19 Mechanisms and Capabilities in Reading[1]

19.1 Introduction

15.4.1A The routine versus learning states of operation

[xxx may want to abbreviate this section and move part concerning reading to follow 15.6.6 & 15.6.7]
Each of the couples defined in **Section 15.6.5** [xxx out of order] can perform in conjunction with two different state diagrams. In the routine case, any signal sensed and reported to the old brain will be interpreted and reported to the appropriate portion of the *neo*-cortex as a percept according to the routine state diagram for that modality. However, whenever a signal is sensed that can not be interpreted by the appropriate couple, a supervisory signal will be sent to the TRN. The TRN will in turn change the operating mode to follow the "learning" state diagram. This state diagram is more involved and employs more elements of the CNS. The resulting procedure is slower in execution but it results in a new interp being stored in the appropriate memory element, and being available for future interpretation via the routine state diagram.

15.4.1A.1 The generic aspects of learning and routine operation EDIT

[xxx edit out ref to cerebellum specifically]
Based on this work, a critically important role for the pulvinar can be defined. That role is one of a major memory function (a very large lookup table). Its primary responsibility is to provide a translation service between low level and high level signaling languages. The role is very similar to that used in modern computers where a translation table is required between machine language (ones and zeros) and the stilted English of higher level languages like C++. Such a table is frequently required to be bi-directional. This is usually achieved by using two separate sections of the memory to perform the translations in different directions. In the visual system, the pulvinar has the additional task, in conjunction with other elements, of learning the required translation routines initially.

To accomplish both the learning and the translation of the appropriate signals, the pulvinar can be considered a special type of memory device called a write once, read many times (or read only) memory unit. At the technical level, such a man made device is called a PROM, a programmable read only device. The use of such a device can greatly reduce the computational requirement associated with routine tasks.

In the case of vision, the main role of the pulvinar is proposed to be the conversion of a mid-level signaling language generated by the two-dimensional correlator of the pretectum (expressing interps) into a higher level language (expressing percepts) that is then transmitted to the cerebrum for high level cognition. This role is suggested by Noback when he observes on page 146, "there are approximately three times as many cerebellar afferent fibers as cerebellar efferent fibers."

In its role in oculo-motor and skeletal-motor operations, it plays the opposite role. It converts high level language signals from the cerebrum into mid level language signals that can be processed by the various motor neuron nuclei into the low level signals required to operate the various muscles. The conversion from a high level to a mid level language introduces a greater degree of detail into the commands. This level of detail results in both smoother and more highly coordinated movements.

---

[1] Released March 4, 2003

The question remains, how does the pulvinar become programmed? A proposed method is illustrated by the state diagram of **Figure 19.1.1-1**. The diagram illustrates two main paths. The upper path is used when the pulvinar fails to recognize a signaling sequence sent to it. This initiates the complex sequence required to learn the meaning of the sequence. The second is the simpler case where the pulvinar recognizes the signal sequence, based on prior learning, and issues the appropriate signals. These commands are typically condensed and higher level in the case of vision, and expanded lower level commands in the case of motor operations. Only the vision case is illustrated.

It is noteworthy that if the pulvinar is damaged following learning, the system can revert to the direct mode of connection between the pretectum and the cerebrum used during initial learning. The system will operate slowly, but it will operate. This mode is characterized by inadequate coordination in skeletal-motor activities. These motions are frequently tentative, not unlike those of a very young child.

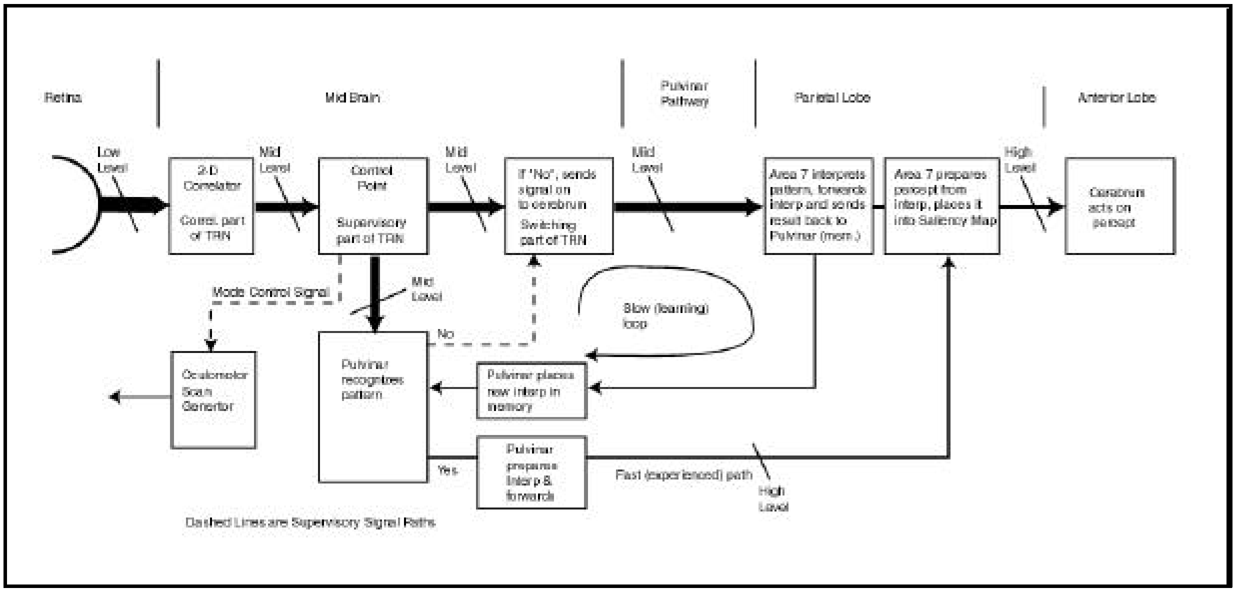

Figure 19.1.1-1 Candidate state diagram describing the operation of the pulvinar and thalamic reticular nucleus. Moving left to right, the information density rises as the data volume decreases. The information loses all spatial correlation to the scene when converted to an interp. A "no" decision by the pulvinar is probably hierarchal. For efficient mode control, a "yes" decision is probably sent to the control point as well. See text.

The pathways shown in this figure are part of the analytical mode of vision. Physically, they utilize the Pulvinar pathway and do not involve areas 17 through 22 of the cerebrum..

15.4.1A.1.1 "Experienced" operation of the pulvinar in vision

When knowledge of a pattern is available from previous experience, the pulvinar can convert a mid-level input into a high level output very quickly, on the order of the cycle time of the correlator (approximately 50 ms). In situations, such as reading, where a fixed symbol set is used, performance of the pulvinar is routine and few errors occur. Little effort is spent on recognition of the syllable stream and the available effort can be spent on analyzing the ideas associated with a paragraph. However, if the symbol set is more diverse or unusual, the pulvinar may issue a "no recognition" signal to the control point. There may also be levels of "no." It may recognize the object as a female, as an older female, as a grandmother; but will it label the scene as "my grandmother?" This type of situation is the source of the comment, "I did not recognize you initially."

15.4.1A.1.2 "Learning mode" of operation of the pulvinar

If the pulvinar issues an emphatic "no recognition," the visual system reverts to the learning mode. This mode is slower and requires greater mental concentration. In this case, the control point must send the mid level data from the correlator to the cerebrum for cognition leading to a definition (in mid-level language) of the object (an interp) before sending a high level definition of the object (a percept) on to the higher cognitive centers for action. Both the interp and the percept must also be sent back to the control point for entry into the proper memory location of the pulvinar. In the figure, the learning loop is shown returning to the pulvinar directly for convenience.

Common examples of these learning cases include;

‣the printing of "Jesus" in white on a black background and without a border. Cognition is required the first time to tell the pulvinar this is the same as a similar word printed black on white and with a white surround.

‣encountering old German text where the f's and the s's are interchanged. Cognition is required the first time to tell the pulvinar to interpret the syllable by interchanging its translation of the individual characters.

If the pulvinar issues a conditional "no recognition" signal, the learning mode will also be entered in an attempt to more succinctly define the object under examination. This type of situation may be as simple as examining a word to determine if it is misspelled or a new word with a totally new meaning.

15.4.7 Higher Levels of Perception–Introduction to reading & verbalization

Although the computational aspects of higher levels of perception are beyond the scope of this work, it is possible to draw some general guidelines for the structural aspects of this level of perception within the brain based on the fact that the retina is generally regarded as an extension of the brain surface. The general conclusions would consist of:

+ the mechanisms of the brain are based on electrolytic chemical principles which allow the brain to achieve its remarkable thermal efficiency by utilizing non-dissapative, reversible thermodynamic processes.

+ the "conductors" of the brain are non-metallic, electrolytic paths which exhibit charge transport velocities on the order of 10 meters per second (versus 10^8 meters per second for metallic conductors).

+ the principle of feedback is utilized frequently in its internal form. External feedback paths around an amplifier are used infrequently if at all.

+ the structural foundation of the brain is the activa, the liquid crystalline based electrolytic transistor.

+ each neuron contains at least one activa at the junction of the dendritic and the axonal structures, which may or may not be located within the soma; the soma is primarily a housekeeping apparatus.

+ the neurons of the brain are interconnected primarily by "gap junctions" which are actually the active regions of additional electrolytic semiconductor devices, activas.

+ the activa of the brain are used primarily in analog signal processing, except where long distance communication is required, on the order of two millimeters or more.

+ the analog activa is easily converted into an action potential generator by the addition of a significant size capacitor to either the dendritic or axonal terminal of the device and the implementation of an larger impedance in the poda lead.

+ the differential amplifier configuration found in the photoreceptor cells of the eye are easily adapted to a two stable state device that could be used for long term storage in the brain.

15.4.7.1 Modes of vision using nomenclature of psychology

Koch & Ullman have provided an analysis of the modes of vision from a largely psychological perspective[2]. They review the literature in terms of a two-stage model. The first stage is described as a "preattentive" mode. The second is an "attentive" mode [their emphasis]. They then concentrate on further analyzing the attentive mode, under the label selective visual attention. They suggest this mode can be further divided into three operational stages. First, a set of elementary features is computed in parallel across the visual field and is represented in a set of topographical maps. Second, a "winner-take-all" (WTA) network localizes and selects the most active unit in the above saliency map. Third, an additional network directs this localized unit to a central representation unit, presumably for detailed analysis.

Figure 19.1.2-1 XXX XXX XXX

The Koch & Ullman approach is quite compatible with this work. However, they do not appear to address the alarm mode or differentiate between the analytical mode operating in response to an alarm or to volition. Their preattentive mode conforms well to the awareness mode of this work. The discuss the "pop-out effect and other aspects of conspicuousness. The pop-out situation appears to be a clear case of autonomous alarm mode operation, although not necessarily due to a high order of danger. In the second case, searching for a feature having several attributes, the subject is clearly commanding a detailed search. This procedure is clearly compatible with the volition mode of this work. In both cases, a central engine is making choices. These can be considered "winner-take-all" with respect to the whole organism or winner-take-all with respect to only certain signal processing sequences. In this work, that engine is the thalmic reticular nucleus. Their interpretation appears to be largely sequential. It does not emphasize the parallel operation of both the awareness (preattentive) mode and the analytic mode (whether invoked by the alarm or volition mode). They do recognize a relationship between their attentive mode and the fovea but do not detail it. They define the non-retinotopic nature of the information following selection by the WTA mechanism and the subsequent attentive processing. Their attentive processing mechanism appears to conform to the combined activity of the perigeniculate nucleus and

---

[2] Koch, C. & Ullman, S. (1985) Shifts in selective visual attention: towards the underlying neural circuitry *Human Neurobiol.* vol. 4, pp 219-227

pulvinar (and potentially the cerebellum) in this work.

Their discussion contains several concepts that can be associated with the introduction of reading as well as general scene element interpretation.

15.4.7.2 Operating modes extended to reading

While the investigations by Koch & Ullman were generic to signal processing in vision, Olshausen, et. al. have focused more specifically on the analysis of symbols[3]. Their model was aimed primarily at the need to accommodate variations in scale, orientation and position in arriving at an object oriented reference framework. They do focus on the use of an associative memory as part of the processing model. Their considerations are mathematical and conceptual but do address how some of their conceptual circuits might be implemented biologically. However, their analysis does not incorporate a dual path visual system. All of their discussion centers on the retino-geniculo-striate pathway to area 17. They do focus on the generally held view that the pulvinar is a major player in controlling the routing of signals. At the next level of detail, this work assigns this role more specifically to the adjacent thalamic reticular nucleus. They summarize their model as a “zero-th order” model. It does not attempt to explain the mechanism for recognizing a symbol. They do attempt to size the spatial dimensions of the analytical (attentive) mode. However, the numbers are of “ball park” precision. They take the “window of attention” to occupy an array of only 30 by 30 elements. This appears far from the capability of the foveola and the associated analytical mode of this work. Such a small area does not appear compatible with the area perceived by the visual system during one gaze in reading.

15.4.7.2.1 Tasks involved in reading

The task of reading involves so many individual procedural steps and so many alternate paths that it can only be drawn at a coarse level of detail here. The diagram is taken from a review of the material by Kennedy in Chapter 7, Rayner, Reichle & Pollatsek in Chapter 11 of Underwood. The material of Murray in Chapter 8 is also of significance but suffers from a lack of specifics concerning the geometry and performance of the fovea and foveola.

15.4.7.2 Operating modes extended to verbalization

Pre-1980 literature generally spoke of verbalization as being centered in the left cerebral hemisphere. As an example, Trevarthen & Sperry discussed the ability to verbalize with respect to perception limited to the left or right visual field[4]. More recent MRI and PET experiments have shown the subject to be far more complex.

19.8 Reading and the interpretation of fine detail

In the following introductory discussion the functions of reading and the analysis of fine detail will be considered one. However, there is a significant difference between analyzing fine detail and analyzing the typical detail in a scene. Because reading involves nearly every facet of the visual system, it has not been described in one of the previous Chapters that focus on specific functional areas. Therefore, this section will be subdivided into four portions to include the scope of the full subject. **Section 19.8.1** will assemble a variety of relevant subjects in a background section. **Section 19.8.2** will present an overview scenario of the total reading process. **Section 19.8.3** will discuss the environment, mechanics and mechanisms involved in reading. **Section 19.8.4** will provide some performance descriptors for specific reading scenarios. There are a series of clinically significant

---

[3] Olshausen, B. Anderson, C. & Van Essen, D. (1993) A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information *J. Neurosci*. Vol. 13, no. 11, pp 4700-4719

[4] Trevarthen, C. & Sperry, R. (1973) Op. Cit.

irregularities related to reading. These are discussed in **Section 18.9**.

Reading is the ultimate mechanism and capability provided by the analytical mode of visual system operation. The process of reading is probably the most sophisticated function in vision, if not in the entire neural system. It involves all of the features of the POS operating in the most sophisticated and choreographed manner. Furthermore, reading clearly involves distinct training, operating and senescent phases. Therefore both the initial introduction to reading and the subsequent time scale become significant factors in discussing and evaluating reading performance. The interconnection of cognition related to the awareness mode of vision and the cognition related to the reading process is highly dependent on the training, memory capability and innate intelligence of the subject. This situation complicates the description of the capability level reached by a subject, and of similar subjects at a given time.

19.8.1 Background

This material will highlight a difference between the performance of selected primates (other than man) and man. As found by many experimentalists, the general premise is that the other primates mature intellectually and visually along a course similar to humans up to an equivalent age of three or four. His capability is clearly evident where chimpanzees and other primates are taught to respond to verbal and graphical materials by selecting images on a television screen, or (reversing the process) expressing their desires by selecting an image on a television screen. Rare animals may even be able to express more complex thought patterns involving a series of images on a television screen. However, these images are more pictures than writing. Change the font of the words on the screen or alternate the colors of the letters in the words and the experiment quickly comes to an end. Humans have little difficulty with such changes beyond the age of six or seven (on average).

The overall capabilities to be developed and discussed in this section as well as the individual mechanisms involved exhibit a wide range within a given human population. In nearly every case, this material will attempt to describe a typical subject unless it is clearly indicated that a special case is being discussed.

19.8.1.1 Glossary

To develop a comprehensive discussion of reading, it is necessary to introduce a variety of terms from a wide assortment of disciplines from biology, computers and the education process. Many of these terms will be defined in greater detail in the following sections.

Comprehensions– High level thoughts reflecting a high degree of cognitive understanding. Generally stored in the highly associative saliency map and only subject to change as a result of new cognitive processes.

Controller– A type of neural engine specialized to control the flow of data into and out of memories located throughout the neural system.

Default mode–In the case of reading, the preprogrammed mode assumed in the absence of feedback from a controller indicating a failure in one of the cognitive steps following initial image scanning.

Engine–A general purpose analog computational facility found in perfusion throughout the brain. They are frequently programmed as special purpose machines based on PROMS. These PROMS are initially programmed through learning and experience.

Interp–A vector description of the symbol(s) imaged on the foveola issued by the pretectum following a nominal 50 ms scanning (at the microsaccades level) and cross correlation process.

Percept– The vector created by the assembly of individual interps generated by the pretectum in response to the scanning and cross correlation process performed by the Precision Optical System.

PROMS–See Write once–read repeatedly memory.

Saliency engine–A higher level mechanism in perception that provides the addresses of pieces of perceived information that must be considered as a group to achieve cognition. The various vectors, relating to the perceived information, when grouped act as individual bytes in the overall cognition vector describing the recognized event.

Saliency Map– The general database employed by the cortex as a archive of all sensory information collected during the life of the subject. Highly associative. Not generally re-writable except under instruction from the controller following significant cognitive processes calling for such a change.

Shift register–A type of memory that is designed to accept short values (consisting of several characters) in part of its memory (defined by an address) and then accept subsequent values at other memory locations with the general intent of assembling a longer, more comprehensive, value from the sequentially received individual values. The comprehensive value can be accessed by addressing only the initial address.

Write once–read repeatedly memory– A type of memory, commonly called a programmable read-only memory (PROM) in computers, that once data is written to a specific address in the memory, it cannot be changed. However, the address can be accessed repeatedly to recover a copy of the stored data.

Vector– A coded multidimensional signal incorporating multiple values derived from multiple sensory or cognitive sources (possibly over a time interval).

19.8.2 Overview scenario of the reading process

[rewrite xxx]
Several sections addressed previously are important to an initial understanding of how the brain manipulates information related to these subjects. **Section 15.2.2.7** xxx addresses the role of the pulvinar in creating percepts of a scene. These percepts depend on the ability of the two-dimensional correlator within the pretectum to generate the initial interps of the elements of the scene. The correlator design and organization are discussed in **Sections 15.3** & **15.4**. A number of consequences associated with improper operation of the POS, the control point, the pulvinar and the cerebrum are highlighted in **Section 18.4, 18.6** & **18.9**.

Reading should be differentiated from the observation of logos, and other large graphics that may contain letters drawn from a local language. Logos are meant to recall a previous image. They are not designed to convey new information.

Reading is a highly stylized method of acquiring information from a written message. It invariably involves a step and repeat process wherein the eyes move progressively along a line of symbols, pausing for 50-200 msec. before proceeding via a small saccades. During the pause, the eyes perform a complex pattern of microsaccades (only poorly categorized in the literature under the heading of tremor). These individual motions are used to convert the spatial pattern of the message into a temporal pattern that can be manipulated by the neural portion of the visual system. The control of these mini and microsaccades emanates from the Precision Optical System. The POS is a group of high performance servomechanisms that control the motions of the eyes in response to the requirements of the portion of the midbrain tasked with extracting intellectual meaning from the symbology projected onto the foveola of the human eye. This portion is called the pretectum and it is a part of the thalamus. The pretectum attempts to extract the fundamental meaning of small groups of interconnected lines we generally call syllables. When the pretectum is able to do this successfully, the POS follows a preprogrammed sequence of step and repeat motions. These motions present a new image to the pretectum every 50-200 msec as required. The pretectum extracts the meaning of each small group of lines in the form of an initial (or basic) interp. These individual interps are assembled sequentially until they represent a higher level interp, typically a word. Multiple interps are then assembled into a percept, typically a short sentence or a phrase. The process continues to build until the subject comprehends the entire meaning of the written message.

The gross steps in reading are accomplished serially, but in a continuous pipeline fashion, and can be defined as;

- sensing, the acquisition of material imaged on the foveola,
- Initial cognition (based largely on prior experience) resulting in an interp,
- Second level cognition (based largely on prior experience) resulting in a percept,
- Third level cognition (based to a significant extent on prior experience) resulting in a theme,
- Higher cognition leading to a comprehension of a message or idea.

Much of the material comprehended is stored in “permanent” memory.

These terms will be defined in detail below and in the Glossary.

19.8.2.1 State of knowledge concerning the circuits of reading

The current state of knowledge concerning the process of reading is spotty. The architecture of the process can be reasonably understood through comparison to man-made circuits of similar function and through reverse engineering (typically involving diagnosis resulting from trauma, disease or intentional lesions). The signaling capabilities of the individual neurons are also well understood. However, the coding used to transmit the messages, and the nature of the messages is totally unknown at this time. As illustrated below, our knowledge of the more gross movements of the eyes in reading are well documented. Our knowledge of the more critical microsaccades of the eyes has not progressed since Yarbus and Ditchburn in the 1970's.

While a large amount of dissection, at the neuron anatomy level, has provided a large amount of detailed information about specific signal paths, the overall signaling patterns have been difficult to define on a global basis. This is changing rapidly with the introduction of nuclear magnetic and nuclear emission imaging techniques, particularly the functional MRI (fMRI) which is fast enough to represent individual areas of gross neural activities simultaneously over a significant area of the neural system. Unfortunately, these techniques are still classed as a tool in traffic analysis from a communications perspective. They do not describe the precise commissure path between engines, the number of parallel neuron circuits in the path, or anything about the messages or the coding of the messages along these paths. To answer these questions will require an increase in the sophistication of recent electrophysiological techniques. Such techniques are available in the available test equipment base. They involve synchronously recording multiple neural paths, each at finite time intervals during variations in the stimulus applied to the subject. Such techniques are readily applicable to the simple neural commands of the oculomotor systems now. By advancing back along the neural circuits of the oculomotor sections of the POS, additional understanding of the codes and messages can be obtained. Hopefully, this will eventually lead to the interception and decoding of neural signals along selected simple paths that relate to various levels of comprehension.

19.8.2.2 Setup procedures prior to reading

While it may seem obvious, the material to be read must be placed to present imagery within a specific size range

(if it is to be read as opposed to just being recognized as a graphic). The smallest character to be analyzed must normally represent an angular size of at least 40 seconds of arc. If the size exceeds five times that height, the full capability of the analytical process to read efficiently will not be realized. The material should be lighted to within the photopic range. Ideally, it should be illuminated with a daylight source (nominally 7053 Kelvin) to insure maximum performance by all of the photoreceptors in the foveola. The material should be of the highest contrast, black on white being preferred.

If serious reading is to be performed, the subject should prepare to read by isolating himself from other distractions that might instigate needless activity in the alarm mode of vision. Such alarm signals will override and interrupt the reading process.

19.8.2.2.1 Achieving focus

The auto-focus capability of the eye is dependent on the POS. The POS in turn is dependent on detail in the image to achieve focus. This detail should consist of at least one high contrast edge in the field of the foveola. As Campbell & Greene noted, a sine wave does not provide a good signal for auto-focus purposes. The focusing operation typically requires 50-100 ms. It will become clear that the visual system does not refocus following each saccade. It relies upon short-term memory of the scene to provide ques to proper focus, only stopping to refocus following abrupt movements of the scene relative to the eyes.

The capability of the memory system is largely unexplored relative to reading. It appears that the visual system relies heavily on memory to provide an estimate of the focus condition required to return to any line of fixation previously achieved. Thus, changing from reading to looking across the room appears to rely on memory for an initial estimate of the focus condition required to focus on the far wall.

19.8.2.3 Typical sequence of minisaccades during reading

The following material is reproduced here from a fuller account of reading dynamics presented in **Section 7.5.3**.

Becker, et. al. have also provided empirical material on the reading mechanism as of 1999[5]. Both Underwood[6] and Becker, et. al. concentrate on observable eye movements (at the minisaccade level) as the key to our understanding of reading. This work takes an entirely different view. It treats observable eye movement as merely a mechanism for imaging individual scene features, symbols and character groups onto the foveola where the actual process of perception is initiated. This perception involves eye movement at a level not normally observed by the clinician or academician, the microsaccades or tremor level.

> Based on the conceptualization developed in this work, **reading can be defined as the act of assembling a sequence of perceptions acquired through the sequential analysis of individual symbols or character groups and interpreting these perceptions in accordance with a set of syntactical rules.** In this definition, symbols include hieroglyphics and other glyphs. The initial interpretation of each symbol or symbol group by the POS results in the generation of an individual "interp." When a series of interps are combined, the resulting interpretation will be called a percept.

Within this work, the foveola and the fovea are defined based on the morphological characteristics of the retina. The foveola is defined as 1.18° in diameter, the fovea is defined as 6.26° in diameter in object space and the parafovea is defined as beyond 6.26° from the fixation point. Therefore the fovea of Kennedy conforms closely to

---

[5] Becker,W. Deubel, H. & Mergner, T. (1999) Current Oculomotor Research. NY: Plenum

[6] Underwood, G. (1998) Eye Guidance in Reading and Scene Perception. NY: Elsevier

the foveola of this work and his parafovea will be assumed to equal the fovea of this work.

Kennedy also states “In reading, each word is inspected by an initial fixation at a particular position resulting from an ‘entry saccade’ of a given size, launched from a particular location in another word.” The conditions, variations and significance of this entry saccade are discussed in some detail on page 152-153. The relationship of such saccades on the empirical model of Rayner, Reichle and Pollatsek in Chapter 11 is also reviewed. The referral to a “location in another word” is about as close to the discussion of characters in a word that is achieved in the overall Underwood text. More material is presented on regressive saccades, from one word back to a previously examined word, than is given to the examination of the characters within a word.

Kennedy discusses the term prompt, with the word gaze in parenthesis following it, as the sum of all fixations prior to the first excursion outside of its boundary (page 157). In the Abstract to the chapter, he says the time to process one of two possible foveal “prompt” words was examined using measure of gaze and fixation duration.

Figure 19.2.1-1 presents a modified, semi-standard figure from Liversedge, Paterson & Pickering in Underwood. It should not be inferred that most of each time interval shown relates to the latency before the next saccade. These saccades may be a part of a planned saccade sequence that does not require a significant latency between the end of analysis within one gaze and the beginning of a saccade to the next gaze location. Thus, a more specific set of subdivisions of the term latency is probably called for here. The times do define the maximum length of time required for the visual system to analyze the structure of the symbols within the foveola adequately for the subject to ascertain their semantic content (probably via a lookup table). The above authors did not address the size of the type (subtense of the height of the characters) used in their experiments but they did say the tracking data was quantized every millisecond.

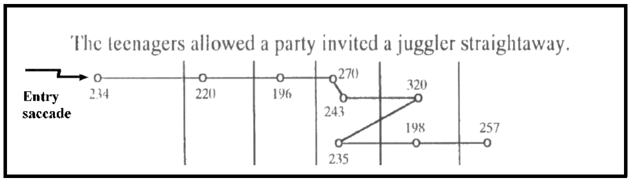

Figure 19.2.1-1 Hypothetical eye movement record showing the time in milliseconds spent in a gaze between saccades. The text being read is at the top of the page. The zig in the line is indicative of a regressive saccade. Total reading time is over two seconds for this single incongruous sentence. The process is highligted by the regressive scan to analyze further the use of the word invited.

Inhoff & Radach, writing in Chapter 2 of Underwood, reported on eye movements when viewing long strings of printed characters. Their data provides good information on the precise nature of the small saccades related to eye movement during reading but little information at the microsaccade level or on the nature of the saccades used to perceive word meaning in the context of reading. Their text introduces a number of potential experimental variables (flexibility of the eyeball leading to transient movements of the lens group–cornea and/or lens) but does not provide a foundation for overcoming or controlling them. Their figure 1shows a number of minisaccades at the 0.1-0.2 character width level in between saccades of one and seven character widths. However, the noise level of their equipment was not specified or shown and they may have excluded microsaccades from their analyses. Their discussion includes presentation of an interesting dichotomy (citing Deubel & Bridgman[7]). First, that the eye is an imager and that small “post saccadic movements will smear the retinal image.” Second, that the post saccadic motions are relatively small and ***principled,*** and the reader may be able to extract useful information during that movement

---

[7] Deubel, H. & Bridgeman, B. (1995) Perceptual consequences of ocular lens overshoot during saccadic eye movements. *Vision Res*. vol. 35, 529-538 & 2897-902.

period [emphasis added].

Radach & McConkie prepared Chapter 4 in Underwood. It discusses the determination of fixation positions in words during reading. One of their conclusions is that, "In all cases (where there are spaces between the character groups), eye movement control during reading appears to be word-based." This control appears to involve two distinct mechanisms, a selection mechanism and a performance mechanism. It is proposed that the selection mechanism determines how one word rather than another is selected as the target of a saccade, and the performance mechanism determines where the eyes actually land given the above selection. The discussion centers on the general likelihood that the saccade is aimed at the center of the selected word. In the data presented, the imagery was presented (a German translation of the initial text in *Gulliver's Travels*) in page format with five to seven double-spaced lines of up to 72 ASCII characters each on a 15 inch VGA monitor. At 80 cm viewing distance, each letter corresponded to approximately 0.25° of visual angle (15 minutes of arc or three times the 20/20 character definition).

One of the conclusion of the above authors is that "where the eyes go with respect to selected saccade target words, is the result of low-level visuo-oculomotor control factors, almost completely unaffected by higher cognitive processes." This is probably true for relatively familiar words not calling for regressive saccades or multiple saccades within one word. The above authors go on to caveat their statement. One of their caveats is "in the case of regressive inter-word saccades, the saccade parameters we have looked at suggest a control mode different from the low-level default routines." The intra-word characteristics of the reading process were not discussed in their chapter. Here again, the word inter-word appears in the index to Underwood but the expression intra-word does not. Heller & Radach made note of two important facts[8]. They noted the work of Dunn-Rankin in 1978 that showed that the initial fixation point on words was not at their center but at positions left of center. They also noted the work of Rayner & Pollatsek in 1981 that showed that the final decision concerning the direction and magnitude of the next saccade was made during a given fixation interval.

Rayner, Reichle & Pollatsek presented Chapter 11 in Underwood. They discuss the effect of limiting the length of time available within a gaze to analyze the text characters. They show that if a given gaze (the traditional fixation interval by name, even if it involves tremor) is interrupted before 50-70 ms have elapsed, the reading process itself is interrupted and comprehension suffers or is lost. They also suggest that a preview of a word, while it is imaged in the area outside the foveola, has a positive impact on reducing the time required to interpret it when it is moved into the foveola.

The above authors review several conceptual models of the eye movement control system required to implement reading. A process model by Morrison is summarized. This model is still conceptual but includes the concept of some preanalysis of a word before it enters the foveola or before it is brought to the point of fixation within the foveola. A logic is provided that controls the length of the gaze and/or the series of interim fixations associated with each word. The model explains two aspects of the eye movement phenomenon in the reading process: (1) the fact that there are fixations that are much shorter than the nominal 195-200 ms saccade latency in simple oculomotor tasks and (2) the occurrence of unusual landing positions, such as between words. A competing model by O'Regan, that they describe as a strategy-tactics model, is also summarized along with a critic of its features.

Finally, the authors summarize their proposed E-Z Reader Model. They describe it as similar to Morrison's process model except more refined through the implementation of two additional facets. First, it decouples the signal to shift attention from the signal to program a saccade. Second, it is better specified in that it has been implemented as a computer simulation program. They refer to the fifth generation of this program as the program under current discussion. They describe the simulation with a schematic representing five basic processes:

---

[8] Becker, et. al. (1999) Op Cit. pg. 341

1. A familiarity check on a word.
2. Completion of lexical access.
3. An early, labile stage of saccadic programming, which can be cancelled by subsequent saccadic programming.
4. A later, non-labile, stage of saccadic programming.
5. The actual saccadic eye movement.

They define the first two steps as products of a single cognitive process which occurs during a preview while the word is still in the fovea (not the foveola) and occurs before the movement of the word to the center of the foveola. They develop the fact that completion of the familiarity check depends on two additional factors. These factors slow the rate of processing as the eccentricity of the word, relative to the point of fixation becomes larger. This factor was added to recognize the rapid falloff of the resolution (acuity?) of the eye with eccentricity.

They also discuss the operational distinction between an interword and an intraword saccade. They suggest the basic eye movement strategy conforms to the “dumb” default strategy: That strategy is to plan to fixate each word from more than one viewing location unless the word’s familiarity indicates that a refixation is unnecessary. The dual fixation strategy could obviously be useful in words that frequently have unusual or multiple suffixes (syllables).

The dumb default strategy emphasizes the importance of having a large vocabulary in the field encompassing the text being read. It also suggests that the average reading rate is dependent on the frequency of occurrence of words found in the vocabulary.

Some authors have argued there is no “magic moment” of word identification, that identification (of individual words or total thought) only comes with a growing amount of data collected on a continuous basis. This appears to be a question of semantics between authors since the motion of the eyes is clearly not continuous with time.

19.8.2.3.1 The familiarity default procedure in reading

The above “dumb default” strategy suggests a major change in the operating mode of the POS during reading. As soon as a word is recognized cognitively (the magic moment?), the POS initiates a saccade and proceeds to the next fixation point. This suggests that the visual system operates in a manner similar to the “auto complete” feature of an INTERNET browser. The browser compares the initial key strokes of an entry with its short term memory and suggests the appropriate completion of the typed entry. If the suggestion is wrong, the typist is free to enter an alternative. An equivalent scenario can be defined for the visual system. After analyzing only a few symbols, initially in the fovea area surrounding the foveola, the system may believe with high probability that it knows what the entire symbol group means. In that case, it will instruct the POS to proceed to the next symbol group. If the following symbol group is recognized but it does not fit logically into a recognized syntax with the first group, the POS may be instructed to perform a regression saccade in order to review the previous symbol group for an alternate suffix or other difference from the assumed meaning. This procedure is illustrated in **Figure 19.2.1-2**. The system attempts to interpret a short sentence. It can methodically perceive and interpret each character group as in A or it can adopt a more aggressive approach and make a guess based on the likelyhood that the character group “dres” is part of the longer word “dressed.” This results in the sequence shown in B and some time is saved as one gaze (fixation) is eliminated from the initial saccades sequence. If however, the assumption was made that “dres” was part of the word “dresses,” the same procedure can be followed until the second or third character group in “yesterday” is reached. At this point in the interpretation, a context conflict is recognized through comparison of the initial concept file with the saliency map of the individual. As a result, a regression saccade is called for back to the word that was actually “dressed” and not “dresses.”

It is instructive to consider the time line of the above activity to determine if it suggests the course of the signals moving through the visual system. Certain time lines may suggest activity limited to the POS, to the POS in conjunction with the pulvinar or to the POS in conjunction with either area 7 or the posterior areas of the cerebrum. **Figure 19.2.1-3**, from Becker, et. al. addresses this subject directly[9]. There is a statistically shorter latency associated with the regressive saccade than with the progressive scan. Apparently, as soon as the system notes an inconsistency in the proposed syntax, it cancels the analysis of that character group and calls for a regressive saccade to reestablish a viable baseline. These experiments were carried out with considerable statistical precision and the original source should be reviewed before proceeding.

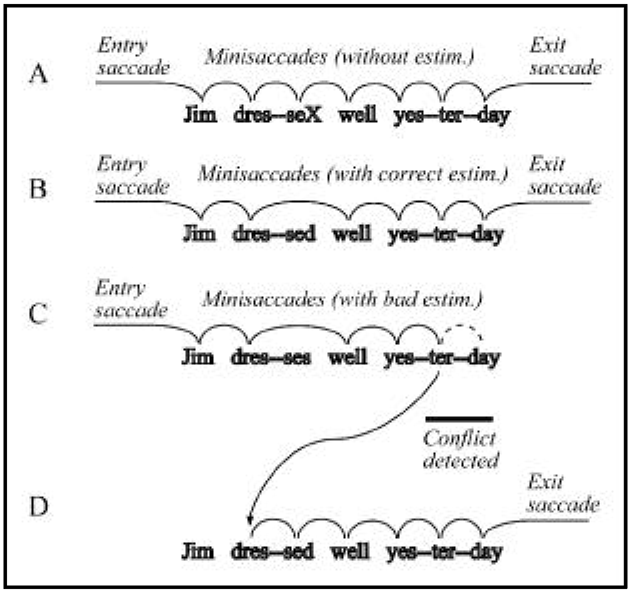

Figure 19.2.1-2 The procedure of perceiving and interpreting a sentence showing three alternatives.

---

[9] Becker, W. Deubel, H. & Mergner, T. (1999) Current Oculomotor Research. NY: Plenum Publishers, pg 320-325.

14 Processes in Animal Vision

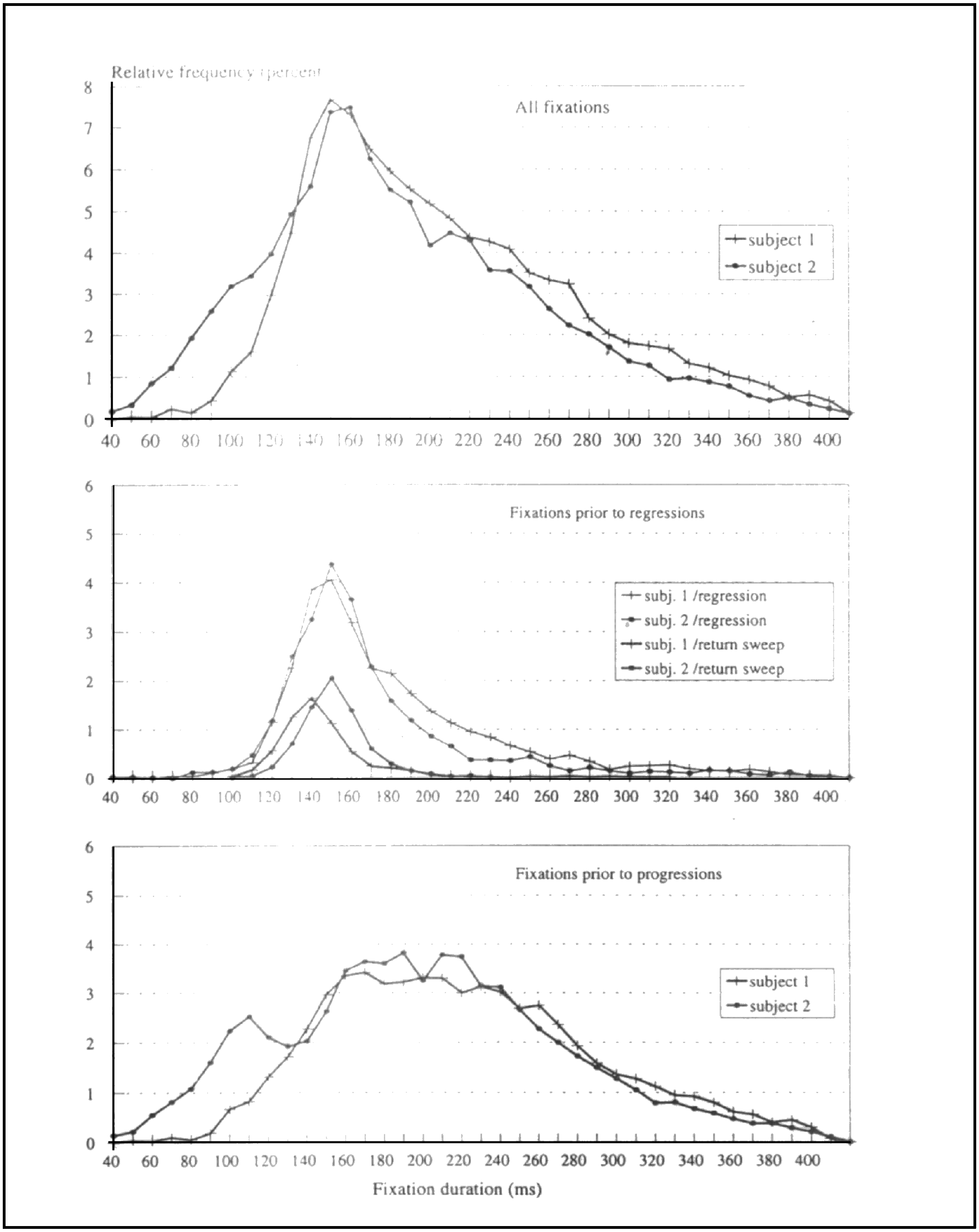

Figure 19.2.1-3 Distribution of fixation duration for two subjects in a standard reading experiment. Top, cumulative distribution. Middle, fixation duration following syntax failure at fixation after a "dumb default." Bottom, distribution of fixations before normal progression to next fixation. See text.

19.8.2.4 The fundamental phenomenon and mechanisms of reading

The above synopsis of much detailed work reported in Underwood and elsewhere leads directly to a discussion of how the semantic meaning of a character group within a word is determined during the individual fixations (or gazes).

It is likely that the reader initially scans textual material as a scene and evaluates its general characteristics, concentrating on paragraph arrangement, figure inserts, indentation etc. The system then calls for an entry saccade to position the line of fixation on the first character group of interest.

[**Figure 19.8.1-4**] demonstrated the importance of considering the size of the text imaged onto the foveola and fovea. The preliminary conclusion was drawn that less than five (nominally three) characters of nine point type were usually imaged on the foveola at one time and this number quantified the amount of information that could be analyzed by the visual system without additional saccadic activity. It also showed that more characters could be imaged across the fovea. However, the resolution of the optics and the different signal processing employed in the foveal region suggests that these characters could not be analyzed with sufficient precision. However, the structural accouterments of the character groups (capitalization, illumination–including capitalized acronyms, hyphens, etc.), and particularly the space between them could be determined using the awareness channels of vision.

The data on progressive intraword and regressive scanning suggests the system is prepared to make estimates of the meaning of an initial character group based on only the first few characters and then repeat the estimation procedure on subsequent character groups until the cumulative estimate fails on the basis of syntax or the analysis and perception of the word is completed successfully. The failure calls for a regressive scan to repeat the interpretation process. This regression may involve several character groups or several words. Success calls for the implementation of a progressive scan to the next word. The empirically measured location of adjacent gazes also suggests that the system is prepared to ignore many if not all two and three letter character groups when they are recognized in the fovea and before they are fixated upon. There is an exception to this policy where the system is expecting such groups such as in highly technical text.

19.8.2.4.1 The mechanism of detailed character group analysis

As a result of the above procedural steps, the critical remaining task is to specify how the visual system recognizes the individual character groups during an individual gaze. **Figure 19.2.1-4** describes the scene imaged on the foveola of the retina for a word imaged at the scale of the 20/20 line on a Snellen Eye Chart. The letters are 3.125 minutes of arc high to reflect the immersed optical system used in vision. This size corresponds to 5 minute of arc characters in object space divided by the index of refraction of the vitreous humor of the eye. The characters are shown in a bold san serif font resulting from the use of a serif font in image space, such as Times New Roman, and passing the image through a resolution limiting optical system such as that of the human eye.

> There appears to be little appreciation in the Ophthalmology community for the difference in performance of the visual system when viewing characters with and without serifs. This is demonstrated by the back to back samples shown without relevant discussion in Miller &

Newman[10]. The discussion of the minimum angle of resolution (MAR) parameter is useful, as is the history of the Snellen Eye Chart.

Each photoreceptor of the retina is shown at its correct nominal size of 2 microns diameter. However, it is shown in a square packed array for ease of discussion when they are actually packed in a hexagonal array. At this resolution, a five equally spaced character group representing a word spans about 43% of the foveola. As discussed earlier, this size image corresponds to about 1.5 point type at 15 inches and most people cannot resolve the individual characters at this scale and the spacing shown. Only when the characters are well separated and in block form can they be perceived under these conditions. The fact that they can be perceived is suggested by the fact that the microsaccades of tremor are sufficiently large as to cause the photoreceptors of the retina to scan across the edges of the individual characters. The degree of scanning is sufficient to determine the contrast edges associated with each character, even with the hexagonal packing of the actual photoreceptors. However, the degree of scanning is or is becoming marginal. As a result, the question arises as to whether the scanning is adequate to generate an outline of the character as in the letter P, or whether only a stick character version, such as the R, is available for perception by the brain. This question becomes moot as the size of the characters is increased. Using nine point type at 15 inches, the individual strokes of a character are about the width of twelve photoreceptors in a line. It is clear that the combination of the photoreceptor size and the scanning mechanism can resolve both edges of a stroke of a character this large. The hexagonal packing of the photoreceptors can only provide additional detail.

Because the complete initial character group falls within the foveola during a single gaze, the parallel processing capability of the neural system connecting the foveola to the brain is able to analyze all of the characters simultaneously (assuming adequate resolution of the space between characters). Hence, the brain is able to assemble an estimate of the three individual characters and to reach a preliminary interpretation of the meaning of this syllable-like group.

Note however, the eyes are not stationary during this nominal 220 ms gaze. They are continually moving with amplitudes of micro-arc-seconds and frequency components in the 100 Hz region. Based on the findings of Rayner, et. al., this gaze interval will be subdivided into individual “glimpse” intervals of about 50 ms. The assumption being that the pretectum is performing its cross correlation calculation based on the edge crossings that occur during the glimpse interval. It then issues an interp based on this calculation (after performing a correlation with the interp values stored previously in the pulvinar).

Having made an estimate of the meaning of the first character group, it has two choices. It can make an assumption as to the meaning of the entire group of characters in the word, and execute a “dumb default” to the next word or it can make a progressive saccade to the next character group within the overall word. In the latter case, it proceeds to build an estimate of the meaning of the entire word essentially one syllable at a time. In the former case, it takes a gamble that is proven good or bad based on the syntax of the previous and following words in the line of text. If the estimate appears to fail, a regressive scan is performed to reestablish a running estimate of the meaning of the multi-word sentence. The more technical the text material, the more likely the estimate is to fail and the more frequently regressive scans will be observed in the experiments.

What is not known at this time is the scanning pattern used by the POS to scan each multiple character group. It appears clear from Shakhnovich and Yarbus that the horizontal and vertical components of each microsaccade can be implemented individually within times of less than 10 (probably less than 5) milliseconds. Experiments are needed to determine the statistical relationship between these two components for a variety of simple symbols and symbol groups as well as conventional alphanumeric character groups.
microsaccades amplitude

---

[10] Miller, N. & Newman, N. (1998) Walsh & Hoyt’s Clinical Neuro-Ophthalmology, 5th ed. Baltimore, MD: Williams & Wilkins, pp 161-162

.

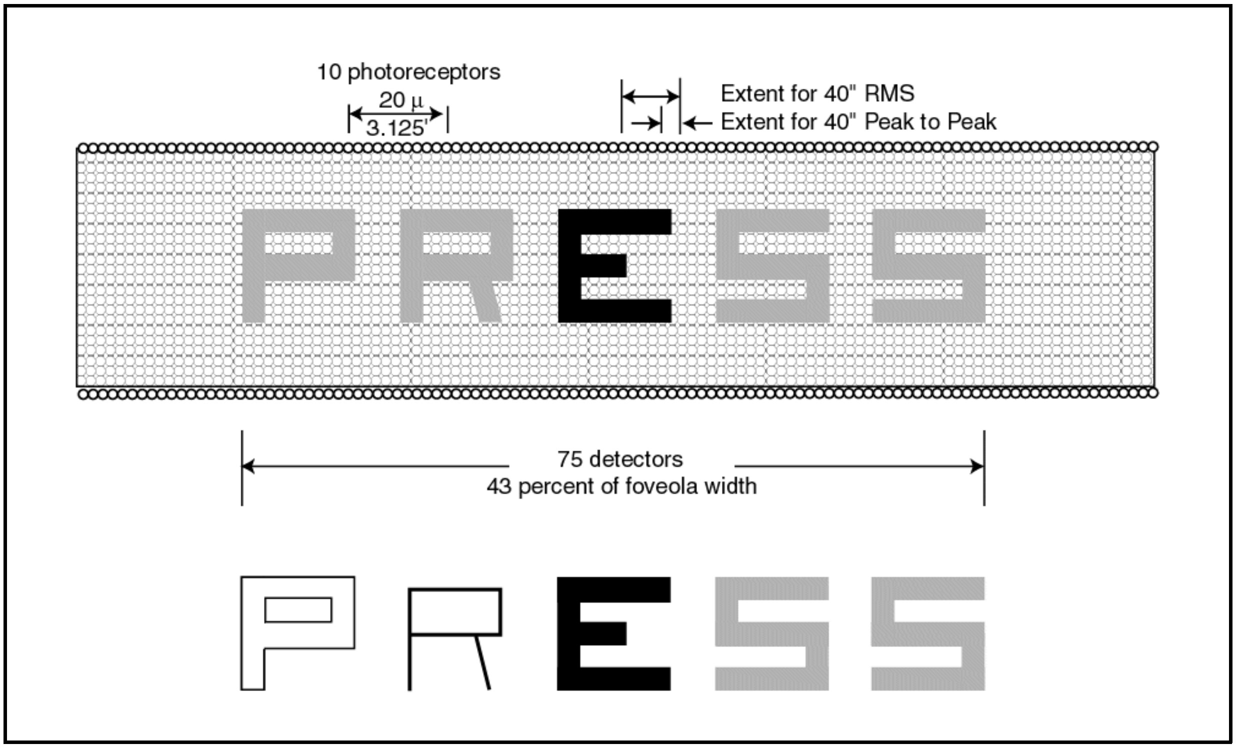

Figure 19.2.1-4 The Stellen E placed on the retina at 20/20. Top frame shows the word PRESS overlayed on the retinal mosaic. The arrows on the upper left show the size of 10 adjacent photoreceptors. The arrows on the upper right show the nominal size of the tremor (microsaccades). The bottom line shows the extent of the tremor if its amplitude is 40 arc seconds peak to peak. The top line shows its extent if its amplitude is 40 arc seconds RMS. Bottom frame presents the question of how a single character is parsed for transmission to the cortex.

19.8.2.4.2 Typical sequences of microsaccades

Only gross information is available at this time concerning the nature of the microsaccades associated with reading. **Section 7.5.3** summarizes what is known. The greatest lack at the current time is the phase relationship between the microsaccades associated with the motion of the eyes in elevation and angle. Currently available instrumentation can provide this information. What is needed is a bio-engineering institution, associated with a strong electronics instrumentation organization, that can look into this subject. Merely producing the Lissajous figures associated with these movements as a function of time (in about 10 millisecond steps) would produce a goldmine of information about vision.

19.8.2.4.3 Suggestions for future research in reading

While excellent insight into the reading process has been achieved during the previous exploratory research, it is appropriate to consider ways to raise the precision of the empirical process. One method of improving the statistical relevance of the data would be to remove the variable of word length from the sample text material in a given experiment. This change would allow more precise conclusions to be drawn with regard to the effect of word length on the statistical properties of the observed saccades. It would also allow more precise determination of the nature of the interword saccades. If the words were randomized, the element of syntax could be removed from the perceptual process, although the element of vocabulary would remain. Ultimately, these experiments may lead to a better understanding of the role of short particles of speech in the interpretation process. When combined with the current data base, they may also lead to a better understanding of the process as a whole.

19.8.2.5 The Strategy Employed in Reading

The strategy employed in reading does not differ significantly from that for general scenes. However, to aid in communications, additional sets of rules have been adopted related to spelling and syntax that are not required for general scene interpretation.

When presented a scene containing text, the observer initially examines the complete scene as discussed above. It notes areas of the scene (objects) that appear to contain a particularly structured texture. If this object appears upon further examination to be text, the visual system performs an "entry saccade" that brings the line of sight to the expected entry point of the text (the upper left hand corner for most western languages).

After locating the first word (and character group of that word) the analytical procedure must initially determine the language of the text and the font of the character set used. It then continues based on the set of style, syntax and spelling rules for that language.

Reading involves the interative process of creating an initial concept file just as when examining a natural scene. However, there are two significant variations. The stylistic rules dictate the direction in which the text is to be scanned. The anticipated rules concerning the shape of individual objects ares replaced by a set of syntax rules unique to the specific language.

The syntactic rules include a stop symbol (the period) that defines the end of the presentation of a a particular concept. By this means, a sentence is formed that is similar to an object within a general scene. The concepts associated with multiple sentences can be grouped into a broader concept. This grouping is lableld a paragraph in pedagogy.

19.8.2.5.1 Examples of the reading process

The following simple examples illustrate the process of perception and interpretation.

CHECKING THE SPELLING ON THE FLY

The following example shows how making a determination of the precise spelling in the first character group of the third word, and checking for that form against the saliency map, changes the concept contained in the initial file created during the reading process.

The boy performs well on cue.
The boy preforms the kit parts.

MODIFYING THE CONCEPT

Each of the three letter, and one two letter, groups in the first line is the target of an individual minisaccade. The character-group as a whole is then scanned by a series of horizontal and vertical microsaccades to determine its meaning. In the second line, one word is too long to be perceived and interpreted completely (at the geometric scale of the text). It is treated as a two-group (syllable)word. The first minisaccade brings the character-group "aro" onto the foveola and the next minisaccade brings the term "und" onto the foveola.

The boy ran to the box

The boy ran around the box

As a result of the saccade sequence, and the continual checking of the spelling of each character-group, these two samples are interpreted slightly differently and result in distinctly different concepts. No conflict has arisen in the perception and interpretation of these simple sentences.

SIMPLE CONCEPTUAL CONFLICT

This example is obvious. Following the second saccade following the entry saccade, the initial file contains a conflict. Boxes do not run.

The boy ran to the box
The box ran to the boy

SHARPENING THE CONCEPT WITH SUFFIXES

This example merely illustrates how additional precision can be added to the concept using a suffix or an adverb.

The boyish girl ran to the box
The boy ran obliquely to the box

MORE COMPLEX CONCEPTUAL CONFLICT

In both of the above examples, the additional character-group within the two longer words would have caused an additional saccade to be added to the initial saccade sequence list. This would slow the reading process. However, the system has adopted a methodology to avoid this problem a high percentage of the time. After interpreting the first character group in boy or in obliquely, it frequently makes an estimate of the impact of the second character group of the word and calls for an immediate saccade to the next word. In the following example, the error is of negligible importance.

Jim dresses well
Jim dressed well

However, proceeding in this manner can result in trouble. In the following example a conceptual conflict arises in the first sentence several saccades later.

Jim dresses well yesterday
Jim dressed well yesterday

When such a conflict is encountered, a *regression saccade* is called for. This saccade takes the image projected on the foveola back to a point where resolution of the conflict is possible. The word dressed must be distinguished from the word dresses by using two saccades.

The frequency of multiple case specific suffixes contributes to the difficulty of a novice trying to read Russian.

7.5.4 Extension of the model of the visual servosystem

The above review of the operation of the visual system during analysis of a bucolic scene, a scene containing local changes and of reading suggests further sophistication of the visual servomechanism subsystem. The subsystem is so complex as to require a state diagram as an aid to understanding the physical structure of the system. The appropriate state diagrams and more detailed circuit diagrams will be presented in **Section 15.2.5**. What is clear from Sections 7.3 through 7.5 is that the servomechanisms of vision, centered on the POS, are highly sophisticated. The overall system involves multiple loops. The outer loop includes the midbrain and the cortex, with initial entry at area 19. The inner loop may include the cortex, particularly area 7, or it may be embodied within the midbrain and the xxx. It is most likely that the outer loop passes through the LGN on the way to area 19 while the inner loop passes through the Pretectum. The system is necessarily a sampled data system since all of the signals from the retina are converted to pulse signals in the signal projection stage between the retina and the midbrain as well as for transmission over the association fibers of the brain.

As shown above, the servomechanisms operate in a variety of modes during vision. Switching between these modes appears to be under the direct control of the Pretectum. However, it is safe to assume that the Pretectum frequently looks to the cortex for advice of a cognitive nature and it is always alert for both stereo-optic signals and alarm signals from the computational capabilities of the LGN.

Besides using sampled data servomechanism techniques, the visual system is able to incorporate a large amount of computational capability within each servo loop. This makes the application of simple analog servo techniques to the understanding of the visual system a questionable procedure. Only under the most circumscribed experimental conditions is it appropriate to use analog concepts in interpreting the visual process. The servo loops of vision incorporate significant lumped constant time delays due to the nature of the signal projection function. While analog concepts have been used in the past to estimate the high frequency capability of the visual system, they have been grossly in error. Even the low frequency region of the overall signal passband cannot be described as a low pass network. Since the system exhibits a zero at zero frequency, even the characteristic of the servo passbands between zero and five Hertz must be described as a bandpass network. In parallel with this bandpass network is the separate high frequency portion of the overall characteristic associated with tremor. This region extends from somewhere above five Hertz, probably above 30 Hertz into the region above 120 Hertz.

Major limitation of the inner servo loop of the visual system are the finite lumped constant delays related to the projection stage. These are located between the retina and the Pretectum, and between the Pretectum and the oculomotor muscles. These two delays sum to a nominal 6.6 ms. With even minimal time delay in the signal processing portion of the loop, it is effectively limited to a 160 Hz maximum frequency response. In terms of these delays, the response of the servo plant, consisting of the oculomotor muscles and the inertia of the eyes, in traveling through a rotation of only 20-40 seconds of arc is minimal.

7.5.5 Extension of the concepts leading to cognition

There is a lack of defined terms describing the process leading to cognition of symbolic material at the level required to understand the mechanisms involved. A specific problem is the lack of a term to describe the signal produced by the pretectum following the cross correlation of the image presented to the foveola in the process of reading. It appears that the pretectum is only able to handle syllables or short words in a single time frame of about 50-70 ms. The output produced by the pretectum is obviously not a complete thought, nor is it likely to be expressed as a complete thought. In this work, the term “interp” will be used as a noun to describe the vector created by the pretectum and the POS in response to stimulation of the foveola by a syllable, group of syllables or a word during a single 50-70 ms of a gaze interval. The pretectum presents to the higher cognitive centers this interp (in vector form) that can be incorporated into the saliency map. Alternately, it could present the vector to a short term saliency map used to collect time sequential vectors (interps) that are collected during a single gaze of approximately 220 ms. These interps would then be assembled into a complete perception of a message (percept) in vector form. The term percept is used in psychology to define the complete message. These individual interps will be accumulated within the POS or within Area 7 of the cerebral cortex. Multiple interps, when accumulated sequentially are presented to the higher cognitive centers as a percept. One or more percepts (possibly from multiple sensory systems) are then analyzed. This analysis leads to the cognition of a specific thought concerning a stimulus input.

> The terms interp and percept are defined here out of expediency. No claim is made as to originality or precedent. They may be defined in a more sophisticated context in psychology.

The definitions used here will eventually be correlated with those other terms.

19.8.2.6 Reading alone is achromatic

As developed in detail in **Chapters 11** and **15**, the signals of the analytical path connecting the foveola to the midbrain are not coded for chromatic information. All of the neural paths reaching the pretectum are treated equally, regardless of spectral representation, in the process of correlation discussed in **Section 15.2.5**. This leads to maximum spatial resolution in the overall process. It also leaves the outcome of the reading process achromatic. This is generally unimportant. If an idea of the color present at a relatively coarse level is desired, it is available from the same signals processed via the awareness mode of vision. It may be that such a coarse indication of color and color contrast is routinely attached to the signals passing through the circuitry of the analytical mode before they are stored in the Saliency Map. If so, the indication would be represented by another value attached to the vectors created during analysis.

While the reading process is achromatic in structure, presenting material in other than black and white causes a loss in signal-to-noise ratio within the data stream.. This condition can result in a significant loss in legibility.

19.8.3 The mechanics and mechanisms of reading

As mentioned above, reading is the ultimate mechanism and capability provided by the analytical mode of visual system operation. In its full implementation, it is only currently found in *homo sapien*. This appears to be primarily due to the higher level of development of the Precision Optical System associated with the midbrain. It is this subsystem that implements most of the uniquely human capabilities of the human visual system. It is this subsystem that goes beyond the simple pattern recognition capabilities shared with all animals, and allows humans to interpret patterns of fine detail that other animals merely observe as "pretty" at best. This position is in agreement with Chomsky and Lenneberg who have argued that human language must be recognized as a species-specific ability based on unique bio-morphological structures[11].

The unique capability of reading is centered on the midbrain, and particularly the thalamus. This structure involves a large number of unique engines that perform a wide range of roles in vision. The roles of interest in reading center on the pretectum and a special engine labeled the control point in this work. These two engines form nodes that are shared with a variety of distinct servomechanisms and cognnitive elements. These roles will be summarized below.

Traffic analysis, reported in Carpenter & Sutin show significant two-way connections between the cerebellum, the thalamus and the cerebrum. These connections form the foundation of the low level cognitive processes associated with reading xxx. It will also be shown that most of the activity along this trunk originates at the control point, proceeds to the xxx, returns to the control point and is then forwarded to the cerebrum, where additional high level cognition takes place.

19.8.3.1 Functional definitions needed to understand the reading process

Lacking clear terminology, this section will discuss two levels of cognition. Low level cognition is based largely on routine operations based heavily on previously learned relationships. It is characterized by the formation of individual and groups of interps, as well as the formation of individual and groups of percepts. An interp is the first-level output of the multi-dimensional correlator located within the pretectum. In reading, it corresponds to the internally coded representation of a syllable imaged on the foveola. A group of interps generally corresponds

---

[11] Berlin, B. & Kay, P. (1969) Basic Color Terms. Berkeley, CA: Univ. Calif. Press, pg 109

to a word. A percept is defined as the internal coded message corresponding to a series of words and constituting an idea. Such a group of words may be either a clause or a complete sentence.

Groups of percepts will be labeled ideas in this discussion. An idea is a complete thought generally equivalent to a complete sentence. For purposes of organization, the interps and percepts will be considered to be encoded in a low level "machine language" in the language of computers. They are assembled into groups of ideas that will be labeled themes. Ideas generally correlate with sentences while themes are broadly equivalent to paragraphs. Ideas and themes will be considered to be "compiled" into a higher level language than the machine language of interps and percepts. These are the languages of the comprehended material. Prior to the first stage of comprehension, the information from the retina is passed to the pretectum in a form that will be described as raw data. There may also be a separate language used by the controllers, specialized neural engines designed to control the collection and dissemination of the collected and comprehended material. For purposes of discussion, this language will be called the control language. It is distinct from the machine and control languages used to handle acquired information. It is generally sent over separate, supervisory, neural channels connecting the various engines of the brain. These terms can be collected into an initial **TABLE 19.8.3-1** for ease of reference and comparison. These terms may or may not be compatible with important documents in the literature. They are subject to change in the process of further definition and coordination.

TABLE 19.8.3-1
SUMMARY OF TERMINOLOGY DESCRIBING SIGNAL FORMATS

Term	Equivalent in Reading	Equivalent in a Graphic
Interp	Syllable	Line or Point
Group of Interps	Word	Feature
Percept	Complex Word or clause	Recognizable object
Group of Percepts	Clause or sentence	Group of objects
Idea	Sentence	Image on foveola
Theme (or Scene)	Paragraph	Image within fovea
Act	An initial comprehension	Environ. within inst. field of view
Play	A complete comprehension	Total environment

Most of the mechanism involved in the cognitive aspects of reading are illustrated in **Figure 19.2.3-1**. This figure will also be used to discuss individual elements of the system in following paragraphs.

19.8.3.2 Major functional steps in the reading process

Reading can be described as a highly stylized process involving a broad collection of neural engines found in the visual system, itself a major subsystem of the central nervous system. Its goal is to acquire information on any abstract subject via a highly stylized presentation of simple groupings of interconnected line elements. The stylistic aspects of the material include direction of sequential presentation of the groups, order of presentation of the groups, spacing of the groups, capitalizations, punctuations, and various methods of emphasis. On a different level, typesetters employ additional techniques to make the overall text pleasing to the eye. These include font design, variable spacing of individual characters within groups and various layout and pagination techniques.

The process of reading is dominated by a complex step and repeat process. The process is virtually automatic and the reader adopts a specific operating strategy that can be considered the default strategy. This default strategy is interrupted when the cognitive processes leading to comprehension are either interrupted or they do not achieve a sufficiently high performance threshold. When the interruption is due to a failure to recognize a simple grouping of line elements (generally representing a syllable), the system normally reverts to a learning mode. This mode involves a more detailed analysis of the group, in an attempt to learn its meaning, followed by an updating of the knowledge base of the individual with regard to that group. When an interruption is due to a failure to accept an idea or theme (generally because of failure of a syntax rule), a different strategy is used. First the control point orders the oculomotor system to return to a previous position and re-examine an interp that describes the tense of a word. In English, this interp appears at the end of a word (frequently a recognizable suffix different from the one assumed during the default reading procedure). In other languages, the interp may appear at other locations in a word or clause. Some languages use multiple suffixes. Any one of these may be the cause of the failure to accept a compiled idea or theme.

The performance level achieved in reading depends on a wide variety of disparate parameters. These range from the homeostasis of the individual to the light level of the material exhibiting the interconnected line element groups. They also include the familiarity of the subject with the language, idiom, semantics and lexical form used in the material. The performance may also be influenced by the quality of both the physiological and neurological optics of the subject.

The default reading process can be tailored by the individual (largely unknowingly) during his early education. Techniques such as omitting the last syllables in long English words can lead to greater reading speed, but with a potential loss in precise comprehension. Similar methods of overlooking short prepositions can lead to higher reading speed but with a loss in precise comprehension. Comprehension of a multi-syllable word normally involves about 200 msec. This suggests that perusing a line of text for less than a full second invariably results in a loss in precise comprehension. “Speed reading” becomes a matter of trading off speed in completing a task versus the level of comprehension achieved. This tradeoff is directly related to the inherent information density of the material. If the material is particularly elementary, many of the individual words can be overlooked by a well educated subject. Only “10 Dow” in a sentence would be adequate to identify the residence of the British Prime Minister. Alternately the “Rose House” might not be enough to identify the residence of the Argentine President on Playa de Mayo near the obelisk in downtown Buenos Aires.

19.8.3.3 Memory is crucial to the reading process

Memory is an absolutely crucial element in the reading process. It is used in a number of discrete locations and for a variety of purposes. In general, the memory associated with reading can be divided into short term and long term types. These types can then be subdivided further with respect to their functional role. While the precise location of some of these memory types is not known, this situation is changing rapidly. The use of the new non-invasive in-vivo imaging techniques are beginning to provide site maps that may be used to further identify memory locations.

Short term memory is critically important in the assembly of information derived from an image into a serial signaling stream that can be used in cognition and comprehension by the brain. Long term memory is used for reference purposes. It is used at a high level, where recent themes can be compared to themes acquired long ago. It is also used at a very low level, as a lookup table in the conversion of raw image data into interps and percepts.

The method of controlling the flow of data and information into and out of these different memories is not understood at this time. However, the capabilities required are well within the capabilities of the typical feature extraction and command generation engines of the brain. It may be appropriate to define an engine specialized for the purpose of controlling the deposition of data in and the access to that data in a memory, a controller–as it is known in computers.

19.8.3.3.1 Types of short term memory

The short term memory used in the visual process can frequently be described as similar to the short term, shift register type memory of computers. It is used to hold specific values and to sequentially assemble short values into longer values.

> It is common to speak of the content of a value in a shift register as a word. However, this will lead to great semantic difficulties here. In this discussion, a value will consist of a number of address slots filled by undefined characters that define a property of a feature. When multiple values are assembled into one larger value, the larger value will generally be considered a vector. Since the coding used to describe these characters, and to define properties to which they relate is still unknown, the discussion must be in a sense totally symbolic. The term word will be restricted to its use in human semantics.

The nature of short term memory in the visual system can be described with the aid of **Figure 19.2.3-1**. The elements along the diagonal line constitute the short term memory elements that can be described in terms of shift registers. They are typically filled during a time interval appropriate to their task (with some details below) and then refilled after the controller has called for the stored data to be forwarded to the next register. In the case of the initial interp assembler, this time interval will be taken as a nominal 50 msec. The time interval associated with the percept assembler will be taken as 200 msec. Both of these times assume the PROM located in the pulvinar has been able to provide a correct response to the individual interps and percepts. If not, the cycle times are extended while further analysis of the visual input from the retina is considered. This is the purpose of the oculomotor path shown. It causes the oculomotor portion of the POS to perform additional scans of the image and generate new or alternates information for processing into an initial interp by the multi-dimensional correlator.

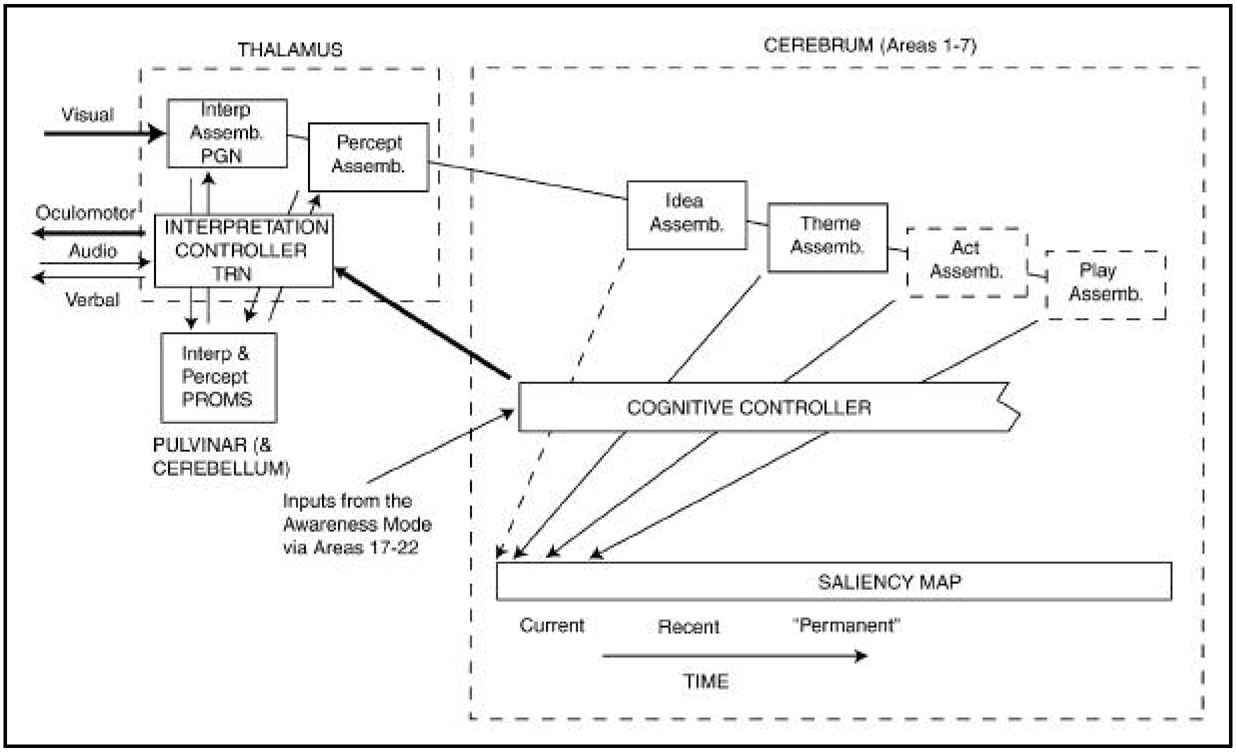

Figure 19.2.3-1 The functional organization of the reading mechanisms including the types of memory employed. The interp assembler receives individual interps from the multi-dimensional correlator of the pretectum. The right hand two boxes are dotted to suggest the task can continue to higher levels of comprehension. The arrow from the idea assembler is also dotted because of its less defined status. The heavy lines are indicative of paths used when learning (adding new comprehensions) or correcting previous comprehensions. See text.

The pretectum contains a multi-dimensional correlator that "digests" individual character groups (or other symbolic groups and described here as a syllable) and creates an interp. This interp is essentially a description of the syllable in terms of a group of straight lines of finite length and specific arrangement. The nature of the code describing these lines is unknown. It is likely that the code group used to describe a specific syllable is very similar to the one used to describe a sound to be created in speech, in response to that syllable. If so, it simplifies the short term memory requirements associated with the overall neurological system. If so, it would also suggest that the same code group would be formed in response to an aural input describing the same sound. This scenario of code group formation is in concert with how children are taught to read and speak. They are essentially taught to assign a verbalization first to what they hear and then (generally a few years later) to the same syllable presented to them in text.

In the context of this discussion, the initial step in reading is to compare the interp created by the correlator to stored interps related to a specific sound. This is also the process used by young readers who speak the sound they have been taught corresponds to the syllable on the paper. As they progress, the young reader begins to verbalize sequences of syllables. In this context, the sequential grouping of interps (syllables) leads to a percept. A percept is a simple concept that is usually devoid of a purpose. The purpose is determined when several percepts are assembled into a theme. Such an assembly may incorporate a verb, in order to form a sentence. Alternately, it may only incorporate a series of words forming a clause. In either of these two cases, the group will be labeled an idea. As the ideas are assembled, they generate a theme. A theme includes a complete thought. In this context, an idea is correlated with a sentence and a theme is correlated with a paragraph. For purposes of illustration, the figure shows the themes being assembled further into an act and the acts into a play. In this analogy, a theme or paragraph could be labeled a scene in a play.

The remainder of the figure is to suggest that the subject is not aware of the individual interps or percepts. In fact, his interest is not in the sentences. His cognitive processes are focused on the themes, acts and plays. This is a description of the cognitive processes involved in seeing. The subject is only conscious of the cognition related to the themes, acts and plays. He has no control over the lower level processes (except the ability to verbalize them in sequence).

19.8.3.3.2 Types of long term memory

19.8.3.4 The input to the cognitive portions of the reading mechanism

19.8.3.4.1 The Precision Optical System (aka Auxiliary Optical System)

19.8.3.5 The role of the control point in reading EMPTY

19.8.3.6 The role of the thalamus in reading EMPTY

19.8.3.7 The role of the cerebellum (and/or pulvinar) in reading EMPTY

19.8.3.8 The ultimate result of the cognitive portion of the reading mechanism

19.8.4 Performance in reading and analysis of fine detail EMPTY

Reading performance is highly dependent on the level and the quality of the light used to image the material. This is shown clearly in the description of the spatial contrast function as a function of signal to noise ratio. The criticality is stressed when the performance at a given signal to noise ratio is restricted to a small, square test stimulus.

CHAPTER 19 Mechanisms & Capabilities in Reading 3/4/03

(Active) SUBJECT INDEX (using advanced indexing option)

(Inactive) DEFINITIONS INDEX (Use individual marks)

Principle of Univariance
Retinal illuminance
Transport delay
net photoreceptor