Chapter 2

BASIC NEUROPSYCHOLOGY

2.1 INTRODUCTION

A human being is a biological entity. Any stimulation from the outside world has to have the effects on the human body before any action can take place. People come from different perspectives to study human mental behavior; therefore, there are neuroscience, human neuropsychology, cognitive psychology and engineering psychology. Neuropsychology is the science of the relationship between brain function and behavior (Kolb, 2000). Cognitive psychology deals with how people perceive, learn, remember and think about information (Sternberg, 1999). Engineering psychology specifies the capacities and limitations of the human cognition for the purpose of guiding a better design (Wickens, 2000). These three research areas are closely connected to each other, but their emphasis and purposes of the study are different. In history, cognitive engineering, or cognitive psychology looked at the brain as a "black box." By observing human behavior when they are given different stimulations/tasks (input) under certain conditions and by observing human subjects' reactions (output) to the stimulation, the researchers try to guess what is going on inside the brain. Neuroscience studies the structure of the brain and tries to allocate different functions in the brain, including the connection of different functions structurally. The purpose is to have better understanding of how a brain works biologically. Up to the '80s we knew too little about how the brain worked biologically, so the knowledge from this area provided very little help for the cognitive engineers for their design purpose.

As far as the technical development, neuroscientists have many useful tools to study the brain function. The research methodologies have been constantly improved in this area. It is probably time again to look at the development in the neuroscience and apply the knowledge to the cognitive engineering studies and to the design. Besides, to have a better understanding of human behavior, especially mental behavior, some basic knowledge from neuroscience and neuropsychology is necessary. In this chapter, I will give some brief summaries of the main theories and findings from neuroscience and psychology regarding the human sensation and perception, human language process, learning and memory.

The anatomy of the human brain is very complicated. It is difficult to describe it in a few sentences. Since the detail of the brain structure and the location of a specific function in the brain is not our interest, if any readers are interested in it, Bryan Klob and Ian Q. Whishaw's book "*An introduction to brain and behavior*" (2001) can be recommended. To prevent from the unnecessary "information overloading," I will try to avoid, as much as I can, using anatomical vocabularies in this chapter.

2.2 GENERAL ASPECTS OF NEUROSCIENCE

The information of neuropsychology came from many disciplines: anatomy, biology, biophysics, ethnology, pharmacology, physiology, physiological psychology and even philosophy. The central focus is to have a better understanding of human behavior based on the function of the human brain.

2.2.1 Basic Neuron Structure and Signal Transfer

The basic structure of the neurons includes the dendrites, cell body and axon. A neuron may have one to twenty dendrites, each of which may have one to many branches. The function of dendrites is to collect information and transfer it to the cell body. The axons from other neurons are connected to the dendrites. The information that appears in the neuron is the action potential on the membranes of the nerve cell. The information from other neurons is collected at dendrites and is processed at the cell body. The cell body integrates the information and then sends the integrated information through axons to the end feet, where it is passed on to a target neuron. Each neuron has a single axon. The information transferred from the axon of one neuron to the dendrite of the other neuron is via chemical synapse where there are synaptic transmitters that synthesize there. When an action potential arrives to the axon terminal, the chemical transmitters are released.

On the membranes of dendrites, there are receptors to receive the chemical transmitters. There are three types of effects that happen on the postsynaptic cell (Kolb, 2001): a) excitatory action; b) inhibitory action and c) initiating other chemical reactions. A schematic diagram of a neuron structure is shown in Figure 2-1. In the figure, "+" symbolizes the excitatory action and "-" symbolizes the inhibitory action.

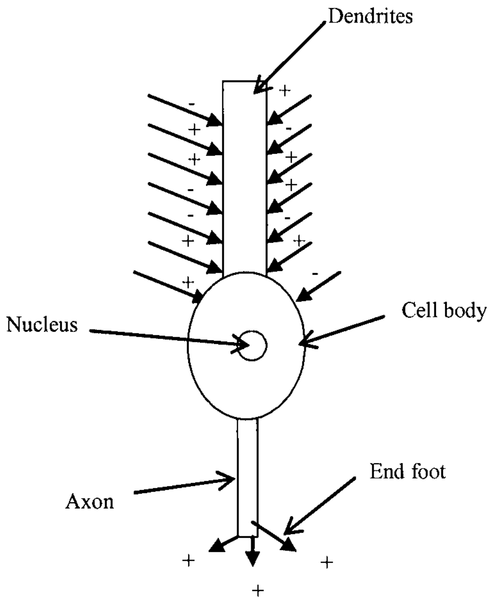

Figure 2-1. Schematic diagram of a neuron structure

There are various types of synapses in the central nervous system. An axon terminal can end on a dendrite, on another axon terminal or on another cell body. It may also end on a blood capillary or end freely in extra-cellular space. Dendrites may also make synaptic connection with each other. These different types of synapses increase the flexibility of behaviors of the nerve system. Even though synapses vary in both structure and location, they all do one of two things: either excite target cells or inhibit them.

2.2.2 The Principles of Brain Functioning

The discussion in Section 2.1 indicated one of the basic principles about how the brain functions: information is processed in a sequence of "in–integration–out." As we know now, most neurons have afferent (incoming) connections with many (tens or even hundreds) other neurons. The information is integrated in the neuron cell body and its efferent (outgoing) connections to many other cells. Basically, the nervous system works through excitation and inhibition.

Anatomically, the nervous system is divided into two main divisions, central (CNS) and peripheral (PNS). The peripheral portion is further divided into somatic and autonomic. The CNS is comprised of the brain and spinal cord, and the autonomic system is divided into the sympathetic and parasympathetic. Basically, the CNS receives input from the senses, processes this information in the context of past experience and inherent proclivities and initiates actions upon the external world. The PNS provides the sensory input and carries out the actions initiated by the CNS.

Sensory and motor divisions exist throughout the nervous system. The peripheral nerves going to and from the spinal cord can be divided into sensory and motor parts. In the spinal cord, there are separate sensory and motor structures. The distinct sensory and motor nuclei are present in different levels of the brain. In the cortex, there are separate regions for sensory input and motor output. Actually, the entire cortex can be viewed as being organized around the sensory and motor distinction. The CNS has multiple levels of function. Similar sensory and motor functions are carried out in various parts of the brain and spinal cord. This can be regarded as another principle of brain functioning.

Functions in the brain are both localized and distributed. In neuroscience, we normally assume that functions can be localized in a specific part of the brain. It is also found that the function is distributed in different parts of the brain. It depends on how we define the function. For example, the function of "language" has many aspects, such as "word" can be spoken, written, signed and even touched. It is not surprise that language function is widely distributed in the cortex.

The brain's circuits are crossed. Each of the brain's halves receives sensory stimulation from the opposite side of the body and controls muscles on the opposite side as well. The brain is both symmetrical and asymmetrical. The left and right hemispheres look like mirror images of each other, but they sometimes handle different tasks. For example, control of mouth movement is only on one side of the brain. Some motor functions are also controlled by one side of the brain to avoid the conflict of having different instructions from different sides of the brain.

Most neuroscientists subscribe to some assumptions: the brain is the source of behavior and there are certain reasons why the physical brain is organized with that of the mind. Regarding the brain functions, neuroscientists assumed that there are many relatively independent modules in the cognitive system. The unit of brain structure and function is the neuron. Each module can function to some extent in isolation from the rest of the processing system. Brain damage will typically impair only some of these modules. Investigation of cognition in brain-damaged patients can tell us much about cognitive processes in normal individuals. Researchers have tried to localize different functions on the brain by studying different functionally impaired patients and by electrophysiological confirmation of the localization. There is a hierarchical organization in brain function. The nervous system has three levels: the spinal cord, the brainstem (which is the central structure of the brain including the hindbrain, midbrain, thalamus and hypothalamus) and the frontal cortex. Each successively higher level that happens in the frontal cortex controls more complex aspects of behavior but does so through the lower levels (such as through the spinal cord).

When it comes to the location of different functions in the brain, there are opposite hypotheses. Neuron hypothesis assumes that the nervous system is composed of discrete, autonomous cells, or units, that can interact but are not necessarily physically connected. Nerve net hypothesis insists that the nervous system is composed of a continuous network of interconnected fibers. These assumptions and hypotheses lead to the research methodology design, results anticipation and theory development in neuroscience. Brain systems are organized both hierarchically and in parallel. It was found that functionally related structures in the brain are not always linked serially. The brain has many serial connections, but many expected connections are not found. This is probably due to certain interconnections that we don't understand clearly.

2.2.3 Methodology of Neuropsychological Studies

One of the traditional methods is to compare the cognitive performance from brain-intact and brain-impaired patients. The information can be used to support or to reject the theories within cognitive psychology. Such theories specify the processes or mechanisms involved in normal cognitive functioning. A typical example is Shallice and Warrington's (1970) investigation on a patient who had suffered damage to a part of the brain specialized for speech perception and production. This patient appeared to have severely impaired short-term memory but essentially intact long-term memory. This investigation supports partly the cognitive theory about the distinction between long-term memory and short-term memory.

In the last few decades, the electronic, digital and computing technologies have achieved fast development, and such development affects the neuroscience study. New measurement technologies appeared and provided the new opportunities and new methods to study the brain function. One of the most interesting factors is that we can study the human mental behavior without opening the brain, or waiting for special damage on the brain to appear. We can carry out neuroscience study on normal human beings. Some commonly used brain function measurement technologies will be briefly introduced here. For detail of their application, one should read some neuropsychology books.

An electroencephalogram (EEG) measures the summarized graded potentials from many thousands of neurons. EEGs reveal some remarkable features of the brain's electrical activity. By interpretation of the EEG records, one can find out that the brain's electrical activity is never silent; it has a large number of patterns, some of which are extremely rhythmical. An EEG's changes are coordinated to the behavior changes.

Among different patterns in an EEG, there is a record called event-related potential (ERP), which has been especially interesting to behavior science. It is a change in the slow-wave activity of the brain in response to sensory stimulus. One problem with ERPs is that they are mixed in with so many other electrical signals in the brain that they are difficult to spot just by visually inspecting an EEG record. One way to detect ERPs is to produce the stimulus repeatedly and average the recorded responses. Averaging tends to cancel out any irregular and unrelated electrical activity, leaving in the EEG records only the potentials that the stimulus generated. This characteristic has limited the ERPs application.

The electrical currents of neurons generate tiny magnetic fields. A special recording device known as a SQUID (superconducting quantum interference device) can record this magnetic field—magnetonencephalogram (MEG). Similar to ERP, the MEG procedure requires that many measurements be taken and averaged.

Positron emission tomography (PET) is a technique to measure the changes in the uptake of compounds such as oxygen or glucose due to the brain metabolism. Magnetic resonance imaging (MRI) is an imaging procedure in which a computer draws a map from the measured changes in the magnetic resonance of atoms in the brain. MRI allows the production of a structural map of the brain without actually opening the skull. For example, when a region of the brain is active, the amount of blood flow to it and oxygen consumption increases. A change in the oxygen content of the blood alters the blood's magnetic properties. This alteration, in turn, affects the MRI signal. Because oxygen consumption varies with behavior, it is possible to map and measure changes that are produced by behavior. The measurement of the magnetic properties of the brain for the distribution of elements, such as oxygen, due to certain functions is called function MRI (fMRI).

The major advantages over PET is that fMRI allows the anatomical structure of each subject's brain to be identified, and brain activity can then be related to a localized anatomical region on the brain image. It has better spatial resolution, but it is very expensive to perform such measurement.

Table 2-1 gives the summary of these electron or magnetic graphics methods. The advantage of these methods is that they provide the opportunities to study the brain functions and to locate the respective function center in the brain without damage to the brain. The disadvantage of these technologies is that the measurement equipment is huge and expensive, and the measurement has to take place in an advanced laboratory. It requires special training to be able to perform, measure and analyze the records.

Table 2-1. The summary of different electron or magnetic graphics methods

Name	Measurement mechanism
Electroencephalogram (EEG)	Measures the summarized graded potentials from many thousands of neurons.
Event-related potential (ERP)	A change in the slow-wave activity of the brain in response to sensory stimulus. It is part of EEG.
Magnetonencephalogram (MEG)	Measures the tiny magnetic fields generated by the electrical currents of neurons.
Positron emission tomography (PET)	Measure the changes in the uptake of compounds, such as oxygen or glucose, due to the brain metabolism.
Magnetic resonance imaging (MRI)	An imaging procedure in which a computer draws a map from the measured changes in the magnetic resonance of atoms in the brain.
function MRI (fMRI)	The measurement of the magnetic properties of the brain for the distribution of elements, such as oxygen, due to certain functions.

Just like many other scientific researches, people employ different research methodologies in neuroscience studies. The methodology that people use, either by studying brain-damaged patients, or by applying different electron or magnetic graphic methods on normal subjects, can influence the research results. Sometimes, by the development of the technologies and applying different research methodologies, some well-established theories may be questioned as well. For example, when studying the onset dynamic of certain activities and allocating it in the brain, the measurement speech of the technology may affect the results (Gage, 1998).

2.3 PERCEPTION

The perception started from receptors, or sensors, that are located on different parts of organs and skin in the human body. They convert the sensory energy (for example, the sound waves or light waves) into neural activity and conduct it through neural relays to the cortex, and the information will be coded in different ways there. The perception is a process of receiving the outside world information by the receptors and transferring the information to the brain. The brain recognizes the meaning of the information.

Different receptors perceive different kinds of energy, for example, visual receptors perceive light waves and auditory receptors perceive sound waves, and turn them into nerve discharges, and they follow certain pathways of nerves to the brain. Neuropsychologists still do not know how human beings or animals recognize that one set of discharges represents pictures and another set is the taste. It is assumed that everything we know comes to us through our senses. But our senses sometimes can deceive us, so some others proposed that we must have some innate knowledge about the world to be able to distinguish between real and imaginary sensations (Kolb, 2000).

2.3.1 The Sensory System

The sensory system starts from the receptor. The receptors are specialized parts of cells. They are different in each sensory system, and they perceive different kind of energy. The receptors act as a filter. They are designed to respond only to a narrow band of energy within the relevant energy spectrum. For example, the human ear can respond only to the sound waves between 20 to 20,000 Hz, and the human nose is not as sensitive as a dog's. Receptors normally receive information from a specific environmental field, such as a visual field. The receptive fields of individual receptors are overlapping. By comparing with neighboring receptors detect, it can also locate the sensory information in space. If one receptor is more activated by a stimulus than another, the object is located more precisely in the receptive field of those receptors. The detection of stimulus is often determined by receptor density and overlap.

Some receptors are easily activated but also stop responding quickly. They are called "rapidly adapting receptors." They react as soon as stimulated energy appears. The reaction can decrease if the stimulated energy does not change its intensity. In other words, they react quickly to the changes of the stimulated energy. For example, our ears can adapt to the noise. When we stay in a noisy environment for a while, we normally do not feel the noise level as high as in the beginning. There are also slow adapting receptors that will provide the information as the stimulated energy is still there. A typical example is the pain receptors. Our pain receptors do not adapt to the stimulation, as pain is a warning signal for the body to indicate certain injuries or damages in the body.

Each sensory system requires three to four neurons, connected in sequence, to get information from the receptor cells to the cortex. There are changes in the code from level to level and it is not a straight-through or point-to-point correspondence. Three important events can occur at the synapses between the neurons in the relay: a) A motor response can be produced, for example, axons from pain receptors stimulate first synapse in the spinal cord, where they can produce a withdrawal reflex. b) The message can be modified. c) Systems can interact with one another.

How does action potential in the neuron cells code the differences in sensations, and how do they code the features of particular sensation? For some sensory modality, we know part of the answer; for most of them, we still don't know. The increase or decrease of the discharge rate of a neuron is correlated to the presence of the stimulus. The amount of change can code the intensity. As in the visual sensory system, redder makes more discharge activity and greener makes less activity. There are different theories for explanations: a) neural areas that process these sensations in the cortex are distinct. b) We learn through experience to distinguish them. c) Each sensory system has a preferred link with certain kinds of movements, which ensures that the systems remain distinct at all levels of neural organization (Kolb, 2000). The distribution of the receptor projecting on the cortex is very complicated. There are at least seven visual systems projecting into different brain regions.

2.3.2 Sensory Modality

Normally, sensory modalities are touch, taste, smell, vision and hearing. There are actually many sub-modalities, or subsystems, within each of these modalities. These sub-modalities are distinctive with respect to their receptors, the size of the fibers that go from the receptors to the brain their connections within the brain, and the actions they produce. Each sub-modality is designed for a very specific function. For example, in the visual system, there are pathways to different locations of the brain for control of different functions as a) daily rhythms of such behaviors as feeding and sleeping in response to day-night cycles; b) changes in pupil size in response to light-intensity changes; c) head orienting, particularly to objects in peripheral visual fields; d) long-term circadian rhythms; e) movement of eyes to compensate for head movements; f) pattern perception, depth perception, color vision and tracking moving objects; and g) voluntary eye movement.

When neuropsychologists try to map the cortex for different functions of the human, it was found that the entire cortex behind the central fissure has some kind of sensory function. A view is emerging that the cortex is fundamentally an organ of sensory perception and related processes. Just as there are many visual systems projecting into different brain regions, there are multiple subsystems projecting to the visual cortex as well. It is demonstrated that some areas in the cortex were identified that had functions in more than one modality, for example, the vision and touch are in the same area. Gesture language and vocal language depend on similar neural systems. These areas, known as polymodal cortex or multimodal cortex, presumably function to combine characteristics of stimuli across different modalities (Kolb, 2000). There are three distinct regions of multimodal cortex. The existence of these three areas implies that there is probably more than one process that requires multimodal information, although it is not known exactly what these processes might be (Kolb, 2000). The knowledge of sensory modality and sub-modalities and the structure of their maps in the brain may lead to us to have a better understanding of how we understand and interact with the real world, and it may also lead us to understanding better about how we define the modality and how does different modalities interact to each other.

At the same time, the integration perception of multisensory input can also be an interesting issue. Viewing an incongruently articulating face categorically alters the auditory speech perception. This is the famous McGurk effect (1976). When it comes to the multisensory-audio/visual-input situation, it was believed that the integration of perception is dominated by one modality. Vision dominates audition is regarded as the most common case. A recent study indicated that either audition or vision domination is a matter of relative issue, not as an all-or-nothing condition (Andersen, 2004).

2.3.3 Cortical Connection and Function

As we discussed above, when the receptors perceive the stimulation, it turns the special stimulated energy (such as light waves, or auditory waves, etc.) into the active potentials and transmits them through different pathways to the cortex. The recognition happens in the cortex. It was also found that there are several locations in the cortex working together for recognition of the stimulation and reaction to it. Different perceptions can also end up on the same cortex location, as the multimodal cortex. Therefore, it is important to study the cortical connection. There are several interesting properties in the cortical connection (Kolb, 2000):

a) There is not a single area in the cortex that could represent entire perceptual or mental states.
b) All cortical areas have internal connections that connect units with similar properties.

c) There is a re-entry of the connection, which means that when a given area A sends information to area B, area B reciprocates and sends a return message to area A.

Now the most agreeable cortical function model is the distributed hierarchical system that is suggested by Felleman and van Essen (Felleman, 1991) as shown in Figure 2-2. There are several levels of processing in the associated area. The areas at each level and across the levels are interconnected with one another.

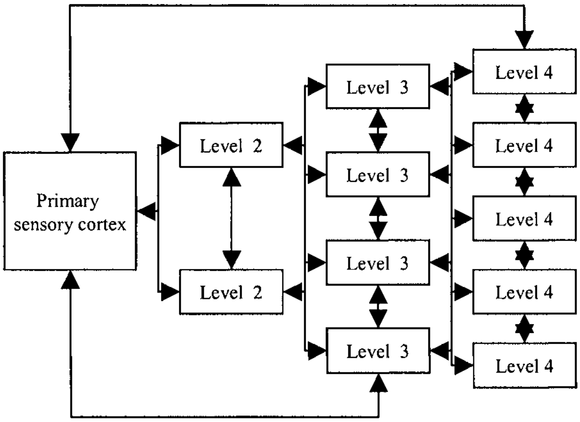

Figure 2-2. Distributed hierarchical mode of central nerve system

2.3.4 Perception and Recognition Theories

There are two different theories regarding perception. a) the *direct perception* (or bottom-up perception) theory: the array of information in our sensory receptors, including the sensory context, is allegedly sufficient to permit the individual to perceive anything. In other words, high-level cognitive processes are not necessary for perception; b) *constructive perception* (top-down perception): the perceiver builds (constructs) cognitive understanding (perception) of a stimulus, using sensory information as the foundation for the structure but also using other sources of information to build the perception (Eysenck, 1995; Sternberg, 1999). In this theory, perception is an active and constructive process. Perception is not directly given by the stimulus input but occurs as the end product of the interactive influences of the presented stimulus and internal hypotheses, expectations and knowledge, as well as motivational and emotional factors. This means that perception is prone to error, since it is affected by hypothesis and expectation (Eysenck, 1995; Sternberg, 1999).

Some studies support the direct perception theory, and some support the constructive perception theory. Nowadays, most psychologists agree that visual perception may be largely determined by bottom-up processes when the viewing conditions are good but may increasingly involve top-down processes as the viewing conditions deteriorate because of very brief presentation time or lack of stimulus clarity.

In word perception, the role of bottom-up processing and top-down processing was varied by altering the exposure duration and the changing the amount of relevant sentence context provided before the word was presented (Eysenck, 1995).

Stoffregen and Bardy (2001) reviewed over three hundred articles related to sensation and perception. They propose that perceptual systems do not function independently. They consider that perception is not a process of picking up information through a group of disparate "channels," and not as a set of interactions among discrete modalities but as the pick-up of information that exists in irreducible patterns across different forms of energy. By this approach, they proposed the concept of "global array" (GA). GA consists of spatial-temporal structures that extend across multiple forms of ambient energy. These patterns are higher-order in the sense that they are super-ordinate to (and qualitatively different from) the patterns that exist within single-energy arrays. Thus, in principle, information in the global array may be detected without prior or concurrent sensitivity to structure in single-energy arrays.

This theory can have broad implications for research on perception and action. In studies of perception, researchers may need to take into account the global information that is always available. They argue that researchers commonly present to subjects only a single form of energy. It is assumed that the application of stimulus energy to only one sense. Methodologically and analytically, it is grounds to ignore the energy available to other senses. However, when an experimenter stimulates a single modality, there is an influence on structure in the global array. If the GA exists, then it is a new challenge for the study of sensation and perception, especially in perceptual-motor learning and development, as well as the challenge of developing new research methodology.

The GA theory has taken the ecological consideration. It questions the validity of most traditional laboratory studies in this area when researchers tried to isolate the stimulation from the environmental context. Ahrens (Stoffregen, 2001) argues that it is impossible to discern the theory of GA. This theory appears to conflate sensation and perception. "They provide no evidence for the synergy between energy arrays they claim is the hallmark of the GA, nor do they provide evidence for novel sensors of higher order patterns; and they discount the behavioral and neurophysiologic evidence that sensory integration occurs within the brain."

The occipital lobe, an anatomical region in the cortex, is involved in the perception of form, movement and color. There are distinguishing locations in the brain that can be identified with different information that is involved in visual object recognition such as visual knowledge, semantic (associative functional) knowledge, and object naming (Eysenck, 1995). At the same time, there are multiple visual systems in the cortex for various aspects of stimulus recognition or for the guidance of movement in space. The multiple visual systems in the cortex are important for visual-spatial recognition. There are different mechanisms behind the facial recognition and word recognition, since facial recognition may take place on the right side of the hemisphere and word recognition may take place on the left side (Kolb, 2000).

One of the major functions of visually guided action is to allow us to interact effectively with our immediate environment. Processing the information about object location and object identification happens on separate parts of brain systems. In the cortex, there are three function zones that can be identified for control of movement. One (anterior zones) is primarily involved in somatosensory functions. The somatosensory system has more than twenty different types of receptor cells covering body tissues, except the brain itself. These include the hair, touch, temperature, nociceptors, joint, and muscle and tendon receptors. One region (superior parietal region) is primarily devoted to the visual guidance of movements of the hands and fingers, limbs, head and eyes. This region can be conceived of as having "spatial" function. We don't know what the nature of this spatial function actually is. The inferior parietal region has a role in processes related to spatial cognition and in what have been described as quasi-spatial processes such as ones used in arithmetic and reading (Kolb, 2000).

2.4 LANGUAGE

Language is one of the special capacities that separate human beings from other animals. The distinction between human language and the "communication" between different animals by sound, such as bird songs, is we use "words" and have the regulation of ordering words in different ways to give the meaning. In language, the fundamental sounds to make words are called *phonemes*. Phonemes are combined to form *morphemes*, which is the smallest meaningful unit of words. A string of words that follows certain grammar is called *syntax*. The stringing together of sentences to form a meaningful narrative is called *discourse*. The collection of all words in a given language makes up the *lexicon*. *Semantics* is the meaning that corresponds to all lexical items and all possible sentences. The vocal intonation that can change the literal meaning of words and sentences is called *prosody*. Language has the ability to represent an extension of the development of multiple sensory channels. We organize sensory inputs by assigning tags to information, which allows us to categorize objects and ultimately concepts. Language involves the motor act to produce *syllables*. A syllable is made up of consonant and vowels. Language requires the ability to impose grammatical rules, which dramatically increases the functional capacity of the system.

2.4.1 Neurology of Language

Neuroscientists can identify four important cortical regions that control language. In the left hemisphere, there are the Broca's area and Wernicke's area. Broca's area is believed to be the set of speech production and Wernicke's area for language comprehension. When we hear a word that matches one of those sound images stored in Wernicke's area, we recognize the word, while to speak words, Broca's area will play the important role, since the motor program to produce each word is stored in this area. If people suffer Wernicke's aphasia, they can speak fluently, but their language is confused and makes little sense (Kolb, 2000). There are another two regions of language use that are located on both sides of the brain: the supplementary speech area and the facial regions of the motor and somatosensory cortex.

How does the brain categorize sensory input and processing language? It was believed that there are three types of systems involved in this process. It was found that there are neural systems to categorize nonlanguage representations of sensory input. In these systems, it not only allows simple sensory stimuli to be combined but also provides a way of organizing events and relationships. In the left cerebral hemisphere, there are smaller regions in neural system that represents phonemes, phoneme combinations and syntactic rules for combining words. There is a third neural system that interacts with the category system and speech system to form the concept (Kolb, 2000).

The neurological models of language include nine facts (Kolb, 2000):

a) Capacity to categorize and to attach phonemic tags to the categories.
b) Visual and auditory language input has different processes.
c) Ability to produce syllables and quickly shift from one to another is important to produce language.

d) Categorizing verbs from movement is the key point for generating grammar in language.
e) Language is a relatively automatic process because it comes from repetition of well-learned responses or the mimicking of input.
f) Memory is required for language process.
g) Both left and right hemispheres are involved in different processes of language.
h) Language representation changes with experience.
i) There are individual differences on the cortical representation of language.

The language process includes the perception of speech or reading the text, word recognition and language comprehension. In the following sessions, we will have a brief introduction for each part of the process.

2.4.2 Speech Perception

Language can be perceived by listening to speech or reading text. Physiologically, speech perception follows the basic mechanism of auditory perception. The sound stimulates the auditory system with its air pressure. The auditory system then transfers changes in air pressure into what we perceive as sound. Sound has three fundamental physical qualities: frequency, amplitude and complexity. Perceptually, these qualities translate into pitch, loudness and timbre. From acoustic perspective, all human language has the same basic structure, despite differences in speech sounds and grammar.

Speech perception and reading text differ in a number of ways (Eysenck, 1995). One is received from auditory channels, and the other is visual. The differences of language perception via speech or reading text are summarized in Table 2-2.

Table 2-2. The differences between speech perception and reading text

Reading the text	Listening to speech
Each word can be seen as a whole	Spoken word is spread out in time
Text is less ambiguous	More ambiguous and unclear signal
The text can be continuously available	Higher memory demand—the information is transient, a word is heard and then it ends.
There is not so much redundant information behind the word itself	Contains numerous hints, such as pitch, intonation, stress and timing, to sentence structure and meaning.
Text is accessible only in the focus visual area	Can take input from any direction

Normally, listening to speech is more difficult than reading a text, as listening to speech demands more from memory, and the words that have already been spoken are no longer accessible. Speech normally contains numerous hints (known as prosodic cues) to sentence structure and the intended meaning in the form of variations in the speaker's pitch, intonation, stress and timing, which makes speech easier to understand. Young children often have good understanding of spoken language but struggle to read even simple stories.

There are several fundamental ways that make the difference between speech input and other auditory input. First, the sounds of speech come largely from three restricted range of frequencies, called *formants*. *Vowel* sounds are in a constant frequency band and *consonants* show rapid changes in frequency. Second, the same speech sounds vary from one context in which they are heard to another, while they are all perceived as being the same. For example, the English letter "d" is pronounced differently in words "deep," "deck" and "duke" when we measure its sound spectrogram. The auditory system must have a mechanism for categorizing sounds as being the same. This mechanism should be affected by experience. So a word's spectrogram is dependent on the words that precede and follow it. Third, speech sounds change very rapidly in relation to each other, and the sequential order of the sounds is critical to understanding.

Listeners use more than acoustic information when interpreting a spoken message during human to human communication. The history behind the conversation, the context during the conversation and the familiarity with the speaker and the subject are all aids to the perception. When available, visual cues, such as lip reading also assist. Lip reading provides the main cue for deaf people it also provides important cues for normal conversation to solve the confusions in the acoustic information. Because of these reasons, it adds certain perception difficulties for people to have telephone conversations and more difficult to listen to synthetic speech, even when the sound is very close to natural human speech.

The functional neuroanatomy of speech perception has been difficult to characterize, due to the fact that the neural systems supporting speech perception vary as a function of the task. Hickok and Poeppel (2000), after reviewing many research articles, proposed that auditory related cortical fields in the posterior half of the superior temporal lobe, bilaterally, constitute the primary substrate for constructing sound-based representations of speech. There are at least two distinct pathways that are involved in task-dependent speech perception, one is a ventral pathway and the other is dorsal pathway, as shown in Table 2-3.

Reading requires several different perceptual and other cognitive processes, as well as a good knowledge of language and grammar. Reading is basically a visual input in which the eye movement and word identification are the important parts. Both bottom-up and top-down processes operate within all three levels: feature, letter and word levels of recognition.

Table 2-3. The pathways for speech perception

	Ventral pathway	Dorsal pathway
Cortex location	Vicinity of the temporal-parietal-occipital junction in both hemispheres, but the computational mechanisms involved is probably different.	Inferior parietal and frontal system.
Function	Interfacing sound-based representations of speech with widely distributed conceptual representations.	Receptive phonetic processing such as performance on syllable discrimination and identification tasks. It involves coding speech in an articulatory context.
Tasks	Requires access to the mental lexicon*.	Requires explicit access to certain sub-lexical speech segments.
Serving	Perception of the semantic content of sensory information.	Interface system between sensory and motor processes. It mediates between auditory and articulatory representations of speech.

*The entries in the mental lexicon include phonemic, semantic and perhaps morph-syntactic information in speech.

2.4.3 Word Recognition

There are several different theories regarding the word recognition. Literatures support the suitability of different theories that apply to different situations. All of the theories have strengths and weaknesses. They can be considered from different angles to understand the process. There has been relatively little research on auditory word recognition and comprehension in the brain-damaged patient.

a) *Phonetic refinement theory*, or bottom-up theory (Pisoni, 1985): When the auditory signal is received on the receptors, it shifts to higher-level processing. The words are identified on the basis of successively paring down the possibilities for matches between each of the phonemes and the words we already know from memory.

b) *Phonemic-restoration effect*, or top-down theory (Warren, 1970): The speech we perceive may differ from the speech sounds that actually reach our ears because cognitive and contextual factors influence our perception of the sensed signal.
c) *Categorical perception* (Sternberg, 1999): Even though the speech sounds we actually hear comprise a continuum of variation in sound waves, we perceive discontinuous categories of speech sounds.
d) *Motor theory* (Lieberman, 1967): The motor signal produced by articulation is more informative than the speech signal, and it is our reliance on the motor signal that allows spoken word recognition to be reasonably accurate.
e) *TRACE model* (McClelland, 1986): It is assumed that bottom-up and top-down processing interact during speech perception. Bottom-up activation proceeds upwards from the feature level to the phoneme level and on to the word level, whereas top-down activation proceeds in the opposite direction from the word level to the phoneme level and on to the feature level.

All of these theories differ in different degrees, yet they all have some common assumptions. First of all, the auditory sensor input activates several candidate words in the early process of word recognition. The activation levels of candidate words are graded rather than being either very high or very low. Several parallel processes may happen, rather than occurring serially. Most of the theories propose that both bottom-up and top-down processes combine in some fashion to produce word recognition.

2.4.4 Language Comprehension

In any language, there is a possibility to generate an infinite number of sentences, but these sentences are nevertheless organized to systematically follow certain roles. These roles are regarded as grammar (or syntax). There are three main levels of analysis in the comprehension of sentences: syntactical (grammatical) structure, literal meaning and interpretation of intended meaning. Comprehension would be impossible without access to stored knowledge. The relationship between syntactic and semantic analysis is not clear. It is possible that syntactic analysis influences semantic analysis. Another possibility is that semantic analysis occurs prior to syntactic analysis. It is also possible that both analyses are carried out independently of each other.

The inference drawing process plays a crucial role. Many researchers are trying to identify more precisely which inferences are normally drawn. Different theories regarding language comprehension are about when and why the inferences are made. These theories are applicable in different situations. There is not an entirely satisfactory research method to study this issue. A few dominant theories are introduced here (Eysenck, 1995):

a) *Constructionist approach*: Numerous elaborative inferences are typically drawn while reading a text.
b) *Minimalist hypothesis*: Inferences are either automatic or strategic (goal-directed). In automatic processes, two kinds of inferences are constructed—concurrently process the parts of a text that are locally coherent representations and the information that is quick and easily available.
c) *Search-after-meaning theory*: Readers engage in a search after meaning based on their assumption of goals, coherence and explanation between the texts.

The minimalist hypothesis is preferable in many different aspects. There is a big gap between the minimalist hypothesis and the search-after meaning theory.

By studying language disorder symptoms, one can distinguish many different language functions. These disorders are: a) comprehension disorder such as poor auditory comprehension and poor visual comprehension; b) language production disorder such as poor articulation; word-finding deficit; unintended words or phrases; loss of grammar and syntax; inability to repeat aurally presented material; low verbal fluency; inability to write; and loss of tone in voice. Different language functions are distributed in a large region of the cortex. Language is organized like other cerebral functions in a series of parallel hierarchical channels. The evolution of language may not represent the development of a single ability but rather the parallel development of several processes such as the ability to categorize and the ability to use gestures for communication.

2.5 LEARNING AND MEMORY

Learning is a process that results in a relatively permanent change in behavior. *Memory* is the ability to recall or recognize previous experience; certain neural structures and circuits are associated with different types of learning and memory. Learning and memory involve a series of stages. When the learning material is presented, the learning process "encodes" the material (encoding stage). Some of the encoded information is stored within the memory system (storage stage). The stored information from the memory system can be retrieved through recovering or extracting.

Theories of memory generally consider both the structure of the memory system and the processes operating within the structure. *Structure* refers to the way in which the memory system is organized, and *process* refers to the activities occurring within the memory system.

2.5.1 Learning Theory

There are multiple forms of learning. The primary distinction can be made between Pavlovian conditioning, in which some environmental stimulus (such as a tone) is paired with a reward; and operant conditioning, in which a response (such as pushing a button) is paired with a reward (Kolb, 2001). Normally a learning theory contains five components (Coe, 1996): a) *Experiences*: learning is the result of experiences. The sensation-perception continuum forms the basis of the experience. b) *Schemata*: A schema is a mental framework or model that we use to understand and interact with the world. Experiences use sensation-perception to either create a new schema or modify an existing one. The easiest way to leverage a user's schemata is to use examples that center on schemata he already has in memory. Simplicity is the way to help users create schema. c) *Habits*: Connections between symbols and their corresponding actions are called habits. Habits have strength and family (a set of related habits). d) *Reinforcement*: It is the process of using events or behaviors to produce learning. There are both positive and negative reinforcement. To be effective, reinforcement must be systematic. Continuous reinforcement is reinforcing a behavior each time it occurs, and this is the quickest route to learning. Intermittent reinforcement is sometimes reinforcing a behavior, sometimes not. Vicarious reinforcement is learning to perform those actions we see others rewarded for and to eschew those actions we see others punished for. e) *Interference*: Old habit families interfere with new learning.

2.5.2 Memory

The study results from neuroscience have strong reasons to imply that every part of the nervous system is connected to learning, even the spinal cord. This would mean that the areas that process auditory information house auditory memory, areas that process visual information house visual memory, areas of the brain involved in producing movement house motor memories and so forth. It would also mean that memory could be further subdivided according to the specialized regions within each which houses the major functional modalities. Neuropsychologists believe that there are many kinds of memories, each of which has its brain location. Different types of memories can be found in the literature (Kolb, 2000):

a) *Short-term memory*: Memory for things that are retained only for a short period of time.
b) *Long-term memory*: Memory for things that are remembered for a long period of time.
c) *Implicit memory*: An unconscious, not intentional form of memory.
d) *Explicit memory*: Memory which involves conscious recollection of previous experiences.
e) *Reference memory*: Memory which refers to knowledge of the rules of a task.
f) *Working memory*: Memory which refers to events that can happen in a trial.

The distinction of the above types of memories is more or less from a functional perspective. There is no specific brain location that holds one specific type of memory. Among them, the long-term memory and the working memory have been especially interesting.

Short-term memory and long-term memory are parallel mechanisms in which material is processed separately and simultaneously (Kolb, 2000). Short-term memory is the memory that we use to hold digits, words, names or other items in memory for a brief period of time. It is sometimes called "working memory." Actually, the working memory and short-term memory is not exactly the same thing. Repeating a list of digits may be useful for putting the list in long-term memory, but it is not the short-term memory trace that is being placed in long-term memory. It is likely, however, that items can be pulled from long-term memory and held for some use in short-term memory. Short-term memory can be doubly dissociated from long-term memory as it can be impaired itself. It can have neural structures. There are probably a number of kinds of short-term memory, each with different neural correlates.

The processes of explicit and implicit memories are housed in different neural structures with different functions. Implicit information is encoded in very much the same way as it is perceived. This is data-driven or "bottom-up" processing. Explicit memory depends upon conceptually driven or "top-down" processing, in which the subject reorganizes the data.

Based on animal and human studies, Petride and Mishkin (1994) have proposed models for explicit and implicit memory. They demonstrated different areas in the cortex that are involved in explicit memory such as object memory, spatial memory, emotional memory, etc. At the same time, they found a brain circuit for implicit memory.

2.5.3 The Mechanism of Memory

Hebb's (1949) *cell assembly theory* has been almost the only theory from the neuron level to explain the mechanism of memory. He proposed that each psychologically important event—whether a sensation, percept, memory, thought or emotion—can be conceived as the flow of activity in a given neuronal loop. The synapses in a particular path become functionally connected to form a cell assembly. It is assumed that if two neurons (A and B) are excited together, they become linked functionally. In Hebb's view (1949), one cell could become more capable of firing another because synaptic knobs grew or became more functional, increasing the area of contact between the afferent axon and the efferent cell body and dendrites. Some studies have proved that there are qualitative changes in the synapses. These include changes in the size of various synaptic components, in the number of vesicles, in the size of postsynaptic thickenings, and in the size of the dendrites spines (Kolb, 2000).

Cell assembly represents the idea that networks of neurons (cell assemblies) could represent objects or ideas, and it is the interplay between those networks that results in complex mental activity (Kolb, 2001).

At least three problems arise (Kolb, 2000): a) A sensory experience could not change every neuron in the relevant systems, otherwise all subsequent experiences would be changed. b) If sensory experiences change sensory systems, thus permitting memories of the events, how do we remember ideas or thought? c) If experiences result in widespread changes in synapses, how do we find specific memories? Even though people are not satisfied with Hebb's theory, there is no better theory and these questions are waiting for answers.

2.5.4 Working Memory

Baddeley and Hitch (1974) suggested replacing short-term memory with working memory. This suggestion has been extensively accepted in cognitive science. The working memory system consists of three components: a) A modality-free *central executive resembling attention.* It has limited capacity and is used when dealing with most cognitive demanding tasks. It is like an attention system; b) an *articulatory or phonological loop* which holds information in a phonological (i.e., speech-based) form. Articulatory loop is determined by temporal duration in the same way as a tape loop. An experimental study showed that the subject's ability to reproduce a sequence of words was better with short words than with long words (Eysenck, 1995). A phonological loop consists of: a) A passive phonological store that is directly concerned with speech perception;

b) An articulatory process that is linked to speech production and gives access to the phonological store. c) A visuo-spatial *sketchpad* that is specialized for spatial and/or visual coding. An auditory presentation of words permits direct access to the phonological store regardless of whether the articulatory control process is being used. In contrast, visual presentation of words only permits indirect access to the phonological store through subvocal articulation (Eysenck, 1995).

The central executive is the most important component in the working memory. There has been relatively little clarification of the role played by the central executive. The capacity is very difficult to measure (Eysenck, 1995). A major function of the phonological loop is to facilitate the reading of difficult material, making it easier for the reader to retain information about the order of words in text during the process of comprehension (Eysenck, 1995). The characteristics of the visuo-spatial sketchpad are less clear than those of articulatory loop. Baddeley (1974) defined it as "a system especially well adapted to the storage of spatial information, much as a pad of paper might be used by someone trying, for example, to work out a geometric puzzle."

There are several advantages with working memory model (Eysenck, 1995):

a) It is concerned with both active processing and transient storage of information.
b) It is better placed to provide an explanation of some findings from neuropsychology with brain-damaged patients. If brain damage affects only one of the three components of working memory, the selective deficits on short-term memory tasks would be expected.
c) It incorporates verbal rehearsal as an optional process that occurs within only one component.

2.5.5 Forgetting

Everybody has the experience that after a certain period of time forgets a lot of things learned, done and talked about. What makes us forget about things? What is the mechanism behind it? Why are there certain things we wish to remember, such as the knowledge we learn the school, but keep forgetting, while some bad experiences that we wish to forget, we remember for the rest of our life? The scientists are not able to answer all these questions yet. There are not studies from neuroscience to demonstrate why forgetting happens.

There are three theories about forgetting; one is interference theory, one is decay theory (Sternberg, 1999) and the third is *functional decay theory.*

Interference theory assumes that our subsequent ability to remember what we are currently learning might be disrupted or interfered with by what we have previously learned (proactive interference) or by what we learn in the future (retroactive interference). *Decay theory* asserts that information is forgotten because of the gradual disappearance, rather than displacement, of the memory trace. This theory is very difficult to test! Whereas interference theory views one piece of information as blocking another, decay theory views the original piece of information as gradually disappearing unless something is done to keep it intact. Evidence exists for both theories, at least in short-term memory. The *functional decay theory* has proposed that forgetting the current task is functional because it reduces interference—if the current task has already decayed by the time the next task comes along, it will cause less interference once the next task is current. Altmann and Gray's (2000) experiment indicates that forgetting helps by preventing stale information from interfering with fresh information. Forgetting is difficult when the world changes faster than memory can decay. Cognition responds to this limitation by starting to forget an item as soon as it is encoded.

2.6 CONCLUSION

Neuropsychology is the science of studying the relationship between brain function and behavior. Within the limited space, we can only introduce some basic concepts about how the brain is constructed and functions in general. We also tried to introduce the recently development in the neuroscience regarding different cognitive behavior such as language process, learning and memory.

The brain consists of a very large number of simple processors and neurons, which are densely interconnected into a complex network. These large numbers of neurons operate simultaneously and co-operatively to process information. Furthermore, neurons appear to communicate numerical values of information from input to output.

The knowledge from neuroscience may tackle some cognitive theories and information processing theories, as they normally consider information is coded as symbolic messages. In the speech and language process studies, research has concentrated on modeling human behavior. The data is taken mainly from cognitive psychology, or linguistics, rather than neuroscience. Even parallel distributed processing, or neural network, or neuro-computing, normally bear no relation to brain architecture and lack biological realism (Christiansen, 1999).

There should be historical reasons behind these phenomena. One of the main reasons is because we still lack enough knowledge about how the brain works. The knowledge gained from traditional methodologies of studying brain-damaged patients is far from enough to support the computational design and applications. The development of the new technologies, especially the electron or magnetic graphics technologies, provides a new possibility for neuroscience to study the functional mechanism of the brain. The results from neuroscience studies based on new methodologies may provide the opportunities for the computational models searching for biological realism. The requirement from computational modeling may also set considerable influence on neuroscience development, especially on language process and memory research.

http://www.springer.com/978-0-387-24155-5

Designing Human Interface in Speech Technology
Chen, F.
2006, XXIV, 382 p., Hardcover
ISBN: 978-0-387-24155-5