cortex into the rudiments of an adult-like organization in modalities.

It is clear that an innate repertoire of precortical instincts relevant to worldly physics, species recognition, and species-typical behavior is functional for the infant at birth. If cortical computational mechanisms are not yet functional, precomputational instincts can carry the infant through the first few months of life while the cortex matures and accommodates to experience. In the meantime, innate physical perceptual/motor mechanisms save the child from corporeal disaster; species-recognition mechanisms orient the child toward caregivers; species-typical motor behaviors encourage the child's caregivers to treat the child as a human. But the innate mechanisms have a further consequence; they shape and partition experience into mentally natural kinds. Thus, the presence of such mechanisms before experience has interacted with cortical structures, must constrain the infant to divide the world into mentally natural modalities. This hypothesis offers an explanation of the innate basis of such functional modalities as "language," "naive physics," " naive psychology" without assuming an innate cortical computational module underlying them. Rather, the infant's set of innate socially adaptive mechanisms channel cortical experiences such that general computational processes operate on appropriately grouped activities. On this view, the functional modalities indeed have innate bases; infantile sensorimotor organizational instincts, selected to promote early survival in a physical, psychological, and social world, before more complex systems are available.

## Evidence for Processing Dimensions in Cognition

The modularity hypothesis is most naturally related to the fact that

mental life appears pre-theoretically to have distinct modalities. But, there is also a direct approach to the discovery of mental computational boundaries—the study of biologically differentiated processing systems. There are two important ways to show evidence for a direct relation between brain structures and particular cognitive processes: Study distinct behaviors in clinical populations with specific neurological damage; study different behaviors of brains in normal populations that can be biologically differentiated. I concentrate on what can be learned from differences in normal populations because of the obvious limitations of clinical data. Brain lesion studies are not experimentally controlled; the patients' overt behavior must be taken as an adaptation to the loss of a structure, not necessarily the direct expression of that loss. Similarly, genetically abnormal brains have the opportunity to adapt to their abnormality over a lifetime. I am not suggesting that nothing can be learned from such investigations; but the discovery of the isolation of an overt cognitive process in a clinical population requires a complex interpretation involving not only an explicit hypothesis about the lost process, but also about the nature of compensatory mechanisms contributing to the overt behavior (see Mehler, Morton, & Jusczyk, 1984; Shallice, 1984).

It may seem eccentric to expect that normal population variation can illuminate fundamental dimensions of cognitive processes. But, in fact, it is a direct consequence of our growing theoretical success in correctly distinguishing the dimensions of cognitive processing. Suppose the following:

1. cognitive processes are differentiated in part because of neurological differences.
2. those differences are in part innate.
3. the genetic code for such differences is complex, leading to relevant co-variation with other genotypes, and consequently phenotypes.

Then if we distinguish cognitive processes correctly, we may find that populations differentiated on the basis of biologically superficial phenotypic traits may also have characteristic differences in cognitive processes. In brief, the better our theoretical cognitive science, the more likely it is that theoretically distinct cognitive processes will appear differentially in biologically coded groups.

Much progress in understanding the physiology of sensation, perception, and cognition has depended on the study of normal populations: A rationale for such a study is that isolation of the component processes intrinsic to a behavior guides the search for relevant physiological bases. A frequent method has involved the theoretical and

empirical study of behaviors in terms of pairs of opponent processes. Color vision is a well-known example: The study of such phenomena as contrast effects, adaptation effects, and afterimages, supported a theory that color vision depends on the interaction of two sensitive systems, one sensitive to a blue/yellow dimension, the other to a red/green dimension. Within each system, activation of one of the dimensions inhibits the other. Thus, blue is the behavioral opposite of yellow, and red of green. The behavioral isolation of these linked systems was an indicator of a physiological dimension. Of course, the physiological distinction between these systems was ultimately confirmed and the chemical basis for the apparent processes became a topic of study.

The existence of opponent processes defines modularity in two different ways. First, opposition of processes implies a dimension that connects the two processes, linked by activating mechanisms (e.g., the red–green system). Second, each of the processes is distinct from the other, and hence is potentially a distinct subsystem of its own. In the following discussion, I apply the technique of isolating linked opponent processes to cognition. I first describe a formal distinction between a pair of processes, single and multiple, then offer some empirical evidence that the theoretical distinction is genetically coded in the biological substrate for three aspects of cognition—that is, the expression of the formal distinction in each of these domains is respected in the behavior of biologically distinguishable brain systems. The results support the validity of physiological relevance of the formal distinction: They also support the view that the associated specific properties of language are a reflection of corresponding general properties of cognition.

## COMPUTATIONAL MODULARITY—PROCESSING SINGLE VERSUS MULTIPLE REPRESENTATION

The concept is quite simple: Certain computational activities within a modality are opaque to each other, because of differences in the kind of mental action they involve, rather than differences in modality, architectural, or informational boundaries. Activities can involve one representation or several. For example, recognition that two right angles are identical is a direct process that involves one representational type. This contrasts with recognizing the part–whole configurational relation between a right angle and a square. Understanding the sentence "dogs chase cats" as a function of interlexical associations is a direct process, which contrasts with computing the formal semantic relations from the word order and inflectional information. Acquiring an abstract hypothesis via successive refinement of it based on evidence, contrasts with successive replacement of it by an internally generated hypothesis.

Learning to negotiate a known neighborhood in terms of an angular relation to a memorized local landmark is a different kind of process from learning to use a cognitive map that sets several landmarks in relation to each other.

These differentiations of mental activities by a number of representations is formal, and may not correspond to a dimension along which mental activities are actually arrayed physiologically. The following empirical discussions show how this distinction is in fact reflected in differences between distinct populations. First, the left hemisphere may be more computationally powerful than the right: This results in its specialization for so-called relational processing while the right hemisphere is specialized for comparatively simple processing. Second, the representation of associative knowledge in right-handers from left-handed families may be more diffuse than for pure right-handers: This results in relatively better differentiated associations in left-handed familials and hence, more reliance on local lexical knowledge during language processing. Third, female humans and rats use local and episodic knowledge to navigate, whereas males use motor and vector representations in spatial behavior. Fourth, the abduction of an abstract representation of an artificial language in female humans may depend primarily on single hypothesis refinement, whereas in males it depends more on competition between hypotheses. There are also some formal and empirical similarities between the population variations in spatial and artificial language learning. In particular, there are similar interactions between gender and familial handedness: In each case, the gender differences in performance are larger for subjects with left-handed families. Such similarities suggest the possibility that a single cognitive dimension underlies both types of behavior.

It is important for our conception of modularity that the three process dimensions are general. That is, they cut across modality domains. If this differentiation is borne out by further research, it will lead to a formulation of modularity in which behavioral modules draw on shared general cognitive capacities.

## THE GENERALITY OF THE PROCESSING DIFFERENCES IN THE CEREBRAL HEMISPHERES

The difference in function between the left and right hemisphere can be taken as an existence demonstration that certain kinds of computational processes can occur in isolation from others. The most stable finding is the relative vulnerability of language to damage in the left compared with the right hemisphere. This fact is consistent with the view that language is

a module, given the apparent special location of its neurophysiological bases. Indeed, numerous researchers have suggested that some special property of the left hemisphere may be the critical biological cause of humans' linguistic capacity (Calvin, 1982; Gazzaniga & Hillyard, 1971; Gazzaniga & Sperry, 1967; Hewes, 1973; Kimura, 1976; Le Doux, 1983; Levy, 1988). For a number of years, it was thought that there was a general partition of modalities between the hemispheres—language and logic to the left, music, vision, and art to the right (Kimura, 1976). Typical supporting experimental evidence was the fact that words are better perceived in the right ear. Typical examples of right-hemisphere superiority were superior recognition of an angle in the left visual field or superior recognition of short melodies in the left ear (Kimura, 1964).

The unique relationship between the left hemisphere and language might instead be due to a general computational property of language, which is better processed in the left hemisphere. Considerable evidence that this is true has accumulated over the last two decades. Language processing is peculiarly dependent on relational compositional processes: The formation of phonological, syntactic, and semantic levels of representation involve setting components into relation with each other. The left hemisphere excels at meeting this computational demand: Hence, language is most strongly represented in the left hemisphere (Bever, 1970, 1975, 1980; Levy, 1969).

A clear way to demonstrate that the left hemisphere is dominant for relational processing in general is to show the corresponding feature for nonlinguistic behaviors. For example, the right visual field (left-hemisphere) is dominant for visual tasks involving relations between images: When a square is followed by a right angle, the decision time that it "contains" the right angle is faster in the right visual field (Hurtig, 1982, see also Kosslyn, 1987, for related demonstrations). The perception of music provides another crucial test case. On the one hand, like language, music is serial and can involve higher-order integrations. On the other hand, basic melody perception is a classic example of holistic gestalt organization, as evidenced by the ease with which transposed melodies are recognized as identical. Various studies have shown that becoming musically sophisticated involves shifting from perceiving melodies as gestalten to perceiving them in terms of isolatable motifs and relations between motifs (Tan, Aiello, & Bever, 1981; Werner, 1948). Many other studies have shown that musicians also process melodies more actively in the left than right hemispheres (Bever & Chiarello, 1974; see Bever, 1980, for a literature review). That is, as the way one processes music shifts from direct to relational the hemisphere dominance shifts from right to left. This does not mean that an entire modality shifts, only the specific relational activities in it. Even though musicians' melody perception is

dominant in the left hemisphere, their recognition of two-note intervals is better in the right hemisphere (Kellar & Bever, 1980).

The hemispheric differentiation of processing styles might be fundamentally caused by a more specific asymmetry. For example, it might be that the left hemisphere has innate structures that are the critical cause for language. In this view, other compositional activities become better processed in the left hemisphere because of that initial predisposition, which generalizes to other tasks (Chomsky, 1965; Gleitman, 1981; Lanneberg, 1967). Other hypotheses suggest that the left hemisphere is innately dominant for integrated motor behaviors (Kimura, 1973); in this view, language becomes dominant in the left hemisphere because speech requires intricate motor behaviors.

It is difficult to disentangle these hypotheses from one based on a fundamental difference between the hemispheres in processing style.

Certainly, adult humans intentionally mediate many complex tasks by language, and language involves both compositional and motor integrations. The potential causal basis, however, for asymmetries is important for this chapter. If the ultimate cause is the presence of an innate neurological module for language in the left hemisphere, then the different language behaviors of the hemispheres is not fundamentally a developmental result of general computational modularity, but rather has a linguistically specific neurological base. Clearly, there must be some asymmetry between the hemispheres: The question is, how specific is that asymmetry? For analytic purposes, we can contrast a modular asymmetry against an asymmetry in computational power (Bever, 1980). Consider the following assumptions:

1.  the left hemisphere is computationally more powerful than the right.
2.  there is complementary inhibition between the hemispheres.
3.  relational tasks are computationally more demanding than associative tasks.

The third assumption is logically necessary: A relational task involves processing the relations between entities. A unary task involves direct processing of a single representation. Thus, relational tasks presuppose unary actions, and are more complex (for some otherwise undifferentiated computational engine). Assumption 2 is widely documented: Activity in one hemisphere inhibits activity in the corresponding area of the opposite hemisphere. If the first assumption is true, then during childhood, the left hemisphere would tend to take over relational activities such as language. This predicts further that unary activities would become asymmetric at a later age because only after language and

other compositional tasks are specialized in the left hemisphere would there be pressure to represent the simpler tasks asymmetrically. The relative late appearance of asymmetries for simple tasks has been noted by various researchers. This hypothesis also explains why it is possible for children with surgically removed left hemispheres to learn language at all (Dennis & Kohn, 1974, 1975; Newman, Lovett, & Dennis, 1986). There is some evidence that their performance is not as skilled as in children with only a left hemisphere, but that difference, as well, is explained on the view that the fundamental difference between the hemispheres is raw computational power.

One way to gain further perspective on the basis for hemisphere asymmetries is to examine their nature in other mammals, for whom natural language is not a theoretically corrupting influence. A recent review suggested that apes have a consistent asymmetry in favor of the right paw for tasks requiring intricate or forceful action, and the left paw for simple and passive actions (McNielage, Studdert-Kennedy, & Lindblom, 1987). This is consistent with the general nature of handedness and asymmetries in humans. Several experimental studies give more particular support. First, we showed that rats learn to recognize tone sequences better in the right than left ear: This difference increases as the length of the sequence increases (O'Connor, Roitblat, & Bever, in press). Such results are consistent with the findings with humans, but leave open the question of whether it is sequencing as such, or complexity of relations that brings out the superiority of the left hemisphere. A separate study with a dolphin suggests that the critical factor is not sequencing as such, but rather whether the task involves relational processing (Morrell-Samuels, Herman, & Bever, 1991). A dolphin had been trained to respond to hand signs, following a set of ordering constraints (Herman, Richard, & Wholz, 1984). There were three kinds of sequences involving several verbs and numerous objects: (3a) single signs involving only the dolphin's action; (3b) two-sign sequences involving an action by the dolphin on an object; (3c) three-sign sequences involving an action by the dolphin setting two objects into a particular relation to each other.

3a. Jump
3b. Hoop tail-touch
3c. Hoop frisbee bring (different meaning from 'Frisbee hoop bring')

Herman, Morrel-Samuels, and Pack (1990) had shown that the dolphin would respond to videorecordings of signs, presented on a 13-inch television through a window in its tank. This allowed us to control which eye could see a sign and to time the speed of the response. We found a significant superiority in the left eye for those sign stimuli that involved

only the dolphin or the dolphin and a single object. We found a right-eye superiority for the sign stimuli involving two objects in relation to each other. Unlike humans, the dolphin's eye is completely connected to the contralateral hemisphere. Thus, we can conclude that the dolphin's left hemisphere is superior for the relational sign stimuli, and the right hemisphere is superior for the stimuli involving at most a single object. The fact that the two-sign sequence is superior in the right hemisphere shows that the superiority of the left hemisphere for the more complex stimuli is not a function of it being a sequence as such. Rather it is a function of a sequence in which the named object's words have a structured relation to each other.

This range of studies suggest that the left-hemisphere superiority for language in humans may not be a reflection of a unique linguistic ability of the left hemisphere. Rather, the specific difference between the hemispheres reflects the logical distinction between relational and unary processing. This difference between the hemispheres holds for language behaviors and nonlanguage behaviors in humans, and for some behaviors in some other mammals. All the phenomena can be explained as the result of a difference in computational power between the hemispheres, which interacts developmentally with the logical distinction between the two kinds of processing. The conclusion here is that we do not have to postulate a left-hemisphere bound linguistic module to explain the presence of language in the left hemisphere of humans. Language is better handled by the left hemisphere because of what language is, but the unique *cause* of what language is lies elsewhere.

## FAMILIAL HANDEDNESS AND VARIATION IN KINDS OF LINGUISTIC KNOWLEDGE

Within language behavior there is a distinction between local-associative and global-computational knowledge. Understanding the sentence "cats sleep" involves accessing the associative and conceptual knowledge connected to the individual words "cat" and "sleep"; it simultaneously involves organizing the words into phrases with global thematic interrelations. Thus, using language involves both accessing lexical information about the reference and association of individual words, and organizing phrases in relation to each other at several levels of representation.

Various modular theorists have canonized this distinction into a claim that there are distinct lexical and syntactic modules that operate independently of each other. Accordingly, the role of lexical knowledge in sentence processing is an important factor in developing empirical

evidence for or against modularity of linguistic components. The distinction turns out to be reflected in a general difference in the way two groups of people process language. It is a common supposition that in normal people there is a large degree of homogeneity in the way language is organized neurologically and processed psychologically. The only exception to this generality is the general acceptance that left-handed people may have some differences in the neurological organization of language. There are also clinical reports that right handers with left-handed family members have relatively more linguistic involvement of the right hemisphere (Brown, 1976, 1978; Brown & Hecaen, 1976; Hecaen, 1976; Subirana, 1958, 1969). Our recent research suggests that the familial handedness of normal right-handed people also influences how language is organized and used (Bever, 1983; Bever, Carrithers, Cowart, & Townsend, 1989). The dimension that differentiates right-handed people with left-handed family members (LHFs) from those with only right-handed family members (RHFs), is exactly that between local lexical knowledge and global syntactic organization. LHF right-handers access local lexical knowledge more directly, whereas RHF right-handers are more immediately sensitive to grammatical knowledge.

The distinction between accessing local and global linguistic information runs through a number of experimental investigations. For example, LF subjects read computer-displayed short discourses about 10% faster when they are presented word by word, than when presented one clause at a time; RHFs read whole-clause presentations about 10% faster than word-by-word presentations. This follows from the view that LHFs naturally give emphasis to individual word recognition, whereas RHFs emphasize overall grammatical structure. This is also true of the behavior when understanding sentences. For example, Carrithers (1989) used a word-by-word reading time measure and had subjects read active and passive sentences: It is traditionally reported in the psycholinguistic literature that passive sentences are more complex, presumably because analyzing their meaning involves a computational step not needed for active sentences. Carrithers confirmed this finding: The reading time for final words was longer for passive than active sentences. This result, however, was true only for RHF readers. LHF readers showed no difference in response to the structural difference between passives and actives.

The difference between RHFs and LHF in sensitivity to local versus global information also effects the processing of single words. This can be shown in a priming lexical decision paradigm that differentiates associative from semantic relations between words. In this paradigm, subjects must decide as quickly as possible whether or not a briefly presented letter sequence is a word. On critical trials, the target word is

preceded by a word that has a specified kind of relation to the target word, either associative or semantic. For example, the relation between a target word "rose" and a prime word "thorn," is associative, that is, the prime word has a real-world connection to the target world; in contrast, the prime word "flower" has a linguistically specified relation to "rose," namely being its category. This distinction appears strongly in the different performance of LHFs and RHFs (Bever et al., 1989). LHFs show a strong priming effect of associative relations between words, and a moderate effect of linguistically semantic relations. In contrast, RHFs show a stronger priming effect for semantic relations and little for associative relations. The significance of this kind of result is that the task involves relations between isolated words, but differentiates between local linguistic and global extralinguistic information about the single words: RHFs are specifically more sensitive to linguistic relations and less sensitive to extralinguistic relations.

Other heterogeneous studies of these two groups confirms the distinct way in which they access information about language (see Bever, Straub, Shenkman, Kim, & Carrithers, 1990). As in the case of cerebral asymmetries, there are several points that follow from this discovery. First, it gives construct validity to the distinction between local-associative and global-structural knowledge about words and sentences. It suggests further that the organization of these kinds of knowledge is related to general innate variables.

This leads us to the question of mechanism: What might differ in the brains of RHF and LHF people that would account for this different behavior in language? One answer derives from the hypothesis that left-handedness is the result of an overexposure of testosterone during gestation (Geschwind & Galaburda, 1987). The idea is that the testosterone exposure occurs and slows all cerebral growth just during the period when the left hemisphere would otherwise be in a growth spurt relative to the right hemisphere. The result is that the hemispheres end up more equipotential, with a greater chance of left-handedness ultimately emerging. The details of this hypothesis and the evidence for it are not critical for this discussion. What is critical is the idea that familial left-handedness may be a marker for a uterine condition, which often, but now always leads to explicit left-handedness. On this view, right-handed people from left-handed families have more equipotential hemispheres at critical stages, but end up right-handed through chance factors. Although right-handed, these people have more bilaterally equal capacity than those from pure right-handed families: This may result in more right-hemisphere representation of lexical association and/or more widespread representation within the left hemisphere. Either way, lexical associative knowledge would be represented more diffusely in LHFs.

Several facts support the idea that LHFs have a more bilateral representation for language than RHFs. For example, LHFs show a greater incidence of crossed aphasia (aphasia resulting from an injury to the right hemisphere; Joanette, Lecours, LePage, & Lamoureaux, 1983; Luria, 1947). Second, although processing language tasks, LHFs show more bilateral evoked potential activity (Kutas, Van Petten, & Bessen, 1988; Kutas & Klueder, personal communication, May 5, 1991). Our results suggest that this bilaterality reflects a more widespread representation of associative linguistic information in familial left-handers. On this interpretation, computational linguistic knowledge must be represented in the left hemisphere regardless of familial handedness because it is so heavily dependent on relational processing. But associative and referential linguistic knowledge can be represented more widely when there is less general asymmetry between the hemispheres.

On this view, left-handed familials can access their associative and conceptual knowledge about individual words with associative processes relatively distinct from grammatical processing: Accordingly, left-handed familials can make more distinct use of associative information connected to individual words, separate from their role in grammatical relation to other words. The research on familial handedness discussed so far is limited to the language domain. The fact that there is consistent variation in how different aspects of linguistic knowledge are accessed demonstrates that the neurological foundations for a linguistic module are not monolithic. However, it remains to be shown that familial handedness mediates behavioral differences in local versus global processing in non-linguistic domains. I return to this later.

## BIOLOGICAL VARIATION IN MODES OF ABDUCTION

The previous two case studies involve distinctions between component mental processes (direct vs. relational) and levels of detail of representations (local-associative versus global–syntactic). A third formal distinction speaks to the way in which humans go about forming abstract concepts, the process Pierce (1957) referred to as *abduction*. Abduction occurs in simple conceptual activities like learning a simple concept, such as "dog"; it also occurs in the learning of more complex abstract structures, such as the grammar of a language or a mental map of a neighborhood. Abduction is the process (neither induction nor deduction) through which internally generated hypotheses interact with data as a person arrives at a correct abstract structure. Pierce noted that there are two contrasting forms of the interaction of data and internalized abstract hypotheses: In one method, new data can be used to refine an

existing hypothesis; in the other, the data can be used to choose between competing alternative hypotheses. Each of these methods can be reduced to an extreme form of the other, but they represent important differences in focus and emphasis. Consider a child developing the concept of what a "dog" is as opposed to a "cat," when presented for the first time with a small dog, such as a Chihuahua. Hypothesis refinement would extend the previous generalization of the distinction between "dogs" and "cats." Hypothesis competition would replace the previous hypothesis (e.g., "cats are small, dogs are big"), with a new one (e.g., "cats meow, dogs bark"). The net result of both procedures is ultimately the same, but hypothesis refinement involves changing a single internal hypothesis, whereas hypothesis competition involves generating a new hypothesis and choosing one of them.

The distinction between the two kinds of abduction raises an interesting question concerning the learning of grammar. Do children emerge with a correct grammar by successive hypothesis refinement or by hypothesis competition? Linguists tend to assume the latter because grammars appear to them to be too com… …x to be formed …; psychologists find hypothesis refinement more compatible with the notion of "learning." It would seem likely that grammar acquisition should rest on the use of both kinds of abductive strategies, but first one must show that the formal distinction corresponds to a mental one. Following the general thesis in this chapter, if we can show that language may be learned by different strategies that correlate with population variables, it will suggest that the abductive procedure for language learning is not univocal within a single language-learning module. Rather, it rests on a general difference in abduction style that itself may cut across different cognitive domains.

To study this, I and my colleagues used the study of population differences in the same way described in the previous cases: We designed a situation in which hypothesis refinement would be a more natural way of acquiring an abstract structure, in contrast with a situation in which hypothesis competition would be more natural. We then found that performance in the two different structure learning conditions indeed corresponded to a biologically coded difference in the subject population. We take this to be a demonstration of the construct validity of the distinction, as well as a result that leads to new ways of thinking about the formation of abstract representations. [1]

We used an artificial language learning study, taken in part from the

---

[1] The new experiments reported in this chapter are all being presented more fully in other publications. All the results reported are statistically significant on standard statistical measures. Co-workers on these studies include, Dustin Gordon, Ralph Hansen, Pietro Michelucci, Ken Shenkman.

literature on what structural properties of languages make them easy to learn. The language we used was a verb-final language, close to that used in studies of artificial language by Meier and Bower (1986). This language allows for the description of arrays of geometric figures displayed in quadrants of a computer screen. The language uses English lexical items, but its own grammatical order, and it has one grammatical morpheme, "te," which is roughly a relative clause introducer. Typical sentences in the language are:

1. triangle circle striped below
   (the triangle is below the striped circle)
2. triangle striped te circle above square left-of
   (the striped triangle that is above the circle is to the left of the square)

Previous studies of artificial language learning typically present subjects with sequences and ask them to learn to discriminate between those that are grammatical in the language and those that are not. This seemed to be unnecessarily artificial: Natural language learners are learning to use a language, to understand and be understood in the language. Accordingly we changed the usual procedure by focusing on situations that force subjects to learn to use the language. For example, in the "production" task, subjects were trained to take as input a geometric array, and produce as output a grammatical sequence that describes that array. The subject was provided with a button for each "word" (e.g., "blue," "triangle," "right-of"); on each trial, the subject typed a sequence as describing the geometric array presented on that trial. If the sequence was correct, he or she was told that; if the sequence was incorrect, the subject was shown a sample of correct sequences. There were 48 different trials with this procedure. Every 12 trials, the subject was stopped and given a forced-choice grammaticality judgment test, in which he or she was to indicate which member of each of 12 pairs of sequences was "grammatically correct" in the test language. These pairs were constructed to give information about a range of structural types of sentences.

A separate procedure involved learning "perception" of the language, in which subjects were given a correct sequence as input and were asked to "draw" the geometric array that the sequence described (we pre-trained subjects in the use of a simple graphics package, designed for this purpose). In this case, the geometric arrays were exactly those used as input to the production version of the study; also, every 12 trials, subjects responded to the "grammaticality" judgment task.

In a third version of the study, subjects were given "two-way" training both perception and production training, alternating from trial to trial. Every 12 trials, subjects responded to the "grammaticality" task. Our initial expectation was that subjects would learn the grammar of the language best in the task that alternated between production and perception because it most closely approximates the child's situation in learning a natural language. At the same time, we noted a fundamental conceptual difference related to the two kinds of abduction, between the two-way condition and the other conditions, each of which was "one-way." The two-way condition naturally stimulated subjects to build up separate mental systems to relate meanings with sequences, one for production that maps meanings onto sequences, and one for perception, which maps sequences onto meanings. If it had any effect at all, two-way training would tend to emphasize the hypothesis competition model of learning, at least because it stimulates subjects to develop separate models that relate meaning and sequence. Each of the one-way conditions would tend more to emphasize the hypothesis-refinement strategy of learning because only one model of the meaning–utterance relationships is required. Hence, we were in a position to find a contrast between subject groups that find hypothesis competition more conducive to learning an abstract structure, and subject groups that find hypothesis refinement more conducive to learning.

To this end, we controlled subject groups stringently on a number of variables. First, they all fell within the same narrow range of verbal and math SAT scores, age, and socioeconomic status. They were all native speakers of English. Gender and handedness background were part of the experimental design, so that we could examine the effects of each. There were 144 subjects who were run in slightly varying versions of the study. The dependent measures were success in grammaticality judgments, and success in the actual mapping tasks.

The results were complex and rich, and I discuss only part of them here, primarily those having to do with the grammaticality judgments, because, *prima facie*, they reflect the formation of an abstract grammar. There was a clear gender difference in the impact of two- versus one-way training. Males learned to discriminate the grammaticality task about 10% better in the two-way condition, than the one-way conditions; females show the exactly opposite pattern. These effects are quite large and statistically reliable. They are not due to differences in handedness background (which were balanced) or SAT scores (which were restricted in range, but which we also co-varied in our statistical analyses).

The first impact of these differences is to support the construct validity of the distinction between the two strategies for abduction. The results suggest that males are more likely to form an abstract structure when

learning distinct hypotheses that bear on that structure. Females are more likely to form an abstract structure when learning a single model that applies the structure. Another way of describing the difference is that females do best at abstracting a structure out of a single model for its use, whereas males do best at abstracting a structure from competing models for its use. Gender is a biologically coded difference, and it would be tempting to argue that this finding shows that the different kinds of abduction may themselves be biologically coded. However, unlike familial handedness, gender is a biological variable with immediate social consequences. It is possible that the different social roles for males and females are expressed in the formation of distinct learning strategies. It is difficult to formulate a specific hypothesis as to how such a socially based differentiation does, or could occur, but the complexity of the social patterns makes us cautious at this point about claiming that the abduction difference is directly biologically caused. The data in the next sections of this chapter strengthens the biological interpretation of these differences.

## SPATIAL LEARNING: TWO WAYS TO GET THERE FROM HERE

The strength of these results made us turn to other areas in which the cognitive performance of the genders might differ in a way related to the different styles of abduction. Both anecdotes and research have supported the claim that women are not as adept as men at reading conventional maps (Maccoby & Jacklin, 1974). Yet there is no evidence that this reflects or causes a fundamental difference in the ability to negotiate familiar and novel neighborhoods. Our first study is a laboratory demonstration of these two observations. We asked 20 men and 20 women to learn to get from one point to another in a set of corridors arranged in a digital-style figure 8. The subjects were started at a colored flag in the middle of a corridor at the top or bottom of the figure 8, and told to wander through the system until they found the differently colored goal flag; the goal flag was located at the middle of the corridor at the opposite end of the figure 8. Subjects were always started in one clockwise or counterclockwise direction by the presence of a "barrier" to the left or the right of their start position (on alternate trials, the starting barrier was on different sides). Several configurations of additional barriers were arranged so that the middle cross corridor of the figure 8 always turned out to be a blind alley: Subjects could turn into the middle corridor, but discovered barriers as soon as they exited at the other end of the middle corridor. Thus, what subjects had to learn was to walk on the perimeter of the figure 8, starting in the direction allowed by

the initial barrier, and to avoid the middle alley. Between each trial, subjects were blindfolded and brought to their starting flag in a wheelchair. Subjects were run for four trials. At the end of the experiment, we asked subjects to choose from a set of 8 prepared stick drawings, the one that depicted the exact arrangement of the corridors and flags (excluding barriers).

Males were significantly better at choosing the correct drawing of the corridor arrangement (87.5% correct for males, 25% correct for females). Overall running performance on the maze task, however, was identical, with respect to average number of turns made and the total time taken on each trial of the maze. This is consistent with the apparent fact that males and females can learn to negotiate a new neighborhood equally well. If that is true, why are the two genders reported to be so different at using a configurational representation of that neighborhood?

Recent investigations of the cognitive components of mental maps in humans and animals casts light on this question. The most consistent use of mental maps in nature is to provide a mechanism that allows an individual to leave a home base and return to it as directly as possible (Gallistel, 1990). For example, a bee leaves its hive and may forage in an irregular path until it finds some food. At that point it makes a beeline back to its hive rather than retracing its exact steps. This can occur when the home base is well out of sight. This clearly shows that the animal has an ability to "dead-reckon," to calculate from local cues where it is in relation to its home base. This computational ability resolves into two kinds of abstract representation: (a) episodic knowledge of individual landmarks and how they look from from different angles in relation to the home base, and (b) configurational knowledge of the spatial relationship between the landmarks. Each of these kinds of knowledge is sufficient in the limit to represent in behavior knowledge of an area: An exhaustive list of landmarks and their angular relation to home base can allow for navigation within the area, as can a list of landmarks and their angular relations to each other and home base. However, although extensionally equivalent, these two kinds of map representations emphasize different aspects of spatial knowledge.

A map is an abstract representation of a space independent of how that knowledge is used—it serves for spatial behavior the same representational function as that served by a grammar for language behavior. Furthermore, how one might best learn the two kinds of spatial knowledge corresponds to the two forms of abduction. Consider the impact on one's current spatial knowledge of an additional landmark location. Learning by hypothesis refinement is naturally suited to respond to the new information, by simply adding it to the existing list of landmarks. Learning through hypothesis competition is better suited to

the exchange of active sets of configurations: If new information to a spatial array is represented in terms of angular relations between locations/landmarks, then adding a new location to elaborate an existing array involves learning a complex set of new angular relations. It is a more direct method to replace the old array of locations with a new one—that is, to replace the previous map with a competitor.

I am arguing that the theoretical differentiation of two computational components of mental maps is like the differentiation of the two modes of abduction. This might resolve the difference in spatial abilities between males and females to the same source as that in artificial language learning. In particular, females may access episodic landmark-based spatial knowledge compared with males, relatively more easily than they access configurational knowledge. If the computational distinction is coded genetically in this way, it serves as a confirmation of the construct validity of the distinction. It suggests that the distinction is genetically represented; it offers an explanation of gender differences in using particular kinds of maps. The remainder of this section is directed at showing the gender difference indeed corresponds to the computational distinction between the two kinds of information involved in using mental maps.

Prior research has shown that females recognize photographs of areas in known neighborhoods better than males (Golledge, 1988). This is consistent with the hypothesis that females have richer access to landmarks from different perspectives in their home territory. The corridors and corners in our study using the figure 8 maze were visually similar, so we did not use such a test. Instead, we varied the reliability of cues for deciding what turns to make by running some subjects in opposite directions on different trials. That is, half the subjects were returned in the wheelchair to their original starting flag for each trial, and always learned the maze in one direction; the other half of the subjects were brought in the wheelchair to what had just been the goal flag, which was now treated as the start flag for the next trial. Accordingly, the latter subjects were exposed to the maze, going in both directions. Because the maze was symmetrical, the actual turns to be made were the same for both sets of subjects.

Our prediction was that the one-way condition would be easier than the two-way condition for females, whereas the opposite would be true for males. This prediction follows from the fact that in the one-way condition, each landmark had a univocal meaning about directions and turns to take; in the two-way condition each landmark has two opposite meanings, depending on one's running direction. Consequently, one-way training, in which landmarks have consistent meaning, should be easier for females, whereas two-way training, which discourages reliance on separate land-

marks alone, should be easier for males. The results were consistent with these predictions. Performance on the first trial revealed large individual differences in number of turns and total running time. Because the first trial was the same for all subjects, it served as a baseline to scale relative performance on the subsequent trials. There were three relative measures: (a) improvement in number of turns made; (b) improvement in time taken; and (c) improvement in the time for each turn. These measures show that females' performance actually deteriorated during the study in the two-way condition, while improving on several measures in the one-way condition. Males' performance became slightly better in the two-way than one-way condition on all measures. (The interaction between running condition and gender was significant on various measures.)

In the corridor study, subjects were free to extract a wide variety of individually idiosyncratic information about local cues. To gain more experimental control over such variables, we implemented a computerized version of the figure 8 maze. In this maze, the walls and floor are displayed in a forced three-dimensional projection on a computer screen; the walls and floor are scored with lines to enhance the perspective effects of movement, but there were no local cues to differentiate one choice point or wall from another. The only unique markings were a dark triangle and a light triangle on the floor in the middle of the top and bottom corridors of the figure 8, corresponding to the start and goal flags in the original study. Subjects could see these triangles only when they were three or fewer steps from there. Subjects stepped forward, backward, right, and left, using the standard cursor movement arrows on a computer keyboard. The design was as before, crossing gender of the subject and one-way versus two-way running conditions. We ran 48 subjects on this paradigm. The results confirmed the corridor study. Females performed better in the one-way than two-way running conditions. We ran 48 subjects on this paradigm. The results confirmed the corridor study. Females performed better in the one-way than two-way condition, whereas males performed better in the two-way than one-way condition. This is consistent with our hypothesis and confirms the reliability of the original corridor study. Most impressive in this case is the fact that all choice points in the maze are visually identical for all subjects (the start and goal diamonds on the floor receded from sight before the first corner is reached). This means that the effect of the one-way/two-way variable on choice points was entirely a function of what each subject thought he or she was doing, not the actual visual configuration.

The conclusion from these studies is first a confirmation of the construct validity of the different mechanisms involved in spatial learning.

The conceptual distinction between relying on the configural location of local cues, and relying on a more global map is reflected in the differential performance of males and females: A genetic difference corresponds to a difference in cognitive style. The mechanism for the transmission of this difference to individuals is unclear: As in the case of the difference in style of learning an artificial language, we do not now whether the difference is a direct function of neurological differences, or one mediated through social interactions. In this case, however, we can consider whether there are corresponding differences in other mammals. In fact, it is the case that female and male rats differ in the kinds of cues they use in spatial learning. In a series of elegant studies, Williams has shown that female rats rely on nearby cues when learning a maze, in contrast, male rats rely more on configurational cues (Williams & Meck, 1991).

We studied this difference in rats further, using a real figure 8 maze, similar to the one we used with humans. As with humans, we had a one-way procedure, in which the animal always ran in the same direction; we contrasted this with a two-way procedure, in which the animal ran in both directions, on different trials. The learning performance of male and female rats in these mazes has the same relative difference as what we found in humans: Male rats perform relatively better right away on the two-way maze; female rats initially perform better on the one-way maze. The neurological basis for such differences awaits further research: It is certainly the case that male and female neonatal brains differ in the distribution and relevance of hormonal receptors (Christensen & Gorski, 1978; Clark, MacLusky, & Goldman-Rakic, 1988; Dawson, 1972; Dawson, Cheung, & Lau, 1975; Juraska, 1984; Linn & Petersen, 1986; Maccoby & Jacklin, 1974; MacLusky, Clark, Naftolin, & Goldman-Rakic, 1987; McGlone, 1980; Plaff, 1980). One can also construct after-the-fact sociobiological stories to "account" for the females' relative dependence on local cues and the males' on configurational knowledge—given that there is a gender difference at all in spatial learning (Gaulin & Hoffman, 1987). For our purposes here, however, the difference in rats serves the same validating function as does the existence of cerebral asymmetries in nonhuman mammals. It suggests that the difference found in humans reflects a fundamental biological property of mammals in general.

## THE RELATION BETWEEN ABDUCTION AND SPATIAL MECHANISMS

Returning to the data from humans, the question arises as to whether there is more than an abstract theoretical relation between the two modes of abduction and the two corresponding ways of forming mental repre-

sentations of spatial knowledge. The finding that there are corresponding gender differences in each domain is suggestive that there is a common basis, but requires more elaboration. We have further evidence on this point from an analysis of the relation between performance and handedness background in relation to gender. Our interpretation of the handedness background variable as expressed in language behavior was that it differentiates local/lexical processing from global/syntactic processing. That is, right-handers from left-handed families (LHFs) tend to process natural language in smaller units than right-handers without left-handed family members (RHFs).

Suppose this difference generalized to other tasks, in particular artificial language learning and spatial learning, what would one predict? If LHFs divide problems into smaller subunits they should tend to exaggerate the effects of other variables (i.e., LHFs should show larger gender differences than RHFs on both the artificial language learning task and the spatial learning tasks). This is exactly what our initial results suggest: The one-way/two-way learning difference between males and females was larger in both kinds of tasks.

Accordingly, the previous two cases suggest that there may be some intrinsic variable that underlies a dimension for both the abstraction of a grammar and the formation of a mental map. Of course, the apparent similarity of the gender variation in both kinds of learning may be misleading: It could be due to a third factor, which itself is related to gender, whereas the learning variations are not. Further research is necessary on this point, in particular to test subjects on both kinds of tasks, as well as other assessments of variables that might be relevant.

In the meantime, however, we can speculate about the bases for the co-variation with gender in the two kinds of learning. One possibility is that different styles in abduction underlie the formation of abstractions, be they in language or spatial representations. The alternative is that spatial representation underlies certain aspects of the representation of abstract systems in general. This hypothesis has been recently advanced in a number of different forms (most notably in the writings of Roger Shepard, 1984; see also Talmy, 1988). Shepard, for example, argued that there has not been enough time for the evolutionary development of a new set of mechanisms for advanced skills such as music, and suggests that language as well, may make use of spatial mechanisms. These arguments are buttressed by consideration of the spatial basis of space terms such as prepositions.

It is hardly surprising to learn that the references of spatial words are conditioned by the mental organization of spatial knowledge. And it is hard to see how that confluence could explain our discovery that gender seems to correspond to similar kinds of variation in learning grammars

and maps. Elissa Newport and I are pursuing a more structural interpretation of the relation between spatial mechanisms and grammar that would explain this relation—in particular, we are developing the hypothesis that spatial mechanisms of representation are deployed in the representation of grammatical mechanisms directly. For example, we are investigating the possibility that laws governing apparent motion are structurally homologous to linguistic constraints on long-distance dependencies in language. This and other data may become the basis for the claim that human abstract knowledge in general recruits spatial representational mechanisms (see also Deane, 1990).

## CONCLUSION AND SUMMARY

In this chapter, I have reviewed some different interpretations of the concept of *modularity*. I pointed out that modular separation of capacities rooted in sensory or motor systems is not controversial. The modularization of cognitive skills such as language is controversial just because their neurological roots are difficult to determine. Furthermore, insofar as such skills assume the interaction of distinct kinds of levels of representation, impenetrability between levels occurs by definition. This follows from the fact that each level has its own internal computational language that is immiscible with that of other levels. Hence, experimental investigations may show modular-like segregation of levels of representation, results that would follow from their informational opacity, not necessarily from architectural segregation.

The lack of an unambiguous argument for the innately determined architectural modularity of skills like language, opens up the possibility of finding more general sources of modalities and their internal structure. One source is developmental. At the outset, postnatal sensorimotor mechanisms regulate much behavior in ways that anticipate normal adult categories of behavior: This has the natural consequence that immediate postnatal cortical experience is packaged in arrays that correspond to adult modalities. This has the natural consequence of shaping and organizing early experience for an otherwise undifferentiated cortex in ways that lead to learning adult modalities. Thus, the basic segregation of natural kinds of knowledge may rest on innate subcortical mechanisms, not specific pre-tuned cortical areas.

A second source of structure within apparent modalities follows from the view that cognitive processes are arrayed in a set of pairs, along dimensions that themselves are innately determined. Processes can be paired in a general sense according to whether they involve single or multiple representations. The interaction of this pairing with different

aspects of language define pairs of representational and processing options that correspond to variables found in nature. Thus, the contrast between relational and unary processing is coordinated in a general way with the difference between the left and right hemispheres, which may explain why language is ordinarily primarily represented in the left hemisphere. The contrast between a strategy of accessing global computational and local associative information about words is correlated with familial handedness, which may reflect a relative symmetry of the representation of associative information about language in people with left-handed families.

The correlation of the previous two contrasts with general biological variation gives them construct validity, and also allows for direct speculation as to how the genetic basis for the dimensions and variation within them lie outside the capacity for language. In addition, our studies suggest that learning an abstract system such as a grammar calls upon a general system of abduction, which can vary in emphasis across a biological dimension—gender. The existence of this systematic variation again lends validity to the construct, and suggests that the ability to discover grammar calls on abduction capacities not limited to language. A similar gender differentiation in how spatial learning in humans and animals occurs, suggests (but only suggests) that the difference in abduction strategies may be quite general. This in turn, supports the argument that the abduction of grammar in particular calls on general learning mechanisms not unique to language.[2]

In brief, I am arguing that there are regularities in general cognition, independent of modality, and that those regularities are genetically coded. I have isolated three cognitive dimensions along which biologically

---

[2] The study of cognitive differences between genetically coded groups in the normal population is just one method to reveal cognitive distinctions that correspond to genetically coded neurophysiological differences. This method must be used with great social and political responsibility: The goal is not to 'scientifically' define cognitive differences in general abilities, which are based in political and social facts. Rather, the goal is to partition general capacities into natural subsets that can differ along genetic lines. As I show in this chapter, careful attention to the distinct multiple cognitive components in an apparently undifferentiated behavior can actually explain an apparent group difference in ability as the result of a group difference in cognitive style. Thus, our research has shown that the previously alleged superiority of males in spatial learning may be a function of how the learning task is presented.

As our theories of the mind and the brain improve, it is not impossible that we will find an increasing number of cases of difference in cognitive style associated with biologically defined differences in populations. We must squarely face the possibility that such discoveries, like all of science, can be misused for political ends: In particular, differences associated with biologically defined groups could be used to justify a form of scientific racism. I hope I have shown by example in this chapter, that such political misuse of population differences is also scientifically corrupt for several reasons.

based populations differ. I have suggested possible mechanisms that might underlie the bases for each of those dimensions. The hypothesis that there are modules for natural kinds of complex behavior is not imperiled by such proposals, but it must be somewhat modified (Shallice, 1984). Stipulating the framework I have outlined for cognition in general, there still may be specific neurological mechanisms that are responsible for recruiting general cognitive capacities in ways specific to each modality. I have suggested the postnatal instinctive mechanisms are certainly such mechanisms: Nothing I have said demonstrates that there are not corresponding higher-order mechanisms as well.

My focus here has been to outline a research program, not to claim its definitive success. Little has been conclusively proven. I have outlined a program aimed at exploring how a complex domain such as language could come to be a modality without resulting from an innately architecturally delimited cortical module. The goal at the least is to show which aspects of language can be explained from general principles and facts about behavior and learning: Those that cannot be so explained become the basis for more specific hypotheses about what is truly uniquely innate to language.

---

First, while group differences in cognitive style can be isolated experimentally, there is no direct implication for the social role of such distinctions. There are atomic differences between steel and aluminum, but either can be the material for an excellent armchair. Groups may differ in the way they tend to solve a problem, but have functionally equal abilities.

Second, the range of differences between groups is generally the same along each cognitive dimension: The difference lies in a group average. Such differences are scientifically instructive, informing us as to how cognitive dimensions are differentiable biologically, in specially constructed experimental situations. But they are socially useless distinctions because they tell us nothing about an individual.

Third, racism is not a scientific concept and can never be supported by scientific investigation. Racism is a political doctrine. In this regard, claiming that one biologically defined group is politically different from another is like the claim that a republic is better than a democracy: There is no scientific domain in which to test such a claim.

Finally, in fact, if science speaks to racism at all, it is to discredit it and show further that it is political in nature. Science can offer us correct ways to redress uniformed racism: Insofar as we discover cognitive differences in biologically defined groups, we will be better able to decide how to make those differences socially and politically unimportant

## REFERENCES

Baillargeon, R., Spelke, E. S. & Wasserman, S. (1985). Object permanence in five month old infants. *Cognition, 20,* 191–208.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and language development* (pp. 277–360). New York: Wiley.

Bever, T. G. (1975) Cerebral asymmetries in humans are due to the differentiation of two incompatible processing mechanisms: Holistic and analytic. In D. Aaronson & R. Rieber (Eds.), *Developmental psycholinguistics and communication disorders* (pp. 76–86). New York: New York Academy of Sciences.

Bever, T. G. (1980). Broca and Lashley were right: Cerebral dominance is an accident of growth. In D. Kaplan & N. Chomsky (Eds.). *Biology and language* (pp. 186–232). Cambridge, MA: MIT Press.

Bever, T. G. (1983). Cerebral lateralization, cognitive asymmetry and human consciousness. In E. Perecman (Ed.), *Cognitive processing in the right hemisphere* (pp. 1939). New York: Academic Press.

Bever, T. G., Carrithers, C., Cowart, W., & Townsend, D. J. (1989). Language processing and familial handedness. In A. Galaburda (Ed.), *From neurons to reading* (pp. 331–360). Cambridge, MA: MIT Press.

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and non musicians. *Science, 185,* 1317–139.

Bever, T. G., Straub, K., Shenkman, K., Kim, J. J., & Carrithers, C. (1990). The psychological reality of NP-trace. *Proceedings of the 1989 Meeting of the Northeast Linguistic Society,* 48–64.

Brown, J. W. (1976). The neural organization of language: Aphasia and lateralization. *Brain and Language, 3,* 482.

Brown, J. W. (1978). Lateralization: A brain model. *Brain and Language, 5,* 258.

Brown, J. W., & Hecaen, H. (1976). Lateralization and language representation. *Neurology, 26,* 183.

Calvin, W. H., (1982). Didn't throwing stones shape HOMINID evolution. *Sociobiology, 3,* 115–124.

Carey, S. (1987). Conceptual change in childhood. Cambridge, MA: Bradford Press.

Carrithers, C. (1989). Syntactic complexity does not necessarily make sentences harder to understand. *Journal of Psycholinguistics Research, 18,* 75–88.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Christensen, L. W., & Gorski, R. A. (1978). Independent masculinization of neuroendocrine systems by intercerebral implants of testosterone or estradiol in the neonatal rate. *Brain Research, 146,* 325–340.

Clark, A. S., Maclusky N. J., & Goldman-Rakic, P. S. (1988). Androgen binding and metabolism in the cerebral cortex of the developing rhesus monkey, *Endocrinology, 123,* 932–940.

Cowart, W. (1982). Autonomy and interaction in the language processing system: A reply to Marslen-Wilson and Tyler. *Cognition, 12*(1), 109–117.

Dawson, J. L. M. (1972). Effects of sex hormones on cognitive style in rats and men. *Behavioral Genetics, 2,* 21–42.

Dawson, J. L. M., Cheung, Y. M., & Lau, R. T. S. (1975). Developmental effects of neonatal sex hormones on spatial and activity skills in the white rate. *Biological Psychology, 3,* 213–229.

Deane, P. D. (1990). *Grammar in mind and brain: Explorations in cognitive syntax.* Unpublished manuscript, University of Central Florida, Orlando.

Dennis, M., & Kohn, B. (1974). Selective impairments of visuo-spatial abilities in infantile

hemiplegics after right cerebral hemidecortication. *Neuropsychologia, 12*(4), 505512.

Dennis, M., & Kohn, B. (1975). Comprehension of syntax in infantile hemiplegics after cerebral hemidecortication: Left-hemisphere superiority. *Brain and Language, 2*(4), 472–482.

Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology.* Cambridge, MA: MIT Press.

Gallistel, C. R. (9190). *The organization of learning.* Cambridge, MA: MIT Press.

Gaulin, S. J. C., & Hoffman, H. A. (1987). Evolution and development of sex differences in spatial ability. In M. Borgerhoff-Mulder, L. Betzig, P. Turke (Eds.L), *Human reproductive behavior. A Darwinian perspective* (pp. 129–152). Cambridge, JA: Cambridge University Press.

Gazzaniga, M. S., & Hillyard, S. A. (1971). Language speech capacity of the right hemisphere. *Neuropsychology, 9*(3), 273–280.

Gazzaniga, M. D., & Sperry, R. (1967). Language after section of the cerebral commisures. *Brain, 90,* 131–148.

Geschwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization.* Cambridge, MA: MIT Press.

Gleitman, L. R. (1981). Maturational determinism, *Cognition, 10,* 103–114.

Golledge, R. G. (1988). *Integrating spatial knowledge.* Santa Barbara: Geographical Press.

Hecaen, H. (1976). Acquired aphasia in children and ontogenesis of hemispheric functional specialization. *Brain and Language, 3,* 114.

Herman, L. M., Richard, D. G., & Wholz, J. P. (1984). Comprehension of Sentences by Bottle-nosed Dolphins. *Cognition, 16,* 129–219.

Herman, L. M., Morrell-Samuels, P., & Pack, A. A. (1990). Bottlenosed colphon and human recognition of veridical and degraded video displays of an artificial gestural language. *Journal of Experimental Psychology: General, 119,* 215–230.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology, 14,* 5–24.

Hurtig, R. (1982). Cerebral asymmetry in the strategies used in processing random line stimuli. *Cortex, 18,* 337–343.

Joanette, Y., Lecours, A. R., LePage, Y., & Lamoureux, M. (1983). Language in Right-Handers with Right Hemispheric Lesions: A preliminary study including anatomical, genetic and social factors. *Brain and Language, 20,* 217–248.

Juraska, J. A. (1984). Sex differences in developmental plasticity in visual cortex and hippocampal dentate gyrus. *Progressive rain Research, 61,* 205–214.

Keil, F. C. (1989). *Concepts, kinds and cognitive development.* Cambridge, MA: Bradford Press.

Kellar, L. A., & Bever, T. G. (1980). Hemispheric asymmetries in the perception of musical intervals as a function of musical experience and family handedness background. *Brain and Language, 10,* 24–38.

Kimura, D. (1964). Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology, 16,* 355–358.

Kimura, D. (1973). The asymmetry of the human brain. *Scientific American, 228,* 70–78.

Kimura, D. (1976). Intrahemispheric interaction between speaking and sequential manual activity. *Neuropsychologia, 14*(1), 2333.

Kosslyn, S. (1987). Seeing and imaging in the cerebral hemispheres: A computational approach. *Psychological Review, 94*(2), 148–175.

Kutas, M., Van Petten, C., Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology, 69,* 218–233.

Le Doux, J. E. (1983). Cerebral asymmetry and the integrated function of the brain. In A. W. Young (Ed.), *Functions of the right cerebral hemisphere* (pp. 203–216). New York: Academic Press.

Lenneberg, E. H. (1967). *Biological foundations of language.* New York: Wiley.

Leslie, A. M. (1987). Pretense and representation: The origins of "Theory of Mind." *Psychological Review, 74,* 412–426.

Levy, J. (1969). Possible basis for the evolution of lateral specialization. *Nature, 21,* 614–615.

Levy, J. (1988). Individual differences and cerebral hemisphere asymmetries; theoretical issues and experimental considerations. In H. Jerison & I. Jerison (Eds.), *Intelligence and evolutionary biology* (pp. 157–173). New York: Springer-Verlag.

Linn, M. C., & Petersen, A. C. (1986). A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement. In. J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender* (pp. 67–101). Baltimore, MD: Johns Hopkins University Press.

Luria, A. R. (1947). *Traumatic aphasia: Its syndrome, psychopathology and treatment.* Moscow: Academy of Medical Sciences.

Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, A: Stanford University Press.

MacLusky, N. J., Clark A. S., Naftolin, F., & Goldman-Rakic, P. S. (1987). Estrogen formation in the mammalian brain: possible role of aromatase in sexual differentiation of the hippocampus and neocortex. *Steroids, 50,* 450–465.

Maratsos, (1982). In T. C. Regret in mental development. Hillsdale, NJ: Lawrence Erlbaum Associates.

Modularity in knowledge representation and natural language understanding (G. P.). Cambridge, MA: MIT Press.

McNeilage, P. F., Studdert-Kennedy, M. G., & Lindblom, B. (1987). Primate handedness reconsidered. *Behavioral and Brain Sciences, 10*(2), 247–303

McGlone, J. (1980). Sex differences in human brain asymmetry: A critical survey. *Behavioral Brain Science, 3,* 215–264.

Mehler, J., & Dupoux (1990). *Naitre Humain [Human nature].* Paris: LaRousse.

Mehler, J., Morton, J., & Jusczyk, P. W. (1984). On reducing language to biology. *Cognitive Neuropsychology, 1,* 83–116.

Meier, R. P., & Bower, G. H. (1986). Semantic reference and phrasal grouping in the acquisition of a miniature phrase structure language. *Journal of Memory and Language, 25,* 492–505.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants initiate facial gestures. *Child Development, 54,* 702–709.

Morrel-Samuels, P., Herman, L., & Bever, T. G. (1991). *The relation between processing demands and cerebral asymmetries in a dolphin.* Unpublished manuscript, University of Michigan, East Lansing.

Murray, L, & Trevarthen, C. (1986). The infant's role in mother-infant communications. *Journal of Child Language, 13*(1), 15–29.

Newman, J. E., Lovett, M. W., & Dennis, M. (1986). The use of discourse analysis in neurolinguistics: Some findings from the narratives of hemidecorticate adolescents. *Brain and Language, 30*(1), 63–80.

O'Connor, K. N., Roitblat, H. L., & Bever, T. G. (in press). Auditory sequence complexity and hemispheric asymmetries of function in rats. In H. L Roitblat (Ed.), *Study in animal behavior.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Pierce, C. S. (1957). The logic of abduction. In C.S. Pierre (Ed.), *Essays in the philosophy of Science* (pp. 235–255). Boston: Liberal Arts Press.

Pfaff, D. W. (1980). *Estrogens and brain function.* New York: Springer-Verlag.

Schwartz, M. F., & Schwartz, B. (1984). In defense of organology. *Cognitive Neuropsychology, 1,* 25–42.

Shallice, T. (1984). More functionally isolable subsystems but fewer "modules?" *Cognition, 17*(3), 24–252.

Shepard, R. N (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imaging, thinking, and dreaming. *Psychological Review, 91*(4), 417–447.

Subirana, A. (1958). The prognosis in aphasia in relation to cerebral dominance and handedness. *Brain,81,* 415.

Subirana, A. (1969). Handedness and cerebral dominance. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of clinical neurology* (Vol. 4, pp. 248–272). Amsterdam: North Holland.

Talmy L. (1988). Force dynamics in language and cognition. *Cognitive Science, 12*(1), 49–100.

Tan, N., Aiello, R., & Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Memory and Cognition, 9,* 533–539.

Townsend, D. J., & Bever, T. G. (1982). Natural units of representation interact during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 21,* 688–703.

Wellman, H. W. (1990). *The child's theory of mind.* Cambridge, MA: Bradford Press.

Werner, H. (1948). *Comparative psychology of mental development.* New York: International University Press.

Williams, C. L., & Meck, W. H. (1991). The organizational effects of gonadal steroids on sexually dimorhic spatial ability. *Psychoneuroendocrinology, 16*(1), 155–176.

## REFERENCE

Bever, T.G. 1992. The Logical and Extrinsic Sources of Modularity in M. Gunnar &
M. Maratsos (eds.) *Modularity and Constraints in Language and Cognition*, Vol 25 of
the Minnesota Symposia on Child Psychology, 1992. Lawrence Earlbaum and Associates

## FOOTNOTES

[1] The new experiments reported in this chapter are all being presented more fully in other publications. All the results reported are statistically significant on standard statistical measures. Co-workers on these studies include, Dustin Gordon, Ralph Hansen, Pietro Michelucci, Ken Shenkman.

[2] The study of cognitive differences between genetically coded groups in the normal population is just one method to reveal cognitive distinctions that correspond to genetically coded neurophysiological differences. This method must be used with great social and political responsibility: The goal is not to 'scientifically' define cognitive differences in general abilities, which are based in political and social facts. Rather, the goal is to partition general capacities into natural subsets that can differ along genetic lines. As I show in this chapter, careful attention to the distinct multiple cognitive components in an apparently undifferentiated behavior can actually explain an apparent group difference in ability as the result of a group difference in cognitive style. Thus, our research has shown that the previously alleged superiority of males in spatial learning may be a function of how the learning task is presented.

As our theories of the mind and the brain improve, it is not impossible that we will find an increasing number of cases of difference in cognitive style associated with biologically defined differences in populations. We must squarely face the possibility that such discoveries, like all of science, can be misused for political ends: In particular, differences associated with biologically defined groups could be used to justify a form of scientific racism. I hope I have shown by example in this chapter, that such political misuse of population differences is also scientifically corrupt for several reasons.

First, while group differences in cognitive style can be isolated experimentally, there is no direct implication for the social role of such distinctions. There are atomic differences between steel and aluminum, but either can be the material for an excellent armchair. Groups may differ in the way they tend to solve a problem, but have functionally equal abilities.

Second, the range of differences between groups is generally the same along each cognitive dimension: The difference lies in a group average. Such differences are scientifically instructive, informing us as to how cognitive dimensions are differentiable biologically, in specially constructed experimental situations. But they are socially useless distinctions because they tell us nothing about an individual.

Third, racism is not a scientific concept and can never be supported by scientific investigation. Racism is a political doctrine. In this regard, claiming that one biologically defined group is politically different from another is like the claim that a republic is better than a democracy: There is no scientific domain in which to test such a claim.

Finally, in fact, if science speaks to racism at all, it is to discredit it and show further that it is political in nature. Science can offer us correct ways to redress uniformed racism: Insofar as we discover cognitive differences in biologically defined groups, we will be better able to decide how to make those differences socially and politically unimportant