1999-01

Brain Learning, Attention, and Consciousness

Brain learning, attention, and consciousness

Adapted from "The link between brain learning, attention, and consciousness"
Consciousness and Cognition

Stephen Grossberg

January 1999

Technical Report CAS/CNS-99-002

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
677 Beacon Street
Boston, MA 02215

BRAIN LEARNING, ATTENTION, AND CONSCIOUSNESS

Adapted from "The Link Between Brain Learning, Attention, and Consciousness", *Consciousness and Cognition*

Stephen Grossberg†
Department of Cognitive and Neural Systems
and
Center for Adaptive Systems
Boston University
677 Beacon Street
Boston, MA 02215

Submitted: January, 1999

Technical Report CAS/CNS-TR-99-002

Keywords Learning, expectation, attention, adaptive resonance, neural network, procedural memory, consciousness, object recognition, speech perception.

---

† Supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409), National Science Foundation (IRI-97-20333), and the Office of Naval Research (ONR N00014-95-1-0657).

Acknowledgments: The author wishes to thank Cynthia E. Bradford, Diana Meyers and Robin Amos for their valuable assistance in the preparation of this manuscript.

Abstract

The processes whereby our brains continue to learn about a changing world in a stable fashion throughout life are proposed to lead to conscious experiences. These processes include the learning of top-down expectations, the matching of these expectations against bottom-up data, the focusing of attention upon the expected clusters of information, and the development of resonant states between bottom-up and top-down processes as they reach an attentive consensus between what is expected and what is there in the outside world. It is suggested that all conscious states in the brain are resonant states, and that these resonant states trigger learning of sensory and cognitive representations. The models which summarize these concepts are therefore called Adaptive Resonance Theory, or ART, models. Psychophysical and neurobiological data in support of ART are presented from early vision, visual object recognition, auditory streaming, variable-rate speech perception, somatosensory perception, and cognitive-emotional interactions, among others. It is noted that ART mechanisms seem to be operative at all levels of the visual system, and it is proposed how these mechanisms are realized by known laminar circuits of visual cortex. It is predicted that the same circuit realization of ART mechanisms will be found in the laminar circuits of all sensory and cognitive neocortex. Concepts and data are summarized concerning how some visual percepts may be visibly, or modally, perceived, whereas amodal percepts may be consciously recognized even though they are perceptually invisible. It is also suggested that sensory and cognitive processing in the What processing stream of the brain obey top-down matching and learning laws that are often complementary to those used for spatial and motor processing in the brain's Where processing stream. This enables our sensory and cognitive representations to maintain their stability as we learn more about the world, while allowing spatial and motor representations to forget learned maps and gains that are no longer appropriate as our bodies develop and grow from infanthood to adulthood. Procedural memories are proposed to be unconscious because the inhibitory matching process that supports these spatial and motor processes cannot lead to resonance.

How Do We Continue to Learn Throughout Life?

We experience the world as a whole. Although myriad signals relentlessly bombard our senses, we somehow integrate them into unified moments of conscious experience that cohere together despite their diversity. Because of the apparent unity and coherence of our awareness, we can develop a sense of self that can gradually mature with our experiences of the world. This capacity lies at the heart of our ability to function as intelligent beings.

The apparent unity and coherence of our experiences is all the more remarkable when we consider several properties of how the brain copes with the environmental events that it processes. First and foremost, these events are highly context-sensitive. When we look at a complex picture or scene as a whole, we can often recognize its objects and its meaning at a glance, as in the picture of a familiar face. However, if we process the face piece-by-piece, as through a small aperture, then its significance may be greatly degraded. To cope with this context-sensitivity, the brain typically processes pictures and other sense data in parallel, as *patterns* of activation across a large number of feature-sensitive nerve cells, or neurons. The same is true for senses other than vision, such as audition. If the sound of the word GO is altered by clipping off the vowel O, then the consonant G may sound like a chirp, quite unlike its sound as part of GO.

During vision, all the signals from a scene typically reach the photosensitive retinas of the eyes at essentially the same time, so parallel processing of all the scene's parts begins at the retina itself. During audition, each successive sound reaches the ear at a later time. Before an entire pattern of sounds, such as the word GO, can be processed as a whole, it needs to be recoded, at a later processing stage, into a simultaneously available spatial pattern of activation. Such a processing stage is often called a working memory, and the activations that it stores are often called short term memory (STM) traces. For example, when you hear an unfamiliar telephone number, you can temporarily store it in working memory while you walk over to the telephone and dial the number.

In order to determine which of these patterns represents familiar events and which do not, the brain matches these patterns against stored representations of previous experiences that have been acquired through learning. Unlike the STM traces that are stored in a working memory, the learned experiences are stored in long term memory (LTM) traces. One difference between STM and LTM traces concerns how they react to distractions. For example, if you are distracted by a loud noise before you dial a new telephone number, its STM representation can be rapidly reset so that you forget it. On the other hand, if you are distracted by a loud noise, you (hopefully) will not forget the LTM representation of your own name.

The problem of learning makes the unity of conscious experience particularly hard to understand, if only because we are able to rapidly learn such enormous amounts of new information, on our own, throughout life. For example, after seeing an exciting movie, we can tell our friends many details about it later on, even though the individual scenes flashed by very quickly. More generally, we can quickly learn about new environments, even if no one tells us how the rules of each environment differ. To a surprising degree, we can rapidly learn new facts without being forced to just as rapidly forget what we already know. As a result, we do not need to avoid going out into the world for fear that, in learning to recognize a new friend's face, we will suddenly forget our parents' faces.

I have called the problem whereby the brain learns quickly and stably without catastrophically forgetting its past knowledge the *stability-plasticity dilemma*. The stability-plasticity dilemma must be solved by every brain system that needs to rapidly and adaptively respond to the flood of signals that subserves even the most ordinary experiences. If the brain's design is parsimonious, then we should expect to find similar design principles operating in all the brain systems that can stably learn an accumulating knowledge base in response to changing conditions throughout life. The discovery of such principles should clarify how the brain unifies diverse sources of information into coherent moments of conscious experience.

This article reviews evidence that the brain does operate in this way. It summarizes several recent brain modeling studies that illustrate, and further develop, a theory called Adaptive Resonance Theory, or ART, that I introduced in 1976 (Grossberg, 1976a, 1976b, 1978, 1980, 1982). In the present article, I will briefly summarize results selected from four areas where ART principles have been used to explain challenging behavioral and brain data. These areas are visual perception, visual object recognition, auditory source identification, and variable-rate speech recognition. On first inspection, the behavioral properties of these visual and auditory phenomena may seem to be entirely unrelated. On a deeper computational level, their governing neural circuits are proposed to incorporate a similar set of computational principles.

I should also say right away, however, that ART principles do not seem to be used in all brain learning systems. Whereas ART learning designs help to explain sensory and cognitive processes such as perception, recognition, attention, reinforcement, recall, working memory, and memory search, other types of learning seem to govern spatial and motor processes. In these latter task domains, it *is* adaptive to forget old coordinate transformations as the brain's control systems adjust to a growing body and to other changes in the body's sensory-motor endowment throughout life.

Sensory and cognitive processes are often associated with the What cortical processing stream that passes from visual cortex through inferotemporal cortex, whereas spatial and motor processes are associated with the Where (or How) cortical processing stream that passes from visual cortex through parietal cortex (Goodale and Milner, 1992; Mishkin, Ungerleider, and Macko, 1983; Ungerleider and Mishkin, 1982). Our research over the years has concluded that many processes in the two distinct streams, notably their matching and learning processes, obey different, and even complementary, laws. This fact bears heavily on questions of consciousness, and helps to explain why procedural memories are not conscious (Cohen and Squire, 1980; Mishkin, 1982; Scoville and Milner, 1957; Squire and Cohen, 1984). Indeed, a central hypothesis of ART since its inception is

ART Hypothesis: All conscious states are resonant states.

As noted in greater detail below, many spatial and motor processes involve a form of inhibitory matching and mismatch-based learning that does not support resonant states. Hence, by the ART Hypothesis, they cannot support a conscious state. Although ART predicts that all conscious states are resonant states, the converse statement, that all resonant states are conscious states, is not yet asserted.

Various other models of cognitive learning and recognition, such as the popular back-propagation model (Parker, 1982; Rumelhart, Hinton, and Williams, 1986; Werbos, 1974), are based on a form of mismatch-based learning. They cannot, therefore, generate resonant states and, in fact, are well known to experience catastrophic forgetting under real-time learning conditions. A comparative survey of ART vs. backpropagation computational properties is provided in Grossberg (1988).

How Do We Perceive Illusory Contours and Brightness?

Let me start by providing several examples of the diverse phenomena that ART clarifies. Consider the images in Figure 1. Figure 1a shows an image called an Ehrenstein figure in which some radial black lines are drawn on a uniformly white paper. Remarkably, our minds construct a circular illusory contour that touches each line end at a perpendicular orientation. This illusory contour is a collective, emergent property of all the lines that only occurs when their positions relative to each other are suitable. For example, no illusory contour forms at the line ends in Figure 1b even though they end at the same positions as the lines in Figure 1a. Note also that the illusory contour in Figure 1a surrounds a disk that seems uniformly brighter than its surround. Where does the brightness enhancement come

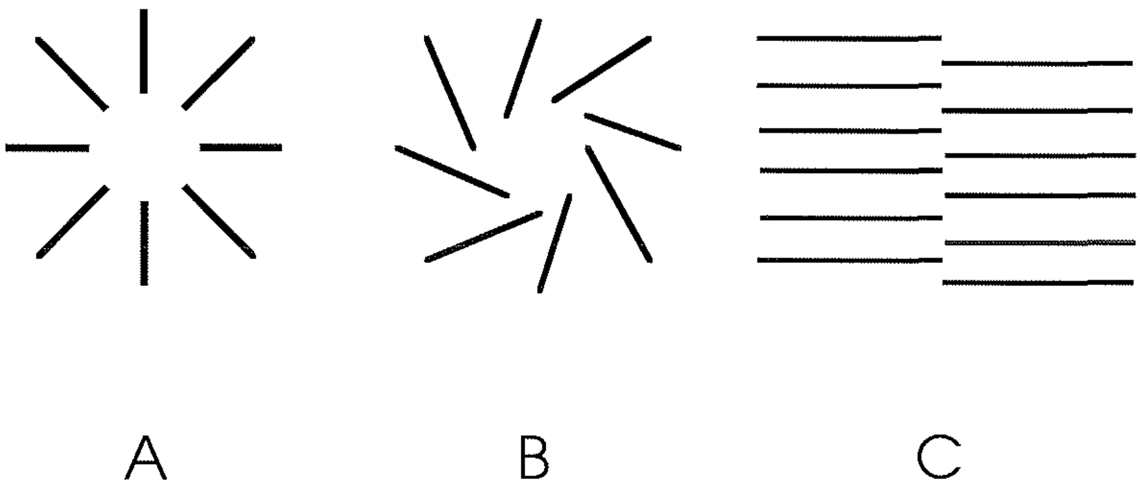


Figure 1. (A) The Ehrenstein pattern generates a circular illusory contour that encloses a circular disk of enhanced illusory brightness. (B) If the endpoints of the Ehrenstein pattern remain fixed while their orientations are tilted, then both the illusory contour and brightness vanish. (C) The offset pattern generates a vertical boundary that can be recognized even though it cannot be seen.

from? It certainly does not always happen when illusory contours form, as can be seen by inspecting Figure 1c. Here a vertical illusory contour can be recognized as interpolating the two sets of offset horizontal lines, even though neither side of the contour seems brighter than the other. How we can consciously *recognize* something that we cannot *see*, and is thus perceptually invisible, is a fascinating aspect of our conscious awareness about which quite a bit is now known. Such percepts are known as *amodal* percepts (Michotte, Thines, and Crabbe, 1964) in order to distinguish them from modal, or visible, percepts. Amodal percepts are experienced in response to many naturalistic scenes, notably in response to scenes in which some objects are partially occluded by other objects. How both modal and amodal percepts can occur is modeled in Grossberg (1994, 1997). Of particular interest from the viewpoint of ART processing is why the Ehrenstein disk looks bright, despite the fact that there are no local contrasts within the image itself that describe a disk-like object. Gove *et al.* (1995) provide an explanation of this illusion using ART mechanisms that are described below.

How Do We Learn to Recognize Visually Perceived Objects?

The Ehrenstein example concerns the process of visual perception. The next example concerns a process that goes on at a higher level of the visual system. It is the process whereby we visually recognize objects. A key part of this process concerns how we learn to categorize specific instances of an object, or set of objects, into a more general concept. For example, how do we learn that many different printed or script letter fonts can all represent the same letter A? Or how do we learn that several different combinations of patient symptoms are all due to the same disease? Moreover, how do we control how general our categories will become? For some purposes, like recognizing a particular face, we need highly specific categories. For others, like knowing that every person has a face, the categories are much more general. Finally, how does our learning and memory break down when something goes wrong in our brain? For example, it is known that lesions to the human hippocampal system can cause a form of amnesia whereby, among other properties, patients find it very hard to learn new information and hard to remember recently learned information, but previously learned information about which their memory has "consolidated" can readily be

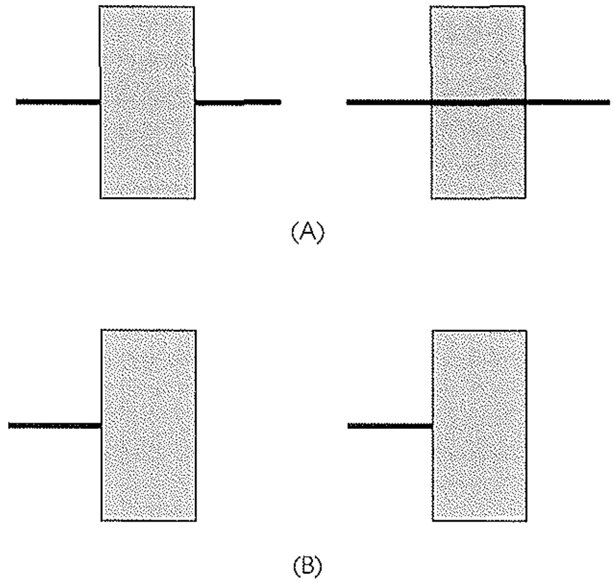


Figure 2. (A) Auditory continuity illusion: When a steady tone occurs both before and after a burst of noise, then under appropriate temporal and amplitude conditions, the tone is perceived to continue through the noise. (B) This does not occur if the noise is not followed by a tone.

retrieved. Thus, an amnesic patient can typically carry out a perfectly intelligent conversation about experiences that occurred a significant time before the lesion that caused the amnesia occurred.

What computational properties do the phenomena of bright illusory disks and amnesic memory have in common? I will suggest below that their apparent differences conceal the workings of a general unifying principle.

How Do We Solve the Cocktail Party Problem?

To continue with our list, let us now consider a different modality entirely; namely, audition. When we talk to a friend in a crowded noisy room, we can usually keep track of our conversation above the hubbub, even though the sounds emitted by the friendly voice may substantially overlap the sounds emitted by other speakers. How do we separate this jumbled mixture of sounds into distinct voices? This is often called the cocktail party problem. The same problem is solved whenever we listen to a symphony or other music wherein overlapping harmonic components are emitted by several instruments. If we could not separate the instruments or voices into distinct sources, or auditory streams, then we could not hear the music as music, or intelligently recognize a speaker's sounds. A striking and ubiquitous property of such percepts, and one which has not yet been understand by alternative modeling approaches, is how future events can alter our conscious percepts of past events in a context-sensitive manner.

A simple version of this competence is illustrated by the auditory continuity illusion (Bregman, 1990). Suppose that a steady tone shuts off just as a broadband noise turns on. Suppose, moreover, that the noise shuts off just as the tone turns on once again; see Figure 2A. When this happens under appropriate conditions, the tone seems to continue right through the noise, which seems to occur in a separate auditory "stream". This example shows that the auditory system can actively extract those components of the noise that are consistent with the tone and use them to track the "voice" of the tone right through the hubbub of the noise.

In order to appreciate how remarkable this property is, let us compare it with what happens when the tone does not turn on again for a second time, as in Figure 2B. Then the first tone does not seem to continue through the noise. It is perceived to stop before the noise. How does the brain know that the second tone will turn on after the noise shuts off, so that it can continue the tone through the noise, yet not continue the tone through the noise if the second tone does not eventually occur? Does this not seem to require that the brain can operate "backwards in time" to alter its decision as to whether or not to continue a past tone through the noise based on future events?

Many philosophers and scientists have puzzled about this sort of problem. I will argue that the process whereby we consciously hear the first tone takes some time to unfold, so that by the time we hear it, the second tone has already begun. To make this argument, we need to ask why does conscious audition take so long to occur after the actual sound energy reaches our brain? Just as important, why can the second tone influence the conscious percept so quickly, given that the first tone could not? Finally, I will indicate what these auditory phenomena have to do with bright Ehrenstein disks and amnesia.

How Do We Consciously Perceive Speech?

The final examples also involve the auditory system, but at a higher level of processing. They concern how we understand speech. In these examples, too, the process whereby conscious awareness occurs takes a long time, on the order of 100 milliseconds or more. An analysis of these percepts will also give us more clues about the nature of the underlying process. The first example is called phonemic restoration. Suppose that a listener hears a noise followed immediately by the words "eel is on the ...". If this string of words is followed by the word "orange", then "noise-eel" sounds like "peel". If the word "wagon" completes the sentence, then "noise-eel" sounds like "wheel". If the final word is "shoe", then "noise-eel" sounds like "heel".

This marvelous example, which was developed by Richard Warren and his colleagues more than twenty years ago (Warren, 1984; Warren and Sherman, 1974), vividly shows that the bottom-up occurrence of the noise is not sufficient for us to hear it. Somehow the sound that we *expect* to hear based upon our previous language experiences influences what we do hear, at least if the sentence is said quickly enough. As in the auditory continuity illusion, it would appear that the brain is working "backwards in time" to allow the meaning imparted by a later word to alter the sounds that we consciously perceive in an earlier word.

I suggest that this happens because, as the individual words occur, they are stored temporarily via STM traces in a working memory. As the words are stored, they activate LTM traces which attempt to categorize the stored sound stream into familiar language units like words at a higher processing level. These list categories, in turn, activate learned top-down expectations that are *matched* against the contents of working memory to verify that the information expected from previous learning experiences is really there. This concept of bottom-up activation of learned categories by a working memory, followed by read-out of learned top-down expectations, is illustrated in Figure 3A.

What is the nature of this matching, or verification, process? Its properties have been clarified by experiments of Arthur Samuel (Samuel, 1981a, 1981b) and others in which the spectral content of the noise was varied. If the noise includes all the formants of the expected sound, then that is what the subject hears, and other spectral components of the noise are suppressed. If some formants of the expected sound are missing from the noise, then only a partial reconstruction is heard. If silence replaces the noise, then only silence is heard. The matching process thus cannot "create something out of nothing". It can, however, selectively amplify the expected features in the bottom-up signal and suppress the rest, as in Figure 3B.

The process whereby the top-down expectation selectively amplifies some features while suppressing others helps to "focus attention" upon information that matches our momentary expectations. This focusing process helps to filter out the flood of sensory signals that would

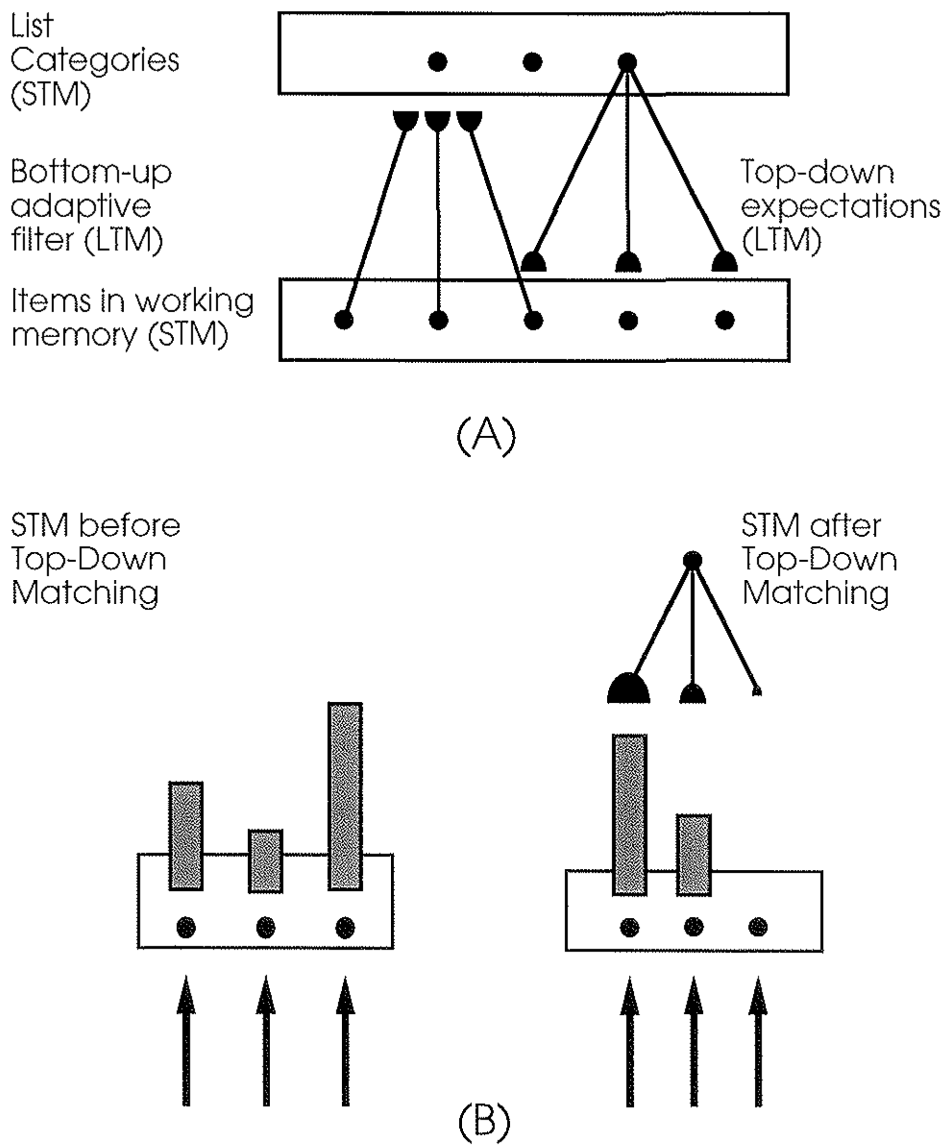


Figure 3. (A) Auditory items activate STM traces in a working memory, which send bottom-up signals towards a level at which list categories, or chunks, are activated in STM. These bottom-up signals are multiplied by learned LTM traces which influence the selection of the list categories that are stored in STM. The list categories, in turn, activate LTM-modulated top-down expectation signals that are matched against the active STM pattern in working memory. (B) This matching process confirms and amplifies STM activations that are supported by contiguous LTM traces, and suppresses those that are not.

otherwise overwhelm us, and to prevent them from destabilizing our previously learned memories. Learned top-down expectations hereby help to solve the stability-plasticity dilemma by focusing attention and preventing spurious signals from accidentally eroding our previously learned memories. In fact, Gail Carpenter and I proved mathematically in 1987 that such an ART matching rule assures stable learning of an ART model in response to rapidly changing environments wherein learning becomes unstable if the matching rule is removed (Carpenter and Grossberg, 1987a).

What does all this have to do with our conscious percepts of speech? This can be seen by asking: If top-down expectations can select consistent bottom-up signals, then what keeps the selected bottom-up signals from reactivating their top-down expectations in a continuing cycle of bottom-up and top-down feedback? Nothing does! In fact, this reciprocal feedback process takes awhile to equilibrate, and when it does, the bottom-up and top-down signals lock the STM activity patterns of the interacting levels into a resonant state that lasts much longer and is more energetic than any individual activation. ART hereby, suggests how only resonant states of the brain can achieve consciousness, and that the time needed for a bottom-up/top-down resonance to develop helps to explain why a conscious percept of an event takes so long to occur after its bottom-up input is delivered.

The example of phonemic restoration also clarifies another key point about the conscious perception of speech. If noise precedes "eel is on the shoe", we hear and understand the meaning of the sentence "heel is on the shoe". If, however, noise is replaced by silence, we hear and understand the meaning of the sentence "eel is on the shoe" which has a quite different, and rather disgusting, meaning. This example shows that the process of resonance binds together information about both meaning and phonetics. Meaning is not some higher-order process that is processed independently from the process of conscious phonetic hearing. Meaning and phonetics are bound together via resonant feedback into a global emergent state in which the phonetics that we hear are linked to the meaning that we understand.

ART Matching and Resonance: The Link Between Attention, Intention, and Consciousness

Adaptive resonance theory claims that, in order to solve the stability-plasticity dilemma, only resonant states can drive new learning. That is why the theory is called *adaptive* resonance theory. I will explain how this works more completely below. Before doing so, let me emphasize some implications of the previous discussion that are worth reflecting about. The first implication provides a novel answer to why, as philosophers have asked for many years, humans are "intentional" beings who are always anticipating or planning their next behaviors and their expected consequences. ART suggests that "stability implies intentionality". That is, stable learning requires that we have expectations about the world that are continually matched against world data. Otherwise expressed, without stable learning, we could learn very little about the world. Having an active top-down matching mechanism greatly amplifies the amount of information that we can stably learn about the world. Thus the mechanisms which enable us to know a changing external world, through the use of learned expectations, set the stage for achieving internal self-awareness.

It should be noted here that the word "intentionality" is being used, at once, in two different senses. One sense concerns the role of expectations in the anticipation of events that may or may not occur. The second sense concerns the ability of expectations to read-out planned sequences of behaviors aimed at achieving definite behavioral goals. The former sense will be emphasized first; the latter towards the end of the article. My main point in lumping them together is that ART provides a unified mechanistic perspective with which to understand both uses of the word.

The second implication is that "intention implies attention and consciousness". That is, expectations start to focus attention on data worthy of learning, and these attentional foci are confirmed when the system as a whole incorporates them into resonant states that include (I claim) conscious states of mind.

Implicit in the concept of intentionality is the idea that we can get *ready* to experience an expected event so that, when it finally occurs, we can react to it more quickly and vigorously, and until it occurs, we are able to ignore other, less desired, events. This property is called *priming*. It implies that, when a top-down expectation is read-out in the absence of a bottom-up input, it can subliminally sensitize the cells that would ordinarily respond to the bottom-up input, but not actually fire them, while it suppresses cells whose activity is not expected. Correspondingly, the ART matching rule computationally realizes the following properties at any processing level where bottom-up and top-down signals are matched:

- **Bottom-Up Automatic Activation:** A cell, or node, can become active enough to generate output signals if it receives a large enough bottom-up input, other things being equal.
- **Top-Down Priming:** A cell can become sensitized, or subliminally active, and thus cannot generate output signals, if it receives only a large top-down expectation input. Such a top-down priming signal prepares a cell to react more quickly and vigorously to subsequent bottom-up input that matches the top-down prime.
- **Match:** A cell can become active if it receives large convergent bottom-up and top-down inputs. Such a matching process can generate enhanced activation as resonance takes hold.
- **Mismatch:** A cell is suppressed even if it receives a large bottom-up input if it also receives only a small, or zero, top-down expectation input.

I claim that this ART matching rule and the resonance rule that it implies operate in all the examples that I have previously sketched, and do so to solve the stability-plasticity dilemma. All the examples are proposed to illustrate how we can continue to learn rapidly and stably about new experiences throughout life by matching bottom-up signal patterns from more peripheral to more central brain processing stages against top-down signal patterns from more central to more peripheral processing stages. These top-down signals represent the brain's learned expectations of what the bottom-up signal patterns should be based upon past experience. The matching process is designed to reinforce and amplify those combinations of features in the bottom-up pattern that are consistent with the top-down expectations, and to suppress those features that are inconsistent. This top-down matching step initiates the process whereby the brain selectively pays attention to experiences that it expects, binds them into coherent internal representations through resonant states, and incorporates them through learning into its knowledge about the world.

Given that such a resonant matching process occurs in the brain, how does the brain react when there is a mismatch situation? The ART matching rule suggests that a big enough mismatch between a bottom-up input and a top-down expectation can rapidly attenuate activity at the matching level. This collapse of bottom-up activation can initiate a rapid *reset* of activity at both the matching level itself and at the subsequent levels that it feeds, thereby initiating a memory search for a more appropriate recognition category or creating a new one.

Resonant Dynamics During Speech Categorization

Many examples of such a reset event occur during variable-rate speech perception. As one example, consider how people hear combinations of vowels (V) and consonants (C) in VC–CV sequences. Bruno Repp at Haskins Laboratories has studied perception of the sequences [ib]–[ga] and [ib]–[ba] when the silence interval between the initial VC syllable and the terminal CV syllable is varied (Repp, 1980). If the silence interval is short enough, then [ib]–[ga] sounds like [iga] and [ib]–[ba] sounds like [iba]. Repp ran a number of conditions, leading to the several data curves displayed in Figure 4. The main point for present purposes is that the transition from a percept of [iba] to one of [ib]–[ba] occurs after 100–150 milliseconds more silence than the transition from [iga] to [ib]–[ga]. One hundred milliseconds is a very long time relative to the time scale at which individual neurons can be activated. Why is this shift so large?

My colleagues Ian Boardman, Michael Cohen, and I have quantitatively simulated these data using a model, called the ARTPHONE model, of how a resonant wave develops due to bottom-up and top-down signal exchanges between a working memory that represents the individual speech items and a list categorization network that groups them together into learned language units, or chunks (Grossberg, Boardman, and Cohen, 1997). We have shown how a mismatch between [g] and [b] rapidly resets the working memory if the silence between them is short enough, thereby preventing the [b] sound from reaching resonance and consciousness, as in Figure 5. We have also shown how the development of a previous resonance involving [b] can resonantly fuse with a subsequent [b] sound to greatly extend the perceived duration of [iba] across a silence interval between [ib] and [ba]. Figure 6A illustrates this property by suggesting how the second presentation of [b] can quickly reactivate the resonance in response to the first presentation of [b] before the resonance stops. This phenomenon uses the property that it takes longer for the first presentation of [b] to reach resonance than it does for the second presentation of [b] to influence the maintenance of this resonance.

If, however, [ib] can fuse across time with [ba], then how do we ever hear distinct [ib]–[ba] sounds when the silence gets long enough? Much evidence suggests that after a resonance fully develops, it spontaneously collapses after awhile due to a habituative process that goes on in the pathways that maintain the resonance via bottom-up and top-down signals. Thus, if the silence is long enough for resonant collapse of [ib] to occur, then a distinguishable [ba] resonance can subsequently develop and be heard, as in Figure 6B.

Such a habituative process has also been used to explain many other data about perception, learning, and recognition, notably data about the reset of visual, cognitive, or motor representations in response to rapidly changing events. Relevant visual data include properties of light adaptation, visual persistence, aftereffects, residual traces, and apparent motion (Carpenter and Grossberg, 1981; Francis and Grossberg, 1996a, 1996b; Francis, Grossberg, and Mingolla, 1994). Abbott *et al.* (1997) have recently reported data from visual cortex that they modeled using the same habituative law that was used in all of these applications. At bottom, such a habituative law is predicted to be found so ubiquitously across brain systems because it helps to rapidly reset and rebalance neural circuits in response to rapidly changing input conditions, notably as part of an opponent process (Grossberg, 1980).

The Repp (1980) data illustrate the important fact that the duration of a consciously perceived interval of silence is sensitive to the phonetic context into which the silence is placed. These data show that the phonetic context can generate a conscious percept of continuous sound across 150 milliseconds of silence – that can be heard as silence in a different phonetic context. Our explanation of these data in terms of the maintenance of resonance in one case, but its rapid reset in another, is consistent with a simple, but revolutionary, definition of silence: Silence is a temporal discontinuity in the rate with which the auditory resonance evolves in time. Various other models of speech perception, having no concept like resonance on which to build, cannot begin to explain data of this type. Several such models are reviewed in Grossberg, Boardman, and Cohen (1997).

Figure 6

If, however, [ib] can fuse across time with [ba], then how do we ever hear distinct [ib]–[ba] sounds when the silence gets long enough? Much evidence suggests that after a resonance fully develops, it spontaneously collapses after awhile due to a habituative process that goes on in the pathways that maintain the resonance via bottom-up and top-down signals. Thus, if the silence is long enough for resonant collapse of [ib] to occur, then a distinguishable [ba] resonance can subsequently develop and be heard, as in Figure 6B.

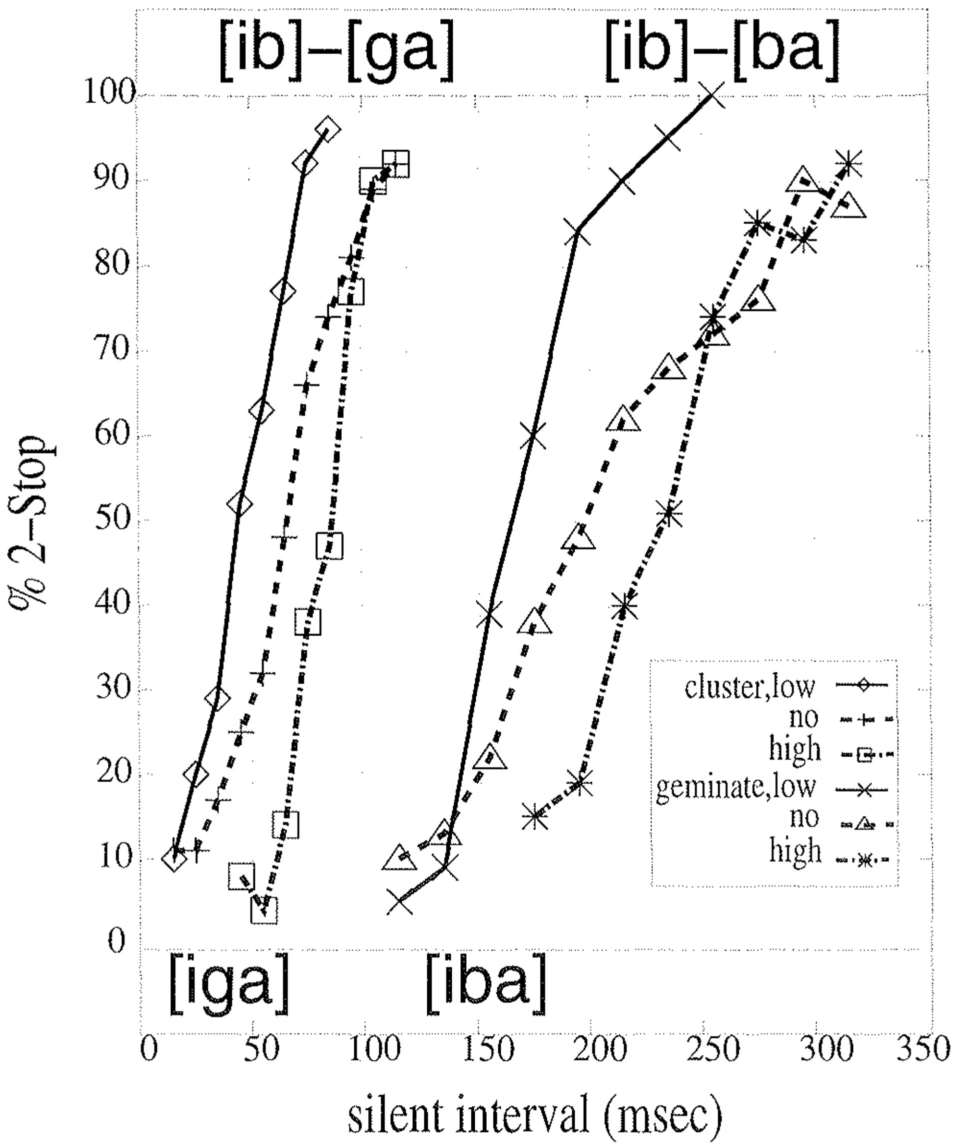


Figure 4. The left-hand curves represent the probability, under several experimental conditions, that the subject will hear [ib]–[ga] rather than [iga]. The right-hand curves do the same for [ib]–[ba] rather than the fused percept [iba]. Note that the perception of [iba] can occur at a silence interval between [ib] and [ba] that is up to 150 milliseconds longer than the one that leads to the percept [iga] instead of [ib]–[ga]. (Data are reprinted with permission from B.H. Repp (1980), Haskins Laboratories Status Report on Speech Research, **SR-61**, 151–165.)

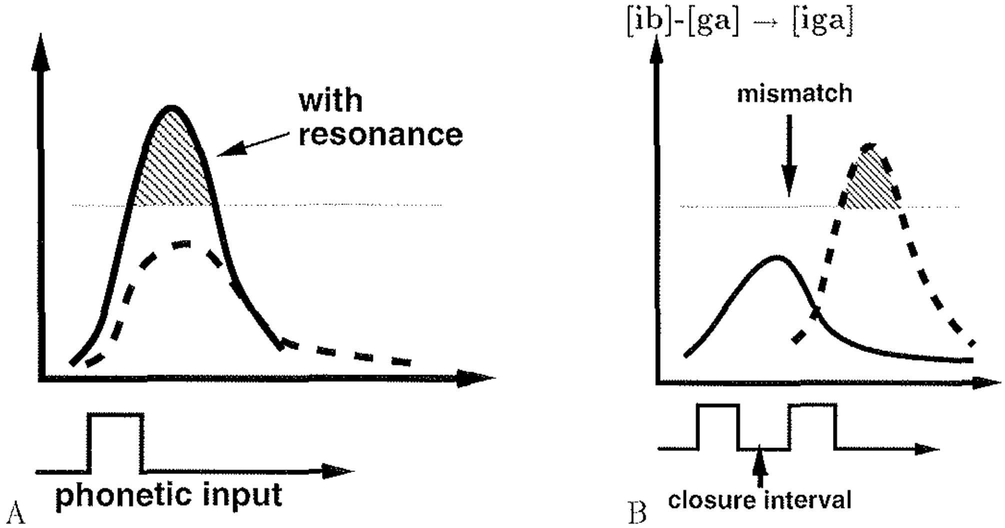


Figure 5. (A) Response to a single stop, such as [b] or [g], with and without resonance. Suprathreshold activation is shaded. (B) Reset due to phonologic mismatch between [ib] and [ga].

Such a habituative process has also been used to explain many other data about perception, learning, and recognition, notably data about the reset of visual, cognitive, or motor representations in response to rapidly changing events. Relevant visual data include properties of light adaptation, visual persistence, aftereffects, residual traces, and apparent motion (Carpenter and Grossberg, 1981; Francis and Grossberg, 1996a, 1996b; Francis, Grossberg, and Mingolla, 1994). Abbott *et al.* (1997) have recently reported data from visual cortex that they modeled using the same habituative law that was used in all of these applications. At bottom, such a habituative law is predicted to be found so ubiquitously across brain systems because it helps to rapidly reset and rebalance neural circuits in response to rapidly changing input conditions, notably as part of an opponent process (Grossberg, 1980).

The Repp (1980) data illustrate the important fact that the duration of a consciously perceived interval of silence is sensitive to the phonetic context into which the silence is placed. These data show that the phonetic context can generate a conscious percept of continuous sound across 150 milliseconds of silence – that can be heard as silence in a different phonetic context. Our explanation of these data in terms of the maintenance of resonance in one case, but its rapid reset in another, is consistent with a simple, but revolutionary, definition of silence: Silence is a temporal discontinuity in the rate with which the auditory resonance evolves in time. Various other models of speech perception, having no concept like resonance on which to build, cannot begin to explain data of this type. Several such models are reviewed in Grossberg, Boardman, and Cohen (1997).

Resonant Dynamics During Auditory Streaming

A similar type of resonant processing helps to explain cocktail party separation of distinct voices into auditory streams, as in the auditory continuity illusion of Figure 2. This process goes on, however, at earlier stages of auditory processing than speech categorization. My colleagues Krishna Govindarajan, Lonce Wyse, Michael Cohen, and I have developed a model, called the ARTSTREAM model, of how distinguishable auditory streams are resonantly formed and separated (Grossberg, 1998b; Govindarajan, Grossberg, Wyse, and Cohen, 1995). Here the two main processing levels (Figure 7) are a spectral stream level at which the frequencies of the sound spectrum are represented across a spatial map, and a pitch stream level at which pitch nodes respond to the harmonics at the spectral stream

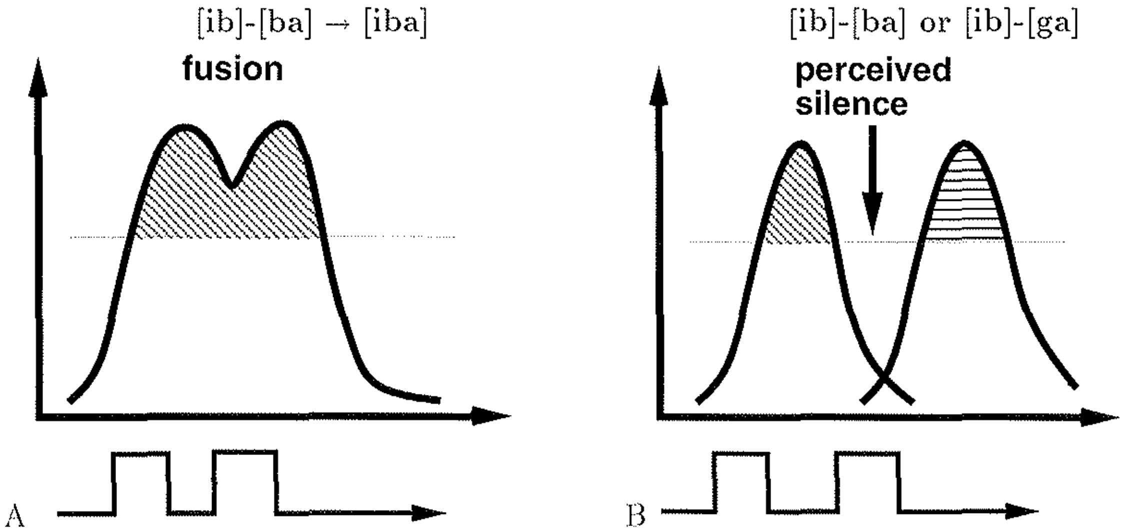


Figure 6. (A) Fusion in response to proximal similar phones. (B) Perceptual silence allows a 2-stop percept.

level that comprise a given pitch. After the auditory signal is preprocessed, its spectral, or frequency, components are redundantly represented in multiple spectral streams; that is, the sound's preprocessed frequency components are represented in multiple spatial maps, each one of which can subserve the percept of a particular auditory stream.

Each of these spectral streams is filtered by bottom-up signals that activate its own pitch stream representation at the pitch stream level; that is, there are multiple pitch streams, one corresponding to every spectral stream. This multiple representation of a sound's spectral components and pitch interact to break up the entire sound stream that is entering the system into distinct acoustic sources or voices. This happens as follows. A given sound spectrum is multiply represented at all the spectral streams and then redundantly activates all of the pitch nodes that are consistent with these sounds. These pitch representations compete to select a winner, which inhibits the representations of the same pitch across streams, while also sending top-down matching signals back to the spectral stream level. By the ART matching rule, the frequency components that are consistent with the winning pitch node are amplified, and all others are suppressed, thereby leading to a spectral-pitch resonance within the stream of the winning pitch node. In this way, the pitch layer coherently binds together the harmonically related frequency components that correspond to a prescribed auditory source. All the frequency components that are suppressed by ART matching in this stream are freed to activate and resonate with a different pitch in a different stream. The net result is multiple resonances, each selectively grouping together into pitches those frequencies that correspond to distinct auditory sources.

Using the ARTSTREAM model, we have simulated many of basic streaming percepts, including the auditory continuity illusion of Figure 2. It occurs, I contend, because the spectral-stream resonance takes a time to develop that is commensurate to the duration of the subsequent noise. Once the tone resonance develops, the second tone can quickly act to support and maintain it throughout the duration of the noise, much as [ba] fuses with [ib] during perception of [iba]. Of course, for this to make sense, one needs to accept the fact that the tone resonance does not start to get consciously heard until just about when the second tone occurs.

A Circuit for ART Matching

Figure 7 incorporates one of the possible ways that Gail Carpenter and I proposed in

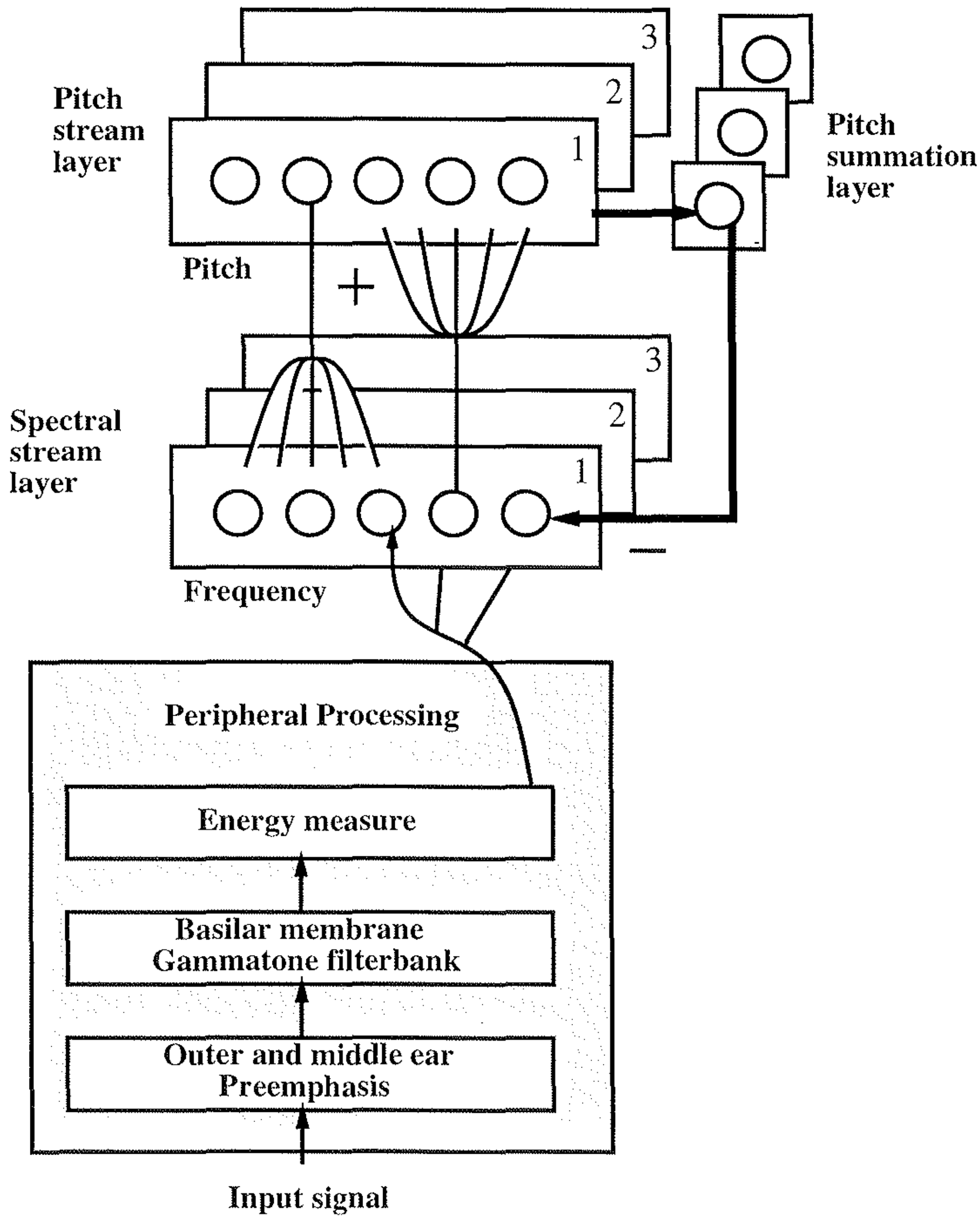


Figure 7. Block diagram of the ARTSTREAM auditory streaming model. Note the non-specific top-down inhibitory signals from the pitch level to the spectral level that realize ART matching within the network.

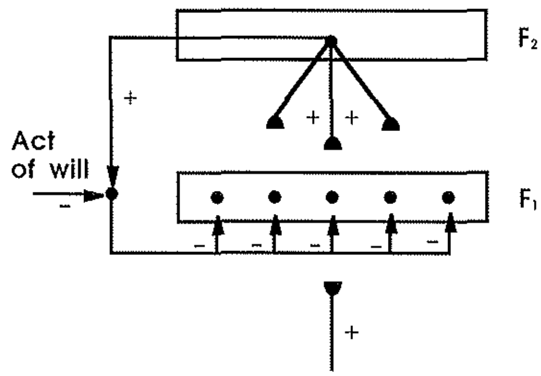


Figure 8. One way to realize the ART matching rule using top-down activation of nonspecific inhibitory interneurons, as in Figure 8. Several mathematically possible alternative ways are suggested in the Appendix of G.A. Carpenter and S. Grossberg (1987a).

the mid-1980's for how the ART matching rule can be realized (Carpenter and Grossberg, 1987a). This matching circuit is redrawn in Figure 8 for clarity. It is perhaps the simplest such circuit, and I have found it in subsequent studies to be the one that is implicated by data time and time again.

In this circuit, bottom-up signals to the spectral stream level can excite their target nodes if top-down signals are not active. Top-down signals try to excite those spectral, or frequency component, nodes that are consistent with the pitch node that activates them. By themselves, top-down signals fail to activate spectral nodes because the pitch node also activates a pitch summation layer that nonspecifically inhibits all spectral nodes in its stream. The nonspecific top-down inhibition hereby prevents the specific top-down excitation from supraliminally activating any spectral nodes. On the other hand, when excitatory bottom-up and top-down signals occur together, then those spectral nodes that receive both types of signals can be fully activated. All other nodes in that stream are inhibited, including spectral nodes that were previously activated by bottom-up signals but received no subsequent top-down pitch support. Attention hereby selectively activates consistent nodes while nonselectively inhibiting all other nodes in a stream.

How Early Does Attention Act in the Brain?

It has classically been thought that attention first acts at higher levels of cortical organization. In vision, for example, it was thought that attention occurred no earlier than the extrastriate visual cortex (McAdams and Maunsell, 1997; Motter, 1994a, 1994b; Reynolds *et al.*, 1995). However, recent experiments have suggested that attentional modulation can occur in primary visual cortex (V1), and even at the earlier Lateral Geniculate Nucleus via top-down cortico-geniculate pathways (Hupé *et al.*, 1997; Ito, Westheimer, and Gilbert, 1997; Johnson and Burkhalter, 1997; Lamme, Zipser, and Spekreijse, 1997; Press and van Essen, 1997). Is there a contradiction here? The answer depends upon how you define attention. If attention refers only to processes that can be controlled voluntarily, then corticogeniculate feedback, being automatic, may not qualify. On the other hand, such top-down feedback does appear to have the selective properties of an "automatic" attention process, even at the earliest stages of sensory processing.

Is there direct experimental evidence for the prediction that corticogeniculate feedback supports ART matching and resonance? In a remarkable 1994 *Nature* article, Sillito and his colleagues (Sillito *et al.*, 1994) published neurophysiological data that strikingly support this prediction. They wrote in particular that "cortically induced correlation of relay cell activity produces coherent firing in those groups of relay cells with receptive field alignments appropriate to signal the particular orientation of the moving contour to the cortex...this increases the gain of the input for feature-linked events detected by the cortex...the cortico-thalamic input is only strong enough to exert an effect on those dLGN cells that are additionally polarized by their retinal input...the feedback circuit searches for correlations that support the 'hypothesis' represented by a particular pattern of cortical activity". In short, Sillito verified all the properties of the ART matching rule.

Attention at All Stages of Sensory and Cognitive Neocortex?

It has, in fact, been suggested how similar automatic attentional processes are integrated within the laminar circuits of visual cortex, notably the circuits of cortical areas V1 and V2 that are used to generate perceptual groupings, such as the illusory contours in Figure 1 (Grossberg, 1998a). In this proposal, the ART Matching Rule is realized as follows. Top-down attentional feedback from cortical area V2 to V1 is predicted to be mediated by signals from layer 6 of cortical area V2. These top-down signals attentionally prime layer 4 of cortical area V1 via an on-center off-surround network within V1 from layer 6 to layer 4. In this conception, layer 6 of V2 activates layer 6 of V1, possibly via a multisynaptic pathway via layers 1 and 5. Layer 6 then activates layer 4 of V1 via an on-center off-surround network from layer 6-to-4. This analysis predicts that the layer 6-to-4 on-center circuit can *prime*, or *modulate*, layer 4 cells, but cannot fully activate them, because the top-down attentional prime, acting by itself, is subliminal. Such a modulatory effect is achieved by appropriately balancing the strength of the on-center and off-surround signals within the layer 6-to-4 network.

Related modeling work has shown how such balanced on-center off-surround signals can lead to self-stabilizing development of the horizontal connections within layers 2/3 of V1 and V2 that subserve perceptual grouping (Grossberg and Williamson, 1997, 1998). It has also been shown how the top-down on-center off-surround circuit from area V1 to LGN can self-stabilize the development of disparity-sensitive complex cells in area V1 (Grunewald and Grossberg, 1998). Other modeling work has suggested how a similar top-down on-center off-surround automatical attentional circuit from cortical area MST to MT can be used to generate coherent representations of the direction and speed with which objects move (Chey, Grossberg, and Mingolla, 1997). Taken together, these studies show how the ART Matching Rule may be realized in known cortical circuits, and how it can self-stabilize development of these circuits as a precursor to its role in self-stabilizing learning throughout life. Grossberg (1998a) has predicted that the same ART matching circuit exists within the laminar organization that is found universally in all sensory and cognitive neocortex, including the various examples of auditory processing that are reviewed above. This prediction does not, of course, deny that these various circuits may be specialized in various ways to process the different types of information with which they are confronted.

Given that the cortical organization of top-down on-center off-surround attentional priming circuits seem to be ubiquitous in visual cortex, and by extension in other types of cortex, it is important to ask: What more does the brain need to add in order to generate a more flexible, task-dependent type of attention switching? This question leads us to consider visual object recognition, and how it breaks down during medial temporal amnesia. Various other models of object recognition, and their conceptual and explanatory weaknesses relative to ART, are reviewed in Grossberg and Merrill (1996).

Self-Organizing Feature Maps for Learned Object Recognition

Let us begin with a two-level network that illustrates some of the main ideas in the simplest possible way. Level $\mathcal{F}_1$ in Figure 9 contains a network of nodes, or cell populations, each of which is activated by a particular combination of sensory features via inputs. Level $\mathcal{F}_2$ contains a network of nodes that represent recognition codes, or categories, which are selectively activated by the activation patterns across $\mathcal{F}_1$. Each $\mathcal{F}_1$ node sends output signals to a subset of $\mathcal{F}_2$ nodes. Each $\mathcal{F}_2$ node thus receives inputs from many $\mathcal{F}_1$ nodes. The thick bottom-up pathway from $\mathcal{F}_1$ to $\mathcal{F}_2$ in Figure 9 represents in a concise way an array of diverging and converging pathways. Let learning take place at the synapses denoted by semicircular endings in the $\mathcal{F}_1 \rightarrow \mathcal{F}_2$ pathways. Pathways that end in arrowheads do not undergo learning. This bottom-up learning enables $\mathcal{F}_2$ category nodes to become selectively tuned to particular combinations of activation patterns across $\mathcal{F}_1$ feature detectors by changing their LTM traces.

Why is not bottom-up learning sufficient in a system that can autonomously solve the stability-plasticity dilemma? Why are learned top-down expectations also needed? To understand this, we consider a type of model that is often called a self-organizing feature map, competitive learning, or learned vector quantization. This type of model shows how to combine associative learning and lateral inhibition for purposes of learned categorization.

In such a model, as shown in Figure 10A, an input pattern registers itself as a pattern of activity, or STM, across the feature detectors of level $\mathcal{F}_1$. Each $\mathcal{F}_1$ output signal is multiplied or gated, by the adaptive weight, or LTM trace, in its respective pathway. All these LTM-gated inputs are added up at their target $\mathcal{F}_2$ nodes. The LTM traces hereby *filter* the STM signal pattern and generate larger inputs to those $\mathcal{F}_2$ nodes whose LTM patterns are most similar to the STM pattern. Lateral inhibitory, or competitive, interactions within $\mathcal{F}_2$ contrast-enhance this input pattern. Whereas many $\mathcal{F}_2$ nodes may receive inputs from $\mathcal{F}_1$, lateral inhibition allows a much smaller set of $\mathcal{F}_2$ nodes to store their activation in STM. These are the $\mathcal{F}_2$ nodes whose LTM patterns are most similar to the STM pattern. These inhibitory interactions also tend to conserve the total activity that is stored in STM (Grossberg, 1982), thereby realizing an interference-based capacity limitation in STM.

Only the $\mathcal{F}_2$ nodes that win the competition and store their activity in STM can influence the learning process. STM activity opens a learning gate at the LTM traces that abut the winning nodes. These LTM traces can then approach, or track, the input signals in their pathways, a process called steepest descent. This learning law is thus often called gated steepest descent, or instar learning. This type of learning tunes the winning LTM patterns to become even more similar to the STM pattern, and to thereby enable the STM pattern to more effectively activate the corresponding $\mathcal{F}_2$ nodes. I introduced this learning law into neural network models in the 1960's (e.g. Grossberg, 1969), and into ART models in the 1970's (Grossberg, 1976a, 1976b, 1978, 1980). Such an LTM trace can either increase (Hebbian) or decrease (anti-Hebbian) to track the signals in its pathway (Table 1). It has been used to model neurophysiological data about learning in the hippocampus (also called long term potentiation and long term depression) and about adaptive tuning of cortical feature detectors during early visual development (Artola and Singer, 1993; Levy, 1985, Levy and Desmond, 1985, Rauschecker and Singer, 1979; Singer, 1983), thereby lending support to ART predictions that these systems would employ this type of learning.

Self-organizing feature map models were introduced and computationally characterized by Christoph von der Malsburg and myself during the 1970's (Grossberg, 1972, 1976a, 1978; von der Malsburg, 1973; Willshaw and Malsburg, 1976). These models were subsequently applied and further developed by many authors, notably Teuvo Kohonen (Kohonen, 1984). They exhibit many useful properties, especially if not too many input patterns, or clusters of input patterns, perturb level $\mathcal{F}_1$ relative to the number of categorizing nodes in level $\mathcal{F}_2$. I proved that, under these sparse environmental conditions, category learning is stable in the sense that its LTM traces converge to fixed values as learning trials proceed. In addition, the LTM traces track the statistics of the environment, are self-normalizing, and oscillate a minimum number of times (Grossberg, 1976a). Also, the category selection rule, like a Bayesian classifier, tends to minimize error. I also proved, however, that under *arbitrary*

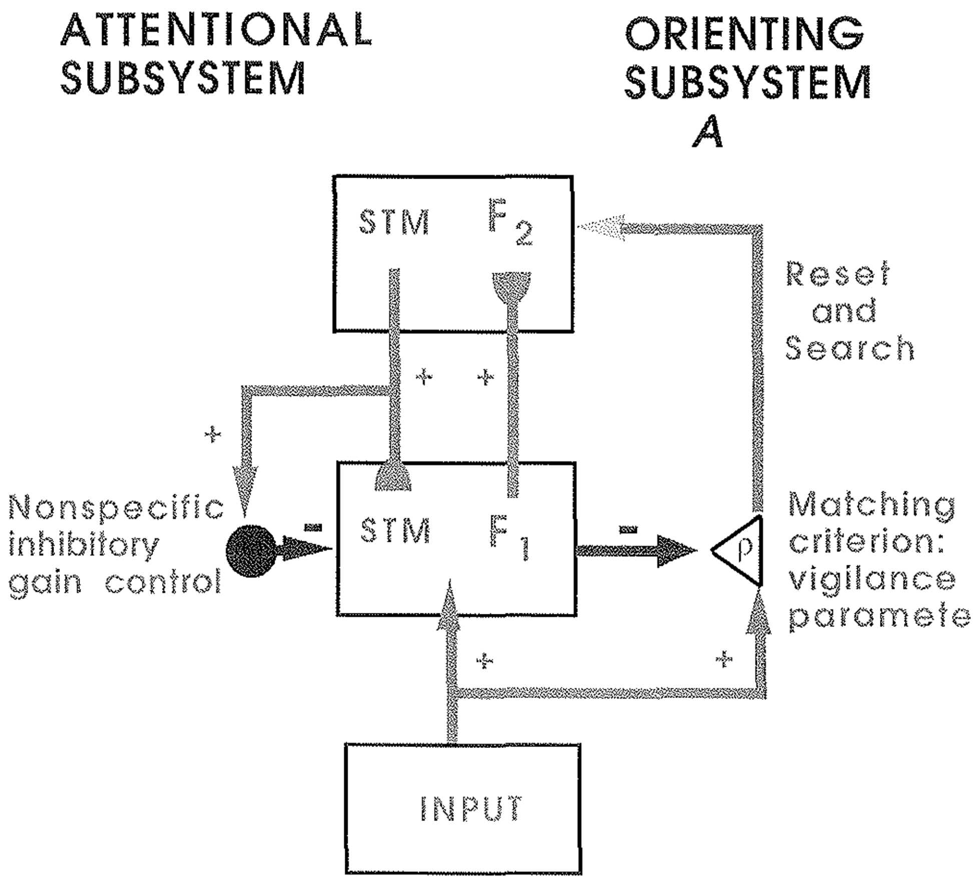


Figure 9. An example of a model ART circuit in which attentional and orienting circuits interact. Level $\mathcal{F}_1$ encodes a distributed representation of an event by a short term memory (STM) activation pattern across a network of feature detectors. Level $\mathcal{F}_2$ encodes the event using a compressed STM representation of the $\mathcal{F}_1$ pattern. Learning of these recognition codes occurs at the long term memory (LTM) traces within the bottom-up and top-down pathways between levels $\mathcal{F}_1$ and $\mathcal{F}_2$. The top-down pathways read-out learned expectations whose prototypes are matched against bottom-up input patterns at $\mathcal{F}_1$. The size of mismatches in response to novel events are evaluated relative to the vigilance parameter ρ of the orienting subsystem $\mathcal{A}$. A large enough mismatch resets the recognition code that is active in STM at $\mathcal{F}_2$ and initiates a memory search for a more appropriate recognition code. Output from subsystem $\mathcal{A}$ can also trigger an orienting response. (A) Block diagram of circuit. (B) Individual pathways of circuit, including the input level $\mathcal{F}_0$ that generates inputs to level $\mathcal{F}_1$. The gain control input g_1 to level $\mathcal{F}_1$ helps to instantiate the matching rule (see text). Gain control g_2 to level $\mathcal{F}_2$ is needed to instate a category in STM.

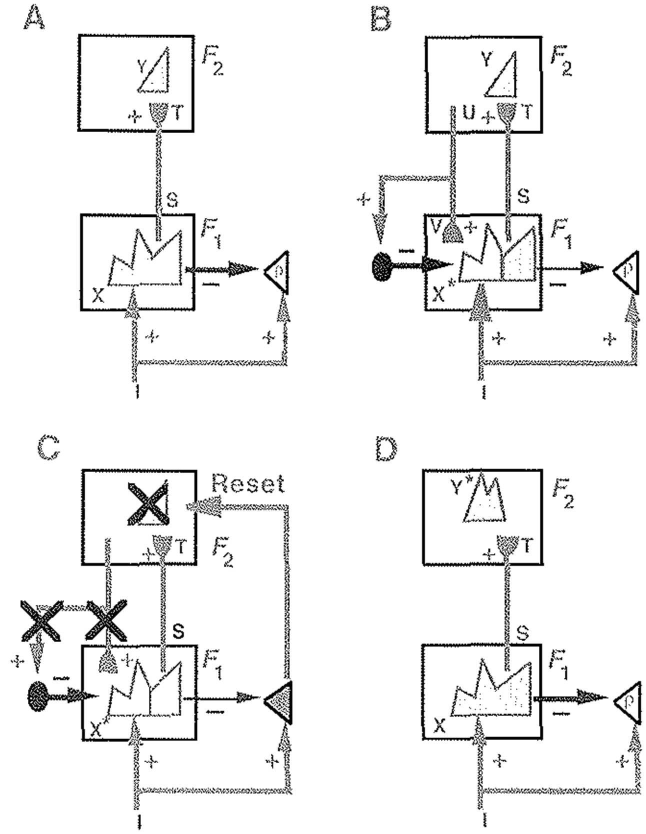
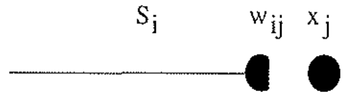


Figure 10. ART search for a recognition code: (A) The input pattern **I** is instated across the feature detectors at level $\mathcal{F}_1$ as a short term memory (STM) activity pattern **X**. Input **I** also nonspecifically activates the orienting subsystem $\mathcal{A}$; see Figure 1. STM pattern **X** is represented by the hatched pattern across $\mathcal{F}_1$. Pattern **X** both inhibits $\mathcal{A}$ and generates the output pattern **S**. Pattern **S** is multiplied by long term memory (LTM) traces and added at $\mathcal{F}_2$ nodes to form the input pattern **T**, which activates the STM pattern **Y** across the recognition categories coded at level $\mathcal{F}_2$. (B) Pattern **Y** generates the top-down output pattern **U** which is multiplied by top-down LTM traces and added at $\mathcal{F}_1$ nodes to form the prototype pattern **V** that encodes the learned expectation of the active $\mathcal{F}_2$ nodes. If **V** mismatches **I** at $\mathcal{F}_1$, then a new STM activity pattern $\mathbf{X}^*$ is generated at $\mathcal{F}_1$. $\mathbf{X}^*$ is represented by the hatched pattern. It includes the features of **I** that are confirmed by **V**. Inactivated nodes corresponding to unconfirmed features of **X** are unhatched. The reduction in total STM activity which occurs when **X** is transformed into $\mathbf{X}^*$ causes a decrease in the total inhibition from $\mathcal{F}_1$ to $\mathcal{A}$. (C) If inhibition decreases sufficiently, $\mathcal{A}$ releases a nonspecific arousal wave to $\mathcal{F}_2$, which resets the STM pattern **Y** at $\mathcal{F}_2$. (D) After **Y** is inhibited, its top-down prototype signal is eliminated, and **X** can be reinstated at $\mathcal{F}_1$. Enduring traces of the prior reset lead **X** to activate a different STM pattern $\mathbf{Y}^*$ at $\mathcal{F}_2$. If the top-down prototype due to $\mathbf{Y}^*$ also mismatches **I** at $\mathcal{F}_1$, then the search for an appropriate $\mathcal{F}_2$ code continues until a more appropriate $\mathcal{F}_2$ representation is selected. Then an attentive resonance develops and learning of the attended data is initiated. [Reprinted with permission from Carpenter and Grossberg (1993).]



	Case 1	Case 2	Case 3	Case 4
State of S_i	+	−	+	−
State of x_j	+	+	−	−
State of w_{ij}	↑	↓	↔	↔

+ = active ↑ = increase

− = inactive ↓ = decrease

↔ = no change

Table 1. The instar learning, or gated steepest descent learning rule, embodies both Hebbian (LTP) and anti-Hebbian (LTD) properties within a single process. (Reprinted with permission from Grossberg and Merrill, 1996).

environmental conditions, learning becomes unstable (Grossberg, 1976b). Such a model could forget your parents' faces when it learns a new face. Although a gradual switching off of plasticity can partially overcome this problem, such a mechanism cannot work in a learning system whose plasticity is maintained throughout adulthood.

This memory instability is due to basic properties of associative learning and lateral inhibition, which are two processes that occur ubiquitously in the brain. An analysis of this instability, together with data about human and animal categorization, conditioning, and attention, led me to introduce ART models to stabilize the memory of self-organizing feature maps in response to an arbitrary stream of input patterns.

How Does ART Stabilize Learning of a Self-Organizing Feature Map?

How does an ART model prevent such instabilities from developing? As noted above, in an ART model, learning does not occur when some winning $\mathcal{F}_2$ activities are stored in STM. Instead, activation of $\mathcal{F}_2$ nodes may be interpreted as "making a hypothesis" about an input at $\mathcal{F}_1$. When $\mathcal{F}_2$ is activated, it quickly generates an output pattern that is transmitted along the top-down adaptive pathways from $\mathcal{F}_2$ to $\mathcal{F}_1$. These top-down signals are multiplied in their respective pathways by LTM traces at the semicircular synaptic knobs of Figure 10B. The LTM-gated signals from all the active $\mathcal{F}_2$ nodes are added to generate the total top-down feedback pattern from $\mathcal{F}_2$ to $\mathcal{F}_1$. It is this pattern that plays the role of a learned expectation. Activation of this expectation may be interpreted as "testing the hypothesis", or "reading out the prototype", of the active $\mathcal{F}_2$ category. As shown in Figure 10B, ART networks are designed to match the "expected prototype" of the category against the bottom-up input pattern, or exemplar, to $\mathcal{F}_1$. Nodes that are activated by this exemplar are suppressed if they do not correspond to large LTM traces in the top-down prototype pattern. The resultant $\mathcal{F}_1$ pattern encodes the cluster of input features that the network deems relevant to the hypothesis based upon its past experience. This resultant activity pattern, called $\mathbf{X}^*$ in Figure 10B, encodes the pattern of features to which the network "pays attention".

If the expectation is close enough to the input exemplar, then a state of resonance develops as the attentional focus takes hold. The pattern $\mathbf{X}^*$ of attended features reactivates the $\mathcal{F}_2$ category $\mathbf{Y}$ which, in turn, reactivates $\mathbf{X}^*$. The network locks into a resonant state through a positive feedback loop that dynamically links, or binds, $\mathbf{X}^*$ with $\mathbf{Y}$. The resonance binds spatially distributed features into either a stable equilibrium or a synchronous oscillation, much like the synchronous feature binding in visual cortex that has recently attracted so much interest after the experiments of Reinhard Eckhorn, Wolf Singer, and their colleagues (Eckhorn *et al.*, 1988; Gray and Singer, 1989); also see Grossberg and Grunewald (1997).

In ART, the resonant state, rather than bottom-up activation, is predicted to drive the learning process. The resonant state persists long enough, at a high enough activity level, to activate the slower learning processes in the LTM traces. This helps to explain how the LTM traces can regulate the brain's fast information processing without necessarily learning about the signals that they process. Through resonance as a mediating event, the combination of top-down matching and attentional focusing helps to stabilize ART learning and memory in response to an arbitrary input environment. The stabilizing properties of top-down matching may be one reason for the ubiquitous occurrence of reciprocal bottom-up and top-down cortico-cortical and cortico-thalamic interactions in the brain.

How is the Generality of Knowledge Controlled?

A key problem about consciousness concerns what combinations of features or other information are bound together into object or event representations. ART provides a new answer to this question that overcomes problems faced by earlier models. In particular, ART systems learn prototypes, rather than exemplars, because the attended feature vector $\mathbf{X}^*$, rather than the input exemplar itself, is learned. Both the bottom-up LTM traces that tune the category nodes and the top-down LTM traces that filter the learned expectation learn to correlate activation of $\mathcal{F}_2$ nodes with the set of all *attended* $\mathbf{X}^*$ vectors that they have ever experienced. These attended STM vectors assign less STM activity to features in the input vector $\mathbf{I}$ that mismatch the learned top-down prototype $\mathbf{V}$ than to features that match $\mathbf{V}$.

Given that ART systems learn prototypes, how can they also learn to recognize unique experiences, such as a particular view of a friend's face? The prototypes learned by ART systems accomplish this by realizing a qualitatively different concept of prototype than that offered by previous models. In particular, Gail Carpenter and I have shown with our students how ART prototypes form in a way that is designed to conjointly maximize category generalization while minimizing predictive error (Bradski and Grossberg, 1995; Carpenter and Grossberg, 1987a, 1987b; Carpenter, Grossberg and Reynolds, 1991; Carpenter, Grossberg, Markuzon, Reynolds, and Rosen, 1992). As a result, ART prototypes can automatically learn individual exemplars when environmental conditions require highly selective discriminations to be made. How the matching process achieves this is discussed below.

Before describing how this is achieved, let us note what happens if the mismatch between bottom-up and top-down information is too great for a resonance to develop. Then the $\mathcal{F}_2$ category is quickly reset and a memory search for a better category is initiated. This combination of top-down matching, attention focusing, and memory search is what stabilizes ART learning and memory in an arbitrary input environment. The attentional focusing by top-down matching prevents inputs that represent irrelevant features at $\mathcal{F}_1$ from eroding the memory of previously learned LTM prototypes. In addition, the memory search resets $\mathcal{F}_2$ categories so quickly when their prototype $\mathbf{V}$ mismatches the input vector $\mathbf{I}$ that the more slowly varying LTM traces do not have an opportunity to correlate the attended $\mathcal{F}_1$ activity vector $\mathbf{X}^*$ with them. Conversely, the resonant event, when it does occur, maintains and amplifies the matched STM activities for long enough and at high enough amplitudes for learning to occur in the LTM traces.

Whether or not a resonance occurs depends upon the level of mismatch, or novelty, that the network is prepared to tolerate. Novelty is measured by how well a given exemplar matches the prototype that its presentation evokes. The criterion of an acceptable match is defined by an internally controlled parameter that Carpenter and I have called vigilance

(Carpenter and Grossberg, 1987a). The vigilance parameter is computed in the orienting subsystem $\mathcal{A}$; see Figure 9. Vigilance weighs how similar an input exemplar $\mathbf{I}$ must be to a top-down prototype $\mathbf{V}$ in order for resonance to occur. Resonance occurs if $\rho|\mathbf{I}| - |\mathbf{X}^*| \leq 0$. This inequality says that the $\mathcal{F}_1$ attentional focus $\mathbf{X}^*$ inhibits $\mathcal{A}$ more than the input $\mathbf{I}$ excites it. If $\mathcal{A}$ remains quiet, then an $\mathcal{F}_1 \leftrightarrow \mathcal{F}_2$ resonance can develop.

Either a larger value of ρ or a smaller match ratio $|\mathbf{X}^*||\mathbf{I}|^{-1}$ makes it harder to satisfy the resonance inequality. When ρ grows so large or $|\mathbf{X}^*||\mathbf{I}|^{-1}$ is so small that $\rho|\mathbf{I}| - |\mathbf{X}^*| > 0$, then $\mathcal{A}$ generates an arousal burst, or novelty wave, that resets the STM pattern across $\mathcal{F}_2$ and initiates a bout of hypothesis testing, or memory search. During search, the orienting subsystem interacts with the attentional subsystem (Figures 10C and 10D) to rapidly reset mismatched categories and to select better $\mathcal{F}_2$ representations with which to categorize novel events at $\mathcal{F}_1$, without risking unselective forgetting of previous knowledge. Search may select a familiar category if its prototype is similar enough to the input to satisfy the resonance criterion. The prototype may then be refined by attentional focusing. If the input is too different from any previously learned prototype, then an uncommitted population of $\mathcal{F}_2$ cells is selected and learning of a new category is initiated.

Because vigilance can vary across learning trials, recognition categories capable of encoding widely differing degrees of generalization or abstraction can be learned by a single ART system. Low vigilance leads to broad generalization and abstract prototypes. High vigilance leads to narrow generalization and to prototypes that represent fewer input exemplars, even a single exemplar. Thus a single ART system may be used, say, to learn abstract prototypes with which to recognize abstract categories of faces and dogs, as well as "exemplar prototypes" with which to recognize individual faces and dogs. A single system can learn both, as the need arises, by increasing vigilance just enough to activate $\mathcal{A}$ if a previous categorization leads to a predictive error. Thus the contents of a conscious percept can be modified by environmentally sensitive vigilance control.

Vigilance control hereby allows ART to overcome some fundamental difficulties that have been faced by classical exemplar and prototype theories of learning and recognition. Classical exemplar models face a serious combinatorial explosion, since they need to suppose that all experienced exemplars are somehow stored in memory and searched during performance. Classical prototype theories face the problem that they find it hard to explain how individual exemplars are learned, such as a particular view of a familiar face. Vigilance control enables ART to achieve the best of both types of model, by selecting the most general category that is consistent with environmental feedback. If that category is an exemplar, then a "very vigilant" ART model can learn it. If the category is at an intermediate level of generalization, then the ART model can learn it by having the vigilance value track the level of match between the current exemplar and the prototype that it activates. In every instance, the model tries to learn the most general category that is consistent with the data. This tendency can, for example, lead to the type of overgeneralization that is seen in young children until further learning leads to category refinement (Chapman *et al.*, 1986; Clark, 1973; Smith *et al.*, 1985; Smith and Kemler, 1978; Ward, 1983). Many benchmark studies of how ART uses vigilance control to classify complex data bases have shown that the number of ART categories that is learned scales well with the complexity of the input data; see Carpenter and Grossberg (1994) for a list of illustrative benchmark studies.

Corticohippocampal Interactions and Medial Temporal Amnesia

As sequences of inputs are practiced over learning trials, the search process eventually converges upon stable categories. Carpenter and I mathematically proved (Carpenter and Grossberg, 1987a) that familiar inputs directly access the category whose prototype provides the globally best match, while unfamiliar inputs engage the orienting subsystem to trigger memory searches for better categories until they become familiar. This process continues until the memory capacity, which can be chosen arbitrarily large, is fully utilized. The process whereby search is automatically disengaged is a form of memory consolidation that emerges from network interactions. Emergent consolidation does not preclude structural consolidation at individual cells, since the amplified and prolonged activities that subserve a resonance may be a trigger for learning-dependent cellular processes, such as protein synthesis and transmitter production. It has also been shown that the adaptive weights which are learned by an ART model at any stage of learning can be translated into IF-THEN rules (e.g., Carpenter *et al.*, 1992). Thus the ART model is a self-organizing rule-discovering production system as well as a neural network.

The attentional subsystem of ART has been used to model aspects of inferotemporal (IT) cortex, and the orienting subsystem models part of the hippocampal system. The interpretation of ART dynamics in terms of IT cortex led Miller, Li, and Desimone (1991) to successfully test the prediction that cells in monkey IT cortex are reset after each trial in a working memory task. To illustrate the implications of an ART interpretation of IT-hippocampal interactions, I will review how a lesion of the ART model's orienting subsystem creates a formal memory disorder with symptoms much like the medial temporal amnesia that is caused in animals and human patients after hippocampal system lesions (Carpenter and Grossberg, 1994; Grossberg and Merrill, 1996). In particular, such a lesion *in vivo* causes unlimited anterograde amnesia; limited retrograde amnesia; failure of consolidation; tendency to learn the first event in a series; abnormal reactions to novelty, including perseverative reactions; normal priming; and normal information processing of familiar events (Cohen, 1984; Graf, Squire, and Mandler, 1984; Lynch, McGaugh, and Weinberger, 1984; Squire and Butters, 1984; Squire and Cohen, 1984; Warrington and Weiskrantz, 1974; Zola-Morgan and Squire, 1990).

Unlimited anterograde amnesia occurs because the network cannot carry out the memory search to learn a new recognition code. Limited retrograde amnesia occurs because familiar events can directly access correct recognition codes. Before events become familiar, memory consolidation occurs which utilizes the orienting subsystem (Figure 10C). This failure of consolidation does not necessarily prevent learning *per se*. Instead, learning influences the first recognition category activated by bottom-up processing, much as amnesics are particularly strongly wedded to the first response they learn. Perseverative reactions can occur because the orienting subsystem cannot reset sensory representations or top-down expectations that may be persistently mismatched by bottom-up cues. The inability to search memory prevents ART from discovering more appropriate stimulus combinations to attend. Normal priming occurs because it is mediated by the attentional subsystem.

Similar behavioral problems have been identified in hippocampectomized monkeys. Gaffan (1985) noted that fornix transection "impairs ability to change an established habit ... in a different set of circumstances that is similar to the first and therefore liable to be confused with it." In ART, a defective orienting subsystem prevents the memory search whereby different representations could be learned for similar events. Pribram (1986) called such a process a "competence for recombinant context-sensitive processing." These ART mechanisms illustrate how, as Zola-Morgan and Squire (1990) have reported, memory consolidation and novelty detection may be mediated by the same neural structures. Why hippocampectomized rats have difficulty orienting to novel cues and why there is a progressive reduction in novelty-related hippocampal potentials as learning proceeds in normal rats is also clarified (Deadwyler, West, and Lunch, 1979; Deadwyler, West, and Robinson, 1981). In ART, the orienting system is automatically disengaged as events become familiar during the memory consolidation process. The ART model of normal and abnormal recognition learning and memory is compared with several other recent models of these phenomena in Grossberg and Merrill (1996).

At this point, it might also be useful to note that the processes of automatic and task-selective attention may not be independent *in vivo*. This is because higher-order attentional constraints, that may be under task-selective control, can in principle propagate downwards through successive cortical levels via layer 6-to-layer 6 linkages. For example, recent modeling work has suggested how prestriate cortical areas may separate visual objects from one another and from their backgrounds during the process of figure-ground separation (Grossberg, 1994, 1997; Grossberg and McLoughlin, 1997). Such constraints may propagate top-down towards earlier cortical levels, possibly even area V1, to modulate the cells that get active there to be consistent with these figure-ground constraints. Still higher cortical processes, such as those involved in learned categorization, may also propagate their modulatory constraints to lower levels. How the strength of such top-down modulatory influences depends upon the source cortical area and on the number of synaptic steps to the target cortical area is a topic that has yet to be systematically studied.

How Universal are ART Processes in the Brain?

In all the examples discussed above—from early vision, visual object recognition, auditory streaming, and speech recognition—ART matching and resonance have played a central role in models that help to explain how the brain stabilizes its learned adaptations in response to changing environmental conditions. This type of matching can be achieved using a top-down nonspecific inhibitory gain control that down-regulates all target cells except those that also receive top-down specific excitatory signals, as in Figure 8. Are there yet other brain processes that utilize these mechanisms?

With my colleagues Mario Aguilar, Dan Bullock, and Karen Roberts, a neural model has been developed to explain how the superior colliculus learns to use visual, auditory, somatosensory, and planned movement signals to control saccadic eye movements (Grossberg, Roberts, Aguilar, and Bullock, 1997). This model uses ART matching and resonance to help explain behavioral and neural data about multimodal eye movement control. The model clarifies how visual, auditory, and planned movement signals use learning to form a mutually consistent movement map, and how attention gets focused on a movement target location after all these signals compete to determine where the eyes will move.

Nobuo Suga and his colleagues (Gao and Suga, 1998; Yan and Suga, 1998) have recently reported neurophysiological data showing how ART-like top-down matching signals can modulate learning in the auditory system.

Recent experiments from Marcus Raichle's lab at Washington University using positron emission tomography (PET) support the idea that ART top-down priming also occurs in human somatosensory cortex (Drevets, Burton, and Raichle, 1995). In their experiments, attending to an impending stimulus to the fingers caused inhibition of nearly cortical cells that code for the face, but not cells that code the fingers. Likewise, priming of the toes produced inhibition of nearby cells that code for the fingers and face, but not cells that code for the toes.

ART models have also been used to explain a great deal of data about cognitive-emotional interactions, notably about classical and instrumental conditioning (Grossberg, 1987b) and about human decision making under risk (Grossberg and Gutowski, 1987). In these examples, the resonances are between cognitive and emotional circuits, and help to focus attention upon, and release actions towards, valued events in the world.

Thus all levels of vision, visual object recognition, auditory preprocessing and streaming, speech recognition, attentive selection of eye movement targets, somatosensory representation, and cognitive-emotional interactions may all incorporate variants of the circuit depicted in Figure 8. These results suggest that a type of "automatic" attention operates even at early levels of brain processing, such as the lateral geniculate, but that higher processing levels benefit from an orienting subsystem that can be used to flexibly reset attention and to facilitate voluntary control of top-down expectations.

Internal Fantasy, Planned Movement, and Volitional Gating

Given this type of circuit, how could top-down priming be released from inhibition to enable us to voluntarily experience internal thinking and fantasies? This can be achieved through an "act of will" that activates inhibitory cells which inhibit the nonspecific inhibitory interneurons in the top-down on-center off-surround network of Figure 8. This operation disinhibits the cells receiving the excitatory top-down signals in the on-center of the network. These cells are then free to generate supraliminal resonances. Such self-initiated resonances can, for example, be initiated by the read-out of top-down expectations from higher-order planning nodes into temporally organized working memories, say in the prefrontal cortex (Fuster, 1996). It is, for example, well-known that the basal ganglia can use such a disinhibitory action to gate the release of individual movements, sequences of movements, and even cognitive processes (Hikosaka, 1994; Middleton and Strick, 1994; Sakai *et al.*, 1998)

These examples also help to understand how top-down expectations can be used for the control of planned (viz., intentional) behavioral sequences. For example, once such planning nodes read-out their top-down expectations into working memory, the contents of working memory can be read-out and modified by on-line changes in "acts of will". These volitional signals enable invariant representations of an intentional behavior to rapidly adapt themselves to changing environmental conditions. For example, Bullock, Grossberg, and Mannes (1993) have modeled how such a working memory can control the intentional performance of handwriting whose size and speed can be modified by acts of will, without a change of handwritten form. Bullock, Grossberg, and Guenther (1994) have shown how a visual target that is stored in working memory can be reached with a novel tool that has never been used before. The latter study shows how a such a model can learn its own parameters through a type of Piagetian perform-and-test developmental cycle.

Thus we arrive at an emerging picture of how the adaptive brain works wherein the core issue of how a brain can learn quickly and stably about a changing world throughout life leads towards a mechanistic understanding of attention, intention, thinking, fantasy, and consciousness. The mediating events are adaptive resonances that effect a dynamic balance between the complementary demands of stability and plasticity, and of expectation and novelty, and which are a necessary condition for consciousness.

What vs. Where: Why Are Not Procedural Memories Conscious?

Although the type of ART matching, learning, and resonance that have been reviewed above seem to occur in many sensory and cognitive processes, they are not the only types of matching and learning to occur in the brain. In fact, there seems to be a major difference between the types of learning that occur in sensory and cognitive processes versus those that occur in spatial and motor processes. In particular, sensory and cognitive processes are carried out in the What processing stream that passes through the inferotemporal cortex, whereas spatial and motor processes are carried out in the Where processing stream that passes through the parietal cortex. What processing includes object recognition and event prediction. Where processing includes spatial navigation and motor control. I suggest that the types of matching and learning that go on in the What and Where streams are different, indeed complementary, and that this difference is appropriate to their different roles. First consider how we use a sensory expectation. Suppose, for example, that I ask you to "Look for the yellow ball, and if you find it within 300 msec., I will give you $1,000,000." If you believed me, you could activate a sensory expectation of "yellow balls" that would make you much more sensitive to yellow and round objects in your environment. As in ART matching, once you detected a yellow ball, you could then react to it much more quickly and with a much more energetic response than if you were not looking for it. In other words, sensory and cognitive expectations lead to a type of excitatory matching.

Now consider how we use a motor expectation. Such an expectation represents where we want to move (Bullock and Grossberg, 1988). For example, it could represent a desired position for the hand to pick up an object. Such a motor expectation is matched against where the hand is now. After the hand actually moves to the desired position, no further movement is required to satisfy the motor expectation. In this sense, motor expectations lead to a type of inhibitory matching. In summary, although the sensory and cognitive matching process is excitatory, the spatial and motor matching process is inhibitory. These are complementary properties. Models such as ART quantify how excitatory matching is accomplished. A different type of model, called a Vector Associative Map, or VAM, model, suggests how inhibitory matching is accomplished (Gaudiano and Grossberg, 1991; Grossberg, Guenther, Bullock, and Greve, 1993; Guenther, Bullock, Greve, and Grossberg, 1994).

As shown in the discussions of ART above, learning within the sensory and cognitive domain is often a type of *match* learning. It takes place only if there is a good enough match of top-down expectations with bottom-up data to risk altering previously stored knowledge within the system, or it can trigger learning of a new representation if a good enough match is not available. In contrast, learning within spatial and motor processes, such as VAM processes, is *mismatch* learning that is used to either learn new sensory-motor maps (e.g., Grossberg, Guenther, Bullock, and Greve, 1993) or to adjust the gains of sensory-motor commands (e.g., Fiala, Grossberg, and Bullock, 1996). These types of learning are also complementary.

Why are the types of learning that go into spatial and motor processes complementary to those that are used for sensory and cognitive processing? My answer is that ART-like learning allows the brain to solve the stability-plasticity dilemma. It enables us to continue learning more about the world in a stable fashion throughout life without forcing catastrophic forgetting of our previous memories. On the other hand, catastrophic forgetting is a good property when it takes place during spatial and motor learning. We have no need to remember all the spatial and motor maps that we used when we were infants or children. In fact, those maps would cause us a lot of trouble if they were used to control our adult limbs. We want our spatial and motor processes to continuously adapt to changes in our motor apparatus. These complementary types of learning allow our sensory and cognitive systems to stably learn about the world and to thereby be able to effectively control spatial and motor processes that continually update themselves to deal with changing conditions in our limbs.

Why, then, are procedural memories unconscious? The difference between cognitive memories and procedural, or motor, memories has gone by a number of different names, including the distinction between declarative memory and procedural memory, knowing that and knowing how, memory and habit, or memory with record and memory without record (Bruner, 1969; Miskin, 1982, 1993; Ryle, 1949; Squire and Cohen, 1984). The amnesic patient HM dramatically illustrated this distinction by learning and remembering motor skills like assembly of the Tower of Hanoi without being able to recall ever having done so (Bruner, 1969; Scoville and Milner, 1957; Squire and Cohen, 1984). We can now give a very short answer to the question of why procedural memories are unconscious: The matching that takes place during spatial and motor processing is often inhibitory matching. Such a matching process cannot support an excitatory resonance. Hence, it cannot support consciousness.

In this regard, Goodale and Milner (1992) have described a patient whose brain lesion has prevented accurate visual discrimination of object orientation, yet whose visually guided reaching behaviors towards objects are oriented and sized correctly. We have shown, in a series of articles, how head-centered and body-centered representations of an object's spatial location and orientation may be learned and used to control reaches of the hand-arm system that can continuously adapt themselves to changes in the sensory and motor apparatus that is used to plan and execute reaching behaviors (Bullock, Grossberg, and Guenther, 1993; Carpenter, Grossberg, and Lesher, 1998; Gaudiano and Grossberg, 1991; Grossberg, Guenther, Bullock, and Greve, 1993; Guenther, Bullock, Greve, and Grossberg, 1994). None of these model circuits has resonant loops; hence, they do not support consciousness.

When these models are combined into a more comprehensive system architecture for intelligent behavior, the sensory and cognitive match-based networks in the What processing stream through the inferotemporal cortex provide self-stabilizing representations with which to continually learn more about the world without undergoing catastrophic forgetting, while the Where/How processing stream's spatial and motor mismatch-based maps and gains can continually forget their old parameters in order to instate the new parameters that are needed to control our bodies in their present form. This larger architecture illustrates how circuits in the self-stabilizing match-based sensory and cognitive parts of the brain can resonate into consciousness, even while they are helping to direct the contextually appropriate activation of spatial and motor circuits that cannot.

REFERENCES

Abbott, L.F., Varela, K., Sen, K., and Nelson, S.B. (1997). Synaptic depression and cortical gain control. *Science*, **275**, 220–223.

Artola, A. and Singer, W. (1993). Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends in Neurosciences*, **16**, 480–487.

Bradski, G. and Grossberg, G. (1995). Fast learning VIEWNET architectures for recognizing 3-D objects from multiple 2-D views. *Neural Networks* **8**, 1053–1080.

Bregman, A.S. (1990). Auditory scene analysis: The perceptual organization of sound. Cambridge, MA: MIT Press.

Bruner, J.S. (1969). Chapter in G.A. Talland and N.C. Waugh (Eds), **The pathology of memory**. New York: Academic Press.

Bullock, D. and Grossberg, S. (1988). Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review*, **95**, 49–90.

Bullock, D., Grossberg, S. and Guenther, F.H. (1993). A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm. *Journal of Cognitive Neuroscience*, **5**, 408–435.

Bullock, D., Grossberg, S. and Mannes, C. (1993). A neural network model for cursive script production. *Biological Cybernetics*, **70**, 15–28.

Carpenter, G.A. and Grossberg, S. (1981). Adaptation and transmitter gating in vertebrate photoreceptors. *Journal of Theoretical Neurobiology*, **1**, 1–42. Reprinted in S. Grossberg (Ed.), **The adaptive brain, II**. Amsterdam: Elsevier Publishing, 1987, 271–310.

Carpenter, G.A. and Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54–115.

Carpenter, G.A. and Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, **26**, 4919–4930.

Carpenter, G.A. and Grossberg, S. (1993) Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, **116**, 131–137.

Carpenter, G.A. and Grossberg, S. (1994). Integrating symbolic and neural processing in a self-organizing architecture for pattern recognition and prediction. In V. Honavar and L. Uhr (Eds.), **Artificial intelligence and neural networks: Steps towards principled prediction**. San Diego: Academic Press, 387–421.

Carpenter, G.A., Grossberg, S. and Lesher, G.W. (1998). The What-and-Where filter: A spatial mapping neural network for object recognition and image understanding. *Computer Vision and Image Understanding*, **69**, 1-22.

Carpenter, G.A., Grossberg, S. Markuzon, N., Reynolds, J.H., and Rosen, D.B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, **3**, 698–713.

Carpenter, G.A., Grossberg, S. and Reynolds, J.H. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, **4**, 565–588.

Chapman, K.L., Leonard, L.B., and Mervis, C.G. (1986). The effect of feedback on young children's inappropriate word usage. *Journal of Child Language*, **13**, 101-107.

Chey, J., Grossberg, S. and Mingolla, E. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America*, **14**, 2570–2594

Clark, E.V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T.E. Morre (Ed.), **Cognitive development and the acquisition of language**. New York: Academic Press, 65-110.

Cohen, N.J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. Squire and N. Butters (Eds.), **The neuropsychology of memory**, New York: Guilford Press, pp. 83–103.

Cohen, N.J. and Squire, L.R. (1980). Preserved learning and retention of a pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, **210**, 207–210.

Deadwyler, S.A., West, M.O., and Lynch, G. (1979). Activity of dentate granule cells during learning: Differentiation of perforant path inputs. *Brain Research*, **169**, 29–43;

Deadwyler, S.A., West, M.O., and Robinson, J.H. (1981). Entorhinal and septal inputs differentially control sensory-evoked responses in the rat dentate gyrus. *Science*, **211**, 1181–1183.

Drevets, W.C., Burton, H., and Raichle, M.E. (1995). Blood flow changes in human somatosensory cortex during anticipated stimulation. *Nature*, **373**, 249.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H.J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, **60**, 121–130.

Fiala, J.C., Grossberg, S., and Bullock, D. (1996). Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eyeblink response. *Journal of Neuroscience*, **16**, 3760–3774.

Francis, G. and Grossberg, S. (1996a). Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception*, **35**, 543–567.

Francis, G. and Grossberg, S. (1996b). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours. *Vision Research*, **36**, 149–173.

Francis, G., Grossberg, S., and Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, **34**, 1089–1104.

Fuster, J.M. (1996). Frontal lobe and the cognitive foundation of behavioral action. In A.R. Damasio, H. Damasio, and Y, Christen (Eds.) **The neurobiology of decision-making**. New York: Springer, 47–61.

Gaffan, D. (1985). Hippocampus: Memory, habit, and voluntary movement. *Philosophical Transactions of the Royal Society of London*, **B308**, 87–99.

Gao, E. and Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences USA* **95**, 12663–12670.

Gaudiano, P. and Grossberg, S. (1991). Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks*, **4**, 147–183.

Goodale, M.A. and Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20–25.

Gove, A., S. Grossberg and E. Mingolla. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, **12**, 1027–1052.

Govindarajan, K.K., Grossberg, S., Wyse, L.L., and Cohen, M.A. (1995). A neural network model of auditory scene analysis and source segregation. Technical Report: CAS/CNS-TR-94-039. Boston, MA: Boston University.

Graf, P., Squire, L.R., and Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 164–178. -183.

Gray, C.M. and Singer, W. 1989. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*, 86:1698–1702.

Grossberg, S. (1964). *The theory of embedding fields with applications to psychology and neurophysiology*. New York: Rockefeller University.

Grossberg, S. (1969). On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, **1**, 319–350.

Grossberg, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49–57.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, 1976, **23**, 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, 1976, **23**, 187–202.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.), **Progress in theoretical biology, Volume 5**. New York: Academic Press, 1978.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **1**, 1–51. How does a brain build a cognitive code? *Psychological Review*, 1980, **1**, 1–51.

Grossberg, S. (1982). **Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control**. Norwell, MA: Kluwer Academic Publishers.

Grossberg, S. (1987a). Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory. *Perception and Psychophysics*, **41**, 117–158.

Grossberg, S. (1987b). **The adaptive brain, I**. Amsterdam: North-Holland.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17–61.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, 1994, **55**, 48–120.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional figures. *Psychological Review*, 1997, **104**, 618–658.

Grossberg, S. (1998a). How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. Technical report CAS/CNS TR-97-023. Boston, MA: Boston University. *Spatial Vision*, in press.

Grossberg, S. (1998b). Pitch based streaming in auditory perception. In N. Griffith and P. Todd (Eds.), **Musical networks: Parallel distributed perception and performance**. Cambridge, MA: MIT Press, in press.

Grossberg, S., Boardman, I., and Cohen, M.A. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, **23**, 481–503.

Grossberg, S. and Grunewald, A. (1997). Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience*, **9**, 117-132.

Grossberg, S., Guenther, F., Bullock, D., and Greve, D. (1993). Neural representations for sensory-motor control, II: Learning a head-centered visuomotor representation of 3-D target position. *Neural Networks*, 1993, **6**, 43–67.

Grossberg, S. and Gutowski, W. (1987). Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review*, 1987, **94**, 300–318.

Grossberg, S. and Kuperstein, M. (1986). **Neural dynamics of adaptive sensory-motor control: Ballistic eye movements**. Amsterdam: North-Holland.

Grossberg, S. and McLoughlin, N.P. (1997). Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks*, **10**, 1583–1605.

Grossberg, S. and Merrill, J.W.L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience*, **3**, 257–277.

Grossberg, S., Roberts, K., Aguilar, M., and Bullock, D. (1997) A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience*, **17**, 9706–9725.

Grossberg, S. and Williamson, J.R. (1997). Linking cortical development to visual perception. *Society for Neuroscience Abstract*, **23**(1), pp. 568, Number 227.9.

Grossberg, S. and Williamson, J.R. (1998). A neural model of how visual cortex develops a laminar architecture capable of adult perceptual grouping. Technical Report CAS/CNS TR-98-022. Boston, MA: Boston University.

Grunewald, A. and Grossberg, S. (1998). Self-Organization of binocular disparity tuning by reciprocal corticogeniculate interactions. *Journal of Cognitive Neuroscience*, **10**, 199–215.

Guenther, F.H., Bullock, D., Greve, D. and Grossberg, S. (1994). Neural representations for sensory-motor control, III: Learning a body-centered representation of 3-D target position. *Journal of Cognitive Neuroscience*, **6**, 341–358.

Helmholtz, H.L.F. von (1962). **Treatise on physiological optics**, J.P.C. Southall (translator). New York: Dover.

Hikosaka, O. (1994). Role of basal ganglia in control of innate movements, learned behavior and cognition–a hypothesis. In G. Percheron, J.S. McKenzie, and J. Feger (Eds.). **The basal ganglia, IV**. New York: Plenum Press, 589-595.

Hupé, J.M., James, A.C., Girard, P. and Bullier, J. (1997). Feedback connections from V2 modulate intrinsic connectivity within. *Society for Neuroscience Abstracts*, **406.15**, 1031.

Ito, M., Westheimer, G. and Gilbert, D.C. (1997). Attention modulates the influence of context on spatial integration in V1 of alert monkeys. *Society for Neuroscience Abstracts* **603.2**, 1031.

Johnson, R.R. and Burkhalter, A. (1997). A circuit for amplification of excitatory feedback input from rat extrastriate cortex to primary visual cortex. *Society for Neuroscience Abstracts*, **651.7**, 1669.

Kohonen, T. (1984). **Self-organization and associative memory**. New York: Springer-Verlag.

Lamme, V.A.F., Zipser, K. and Spekreijse, H. (1997). Figure-ground signals in V1 depend on consciousness and feedback from extra-striate areas. *Society for Neuroscience Abstracts*, **603.1**, 1543.

Land, E. (1977). The retinex theory of color vision. *Scientific American*, **237**, 108–128.

Land, E. (1986). Recent advances in retinex theory. *Vision Research*, **26**, 7–21.

Levy, W.B. (1985). Associative changes at the synapse: LTP in the hippocampus. In W.B. Levy, J. Anderson, and S. Lehmkuhle (Eds.), **Synaptic modification, neuron selectivity, and nervous system organization**. Hillsdale, NJ: Erlbaum, 5–33.

Levy, W.B. and Desmond, N.L. 1985. The rules of elemental synaptic plasticity. In W.B. Levy, J. Anderson, and S. Lehmkuhle (Eds.), **Synaptic modification, neuron selectivity, and nervous system organization**, Hillsdale, NJ: Erlbaum Associates, 105–121.

Livingstone, M.S. and Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, **4**, 309–356.

Logothetis, N.K., Leopold, D.A., and Sheinberg, D.L. (1996) What is rivalling during binocular rivalry? *Nature*, **360**, 621–624.

Lynch, G., McGaugh, J.L., and Weinberger, N.M. (Eds.) (1984). **Neurobiology of learning and memory**. New York: Guilford Press.

McAdams, C.J. and Maunsell, J.H.R. (1997). Spatial attention and feature-directed attention can both modulate neuronal responses in macaque area V4. *Society for Neuroscience Abstracts*, **802.5**, 2062.

Michotte, A., Thines, G., and Crabbe, G. (1964). **Les complements amodaux des structures perceptives**. Louvain: Publications Universitaires de Louvain.

Middleton, F.A. and Strick, P.L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, **166**, 458–461.

Miller, E.K., Li, L., and Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, **254**, 1377–1379.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions Royal Society of London, B*, **298**, 85–95.

Mishkin, M. (1993). Cerebral memory circuits. In T.A. Poggio and D.A. Glaser (Eds.), **Exploring brain functions: Models in neuroscience** New York: Wiley and Sons, 113–125.

Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, **6**, 414–417.

Motter, B.C. (1994a). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience*, **14**, 2178–2189.

Motter, B.C. (1994b). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*, **14**, 2190–2199.

Murphy, P.C. and Sillito, A.M. (1987). Corticofugal feedback influences the generation of length tuning in the visual pathway. *Nature*, **329**, 727–729.

Parker, D.B. (1982, October). Learning-logic (Invention Report 581-64, File 1). Office of Technology Licensing. Stanford University.

Press, W.A. and van Essen, D.C. (1997). Attentional modulation of neuronal responses in macaque area V1. *Society for Neuroscience Abstracts*, **405.3**, 1026.

Pribram, K.H. (1986). The hippocampal system and recombinant processing. In R.L. Isaacson and K.H. Pribram, (Eds.), **The hippocampus, Volume 4**, New York: Plenum Press, 329–370.

Rauschecker, J.P. and Singer, W., (1979). Changes in the circuitry of the kitten's visual cortex are gated by postsynaptic activity. *Nature*, **280**, 58–60.

Redies, C., Crook, J.M., and Creutzfeldt, O.D. (1986). Neuronal responses to borders with and without luminance gradients in cat visual cortex and dLGN. *Experimental Brain Research*, **61**, 469–481.

Repp, B. H. 1980. A range-frequency effect on perception of silence in speech. *Haskins Laboratories Status Report on Speech Research*, **SR–61**, 151–165.

Reynolds, J., Nicholas, J., Chelazzi, L., and Desimone, R. (1995). Spatial attention protects macaque V2 and V4 cells from the influence of non-attended stimuli. *Society for Neuroscience Abstracts*, **21.3**, 1759.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhard and J.L. McClelland (Eds.). **Parallel distributed processing**. Cambridge, MA: MIT Press.

Ryle, G. (1949). **The concept of mind**. Hutchinson Press.

Sakai, K. Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y, and Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. *Journal of Neuroscience*, **18**, 1827–1840.

Samuel, A.G. (1981a). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 1124–1131.

Samuel, A.G. (1981b). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110**, 474–494.

Schiller, P.H. (1994). Area V4 of the primate visual cortex. *Current Directions in Psychological Science*, **3**, 89–92.

Scoville, W.B. and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesion. *Journal of Neurology, Neurosurgery, and Psychiatry*, **20**, 11–21.

Sillito, A. M., H. E. Jones, G. L. Gerstein and D. C. West. (1994). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, **369**, 479–482.

Singer, W. (1983). Neuronal activity as a shaping factor in the self-organization of neuron assemblies. In E. Basar, H. Flohr, H. Haken, and A.J. Mandell (Eds.), **Synergetics of the brain**. New York: Springer-Verlag.

Smith, C., Carey, S., and Wiser, M. (1985). On differentiation: A case study of the development of the concept of size, weight, and density. *Cognition*, **21**, 177–237.

Smith, L.B. and Kemler, D.G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, **10**, 502–532.

Squire, L.R. and Butters, N. (Eds.) (1984). **Neuropsychology of memory**. New York: Guilford Press.

Squire, L.R. and Cohen, N.J. (1984). Human memory and amnesia. In G. Lunch, J. McGaugh, and N.M. Weinberger (Eds.), **Neurobiology of learning and memory**, New York, NY: Guilford Press, 3–64.

Ungerleider, L.G. and Mishkin, M. (1982). Two cortical visual systems: Separation of appearance and location of objects. In D.L. Ingle, M.A. Goodale, and R.J.W. Mansfield (Eds.), **Analysis of visual behavior**. Cambridge, MA: MIT Press, 549–586.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, **14**, 85–100.

Ward, T.B. (1983). Response temp and separable-integral responding: Evidence for an integral-to-separable processing sequencing in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, **9**, 1029–1051.

Warren, R.M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, **96**, 371–383.

Warren, R.M. and Sherman, G.L. (1974). Phonemic restorations based on subsequent context. *Perception and Psychophysics*, **16**, 150–156.

Warrington, E.K. and Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychology*, **12**, 419–428.

Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. Unpublished doctoral thesis, Harvard University, Cambridge, MA.

Willshaw, D.J. and Malsburg, C. von der. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London (B)*, **194**, 431–445.

Yan, W. and Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, **1**, 54–58.

Zola-Morgan, S.M. and Squire, L.R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science*, **250**, 288–290.