

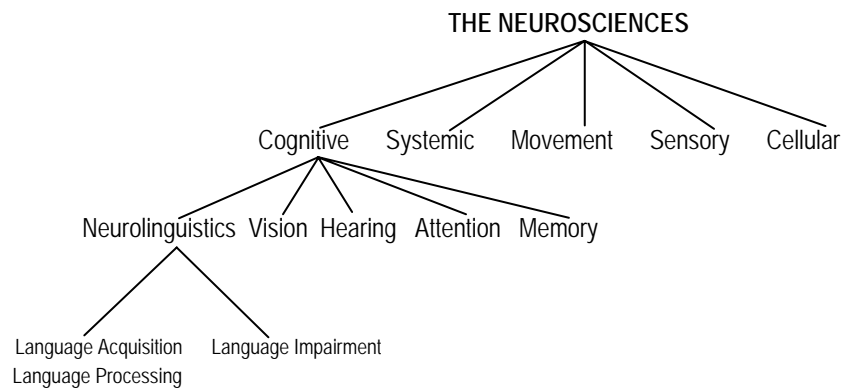
Introduction to Neurolinguistics

Aniela Improta França
UFRJ (Federal University of Rio de Janeiro, Brazil)

This is an account of *Introduction to Neurolinguistics*, a week-long course that took place during the First South American Summer School in Formal Linguistics (EVELIN '04), UNICAMP, São Paulo, Brazil. Classes ranged from an overview of brain anatomy and an introduction to neurolinguistics protocols to the comparative analysis of neurolinguistics and medical research about the Language Faculty.

1. About the course

Neurolinguistics is a branch of Cognitive Neuroscience, that, on its turn, together with Systemic, Movement, Sensory and Cellular, is a branch of a larger domain named the Neurosciences. Neurolinguistics can still be divided into two areas: language acquisition and processing and language impairment.



The focus on language impairment is a historic one, dating from 400 b.C., with Hipocrates' accounts on infirmities that produced lack of language. Contrastingly, the questions about the healthy Faculty of Language – how we come to acquire and use our mother tongue – have been systematically taken for granted through a long stretch of history, despite the fact that language is the one cognition that definitely sets us apart from other animals on this planet. In reality, language investigation has only taken a definite bio-linguistic course in the 1950's with the advent of Noam Chomsky's Generative Grammar (Chomsky, 1957, 1965). And the neurophysiological characterization of the healthy Faculty of Language, that is, the understanding of language-brain

relations at work, only started being investigated specially in the late 1980's, with the introduction of non-invasive cognitive assessment techniques that brought new and exciting perspectives into the field. These very perspectives should be the core of the course *Introduction to Neurolinguistics* that I was invited to teach during the First South American Summer School in Formal Linguistics (EVELIN '04), that took place at UNICAMP, São Paulo, Brazil.

The fact that these are new perspectives should give me substantial teaching materials to fill my five two-hour classes. But the fact that they are exciting, or even, *very* exciting, might bring in a challenge: avoiding to lead my audience of novice students from South America, with little or no training in linguistics, to what my former advisor, Miriam Lemle, has termed “a dangerous scientific shortcut” to the understanding of language in the brain. When there were no direct cognitive assessment techniques available, outputs from an assortment of languages were thoroughly analyzed by linguists and psycholinguists, and a lot came to be revealed about the nature and complexity of this cognition. Now we can look inside the brain and see it decoding and processing language! But for anything to make further sense than just the appealing sight of colorful parts of the brain at work, or of twitching synapses or of squiggled lines representing cortical electricity, one has to know exactly *what* cognitive task is related to such effects. And then how. This means resorting to a complex blueprint carefully thought out and consolidated by linguistics theory during the past fifty years. And here there are no possible shortcuts to take, no five-day introductory courses and no cracking handbooks to shorten this ride.

Whenever possible, the rule of thumb I used was to channel the excitement that the mind-brain field naturally provokes at first hand into motivating students to take the long ride, all the way from linguistics into the neuroscience of language.

Inspiration to format my course was also taken from David Poeppel's series of lectures on neurolinguistics at the Federal University of Rio de Janeiro (UFRJ), in 2001. His visit marked the debut of neurolinguistics of healthy language in Brazil and was the starting point for much of the interdisciplinary research in this area involving the linguistics and the biomedical engineering departments and the medical school at UFRJ.

2. The mind-brain relationship (Day 1)

2.1 Acquiring knowledge about what we don't see

Although everything we see in our visual field appears to be seamless, that is, objects imperceptibly blend into one another, we have a clear notion of the boundaries, shapes and colors of the objects we see, even if some of their facets are concealed from our point of view. This very efficient sense is a result of an incredibly coordinated work between the eyes and the brain. Each of our eyes sees one image. Then the brain combines the two images into one to make vision. However, these images are slightly different from each other. The brain uses these small differences to figure out how far away and how deep an object is. During the process, the eyes move continuously, on a fleeting time scale, to update the information that is being sent to the brain, including details of edge,

shape, motion and color. In a few milliseconds we are able to acquire in our brains an accurate mental representation that matches what is in our visual field.

Our curiosity about the contents of opaque containers certainly derives in part from this accurate binocular vision. When we look at three-dimensional objects, we infer that there is a space underlying their surface. Is this space filled? Infants shake closed objects and tap them against the cradle. Toddlers try to crack open clocks and toys to see the inside. Children are specially attracted to objects that do something: produce a sound, move around, light up. Such objects establish a cause-effect relationship that motivates children to theorize and look for the source of motion inside and not on the surface of what they see.

Adults do not need to open every opaque container they see. Based on their knowledge of the world already represented in their minds, they know or are able to make educated guesses about the contents of closed containers. This system of world knowledge acquisition that maps objects to abstract, mental representations gradually breaks children's innocence in relation to the contents in the world around them.

Since world concepts and objects are constantly updated along with collective history, the break of a child's innocence happens in accordance with the technology and the scientific bias of a specific time. While a child from our time has to build a mental representation of the cell phone, one from the middle ages must have built a mental representation of a knight's armor. If we apply this reasoning further, to the dawn of mankind, we might infer that the curiosity about the contents of the skull - a round hard human part, with holes, through which smell, sound and taste are sensed - might have led our prehistoric ancestors to tamper with the inside of the mysterious container. In fact, proof of that blunt curiosity came with the finding of the 5,000 year-old remains of a man whose skull was carefully opened while he was still alive (fig. 1). Although we can only speculate about the reasons that led this surgeon to perform the operation, it is legitimate to guess that he was, at least in part, motivated to get to know what was underlying the bone.

As we will see in chapter 3, the evolution of neurology is closely connected with man's direct access to study the brain and with the possibility to relate anatomical and functional findings to the outputs of the neurological mechanisms, that is, the establishment of a solid mind-brain relationship.

2.2 Some essential brain anatomy to the understanding of the mind-brain

More specifically, understanding this complex relationship means connecting the subtle effects of cognition – speech, hearing, vision, motor coordination, memory and others – to the brain, the material portion that gives rise to cognition.

2.2.1 The brain and its protections

The brain or encephalon (fig. 2) is a central organ, encapsulated by the skull or cranium, the uppermost portion of the human skeleton. The soft tissue enveloping the cranial vault is the scalp. Underneath the skull, the brain is still

protected by three structures called meninges (plural of *meninx*, from the Greek, *covering*).

The outside meninx, called the dura mater, is the most resistant of the three. It is dense, leatherlike and inelastic. The center membrane is the arachnoid. It lies under the dura mater and is sparsely traversed by connective-tissue spikes named *trabeculae*. They help cushion the brain from impact against the skull. Under the arachnoid lies the innermost meninx, the pia mater. It is a thin vascular membrane, consisting of a minute plexus of blood vessels held together as they perforate the meninx tissue. The pia mater tightly coats the cortex, following its irregular surface. It is separated from the arachnoid by the subarachnoid cavity, which is the space between the arachnoid and the pia mater, filled with cerebrospinal fluid (fig. 3).

2.2.2 *The hemispheres and the lobes*

Underneath the meninges is the cerebrum divided in two cerebral hemispheres – the right and the left (fig. 4). The cerebrum constitutes the largest part of the brain. Specific regions from both hemispheres are known to be recruited to analyze sensory data, perform memory functions, learn new information, form thoughts, make decisions and articulate language, although there is a tendency for these and other cognitive functions to sit, predominantly, in one of the two hemispheres.

The right and left sides communicate with one another through a bundle of nerve fibers, named *the corpus callosum*. More internally, another smaller bundle of fibers makes the connection: the *anterior commissure*.

The two cerebral hemispheres can be further divided into analogous sections on the left and right sides, named lobes. There are four pairs of brain lobes: frontal, parietal, temporal and occipital, each of which specializing in different functions (fig. 5).

The frontal lobe bilaterally makes up the anterior (front) portion of the cerebral cortex and is positioned right behind the forehead. In general terms, the frontal lobe is connected with the orchestration of higher cognitive functions. Its anterior part (front), named pre-frontal cortex, is responsible for the articulation of speech (Broca's area), selective attention, planning, organizing and problem solving. The posterior (back) section consists of the pre-motor cortex, that stores motor patterns and voluntary activities, and the motor cortex, which regulates voluntary movement. The frontal lobe is separated from the parietal lobe by the central sulcus (fig. 6).

Just behind the frontal lobe is the parietal lobe. The anterior parts of the parietal lobe, touching the central sulcus, is the primary sensory cortex. Information about taste, touch, temperature, written language is processed in this cortex, as input comes from the sensory organs. The parietal lobe is recruited in the processing and discrimination of all sensory input and body orientation. It is also required to decode written language.

The occipital lobe is situated just behind the parietal lobe. Its posterior part makes up the primary visual cortex, where the first visual information coming from the eyes is processed. Then, toward the anterior portion, lie the secondary and tertiary visual cortices, in which higher visual processing takes

place, following two visual pathways: the dorsal, for movement, and the ventral, for shape and color. Damage to the occipital lobes can cause blindness (fig 7).

Finally, the temporal lobe lies in front of the visual area, nested under the parietal and frontal lobes and separated from them by the Sylvian fissure. The temporal lobe receives and processes sound information directly from the ear, since it hosts the primary auditory cortex and the Wernicke's area, that is connected in many ways with language decoding. A lot of sensory processing also takes place in the temporal lobe and also do information storage and maintenance, specially in the undersides of this lobe. Other temporal parts process music and underlie the integration of memory with sensory sensations of sight, touch, sound and taste (fig. 7).

2.2.3 The brain cells

On the microscopic level, the brain is constituted by two cell types: neuroglia (glia, the Greek for *glue*) and nerve cells. Neuroglia are about fifty times more numerous than neurons in the brain. Neuroglial cells, or simply glial cells, are peculiarly branched cells arranged in a fine web of tissue. Traditionally, they were only regarded as the supporting structure of nervous tissue. However, neuroglia are now known to be involved in neuronal growth and migration. There are three types of glial cells in the brain: astrocytes, oligodendrocytes and microcytes (fig. 8).

Astrocytes are star-shaped cells that, besides a major supportive role in the central nervous system, are also recruited in the formation of the network on which neurons grow. Filling in the space within neurons, astrocytes hold neurotransmitters, that are released by neurons, and provide part of the fundamental control of chemical concentrations in extracellular space. Astrocytes also clean up brain waste and digest dead neurons. More recently, these cells have been connected with the controversial issue of adult neurogenesis, that is, the growth of new neural tissue.

Oligodendrocytes (astroglia and oligodendroglia) are involved in myelin formation, which functions as insulation to the nerve fibers. It also provides transport of material to neurons and the control of the ionic environment of neurons.

Microcytes (microglia) are the smallest of the glial cells. They phagocytize waste products of nerve and play an important role in the immune system of the brain by protecting it from invading microorganisms.

The other type of cell in the brain is its primary functional unit: the neuron. Until recently, the official estimate was that the brain held 100 billion neurons. This figure was challenged by research with more advanced cell count techniques that estimated the cell count of 100 billion just for the cerebellum (Andersen, Korbo and Pakkenberg, 1992).

Neurons operate in large sets forming neuronal circuits or neural nets. The most striking feature of neurons is that they produce electrical signals that operate as information bits. All data reaching or leaving the body is transformed into electrical signals that are transported and processed by neurons (fig. 9).

Each neuron can be divided into three parts: cell body or soma, dendrites and axons. The cell body contains the nucleus and other cell organelles such as mitochondrion, rough and smooth endoplasmic reticula, lysosomes, ribosomes and Golgi complex, that are responsible for cell metabolic maintenance. Departing from the cell body are dendrites and axons: cell extensions, also named *processes*. Dendrites are tree-like structures that receive messages from other nerve cells. Neurons are electrically excitable and are known to be able to transmit this excitation. So, an electrical signal reaches a dendrite, goes through the cell body and travels away from the cell body down the axon as an impulse called action potential. The axonal process is coated by an insulating glial tissue named myelin. This special coating allows the action potential to move fast toward the end of the axon, that branches out and makes connections (synapses) with the dendrites of other neurons.

2.2.4 The two tissue layers

The special bundling of neuron and glial cells together forms two layers of brain tissue that fill the cerebral cavities. The outermost layer is the gray matter. Under it lies an inner core of white matter, a much thicker layer (fig. 10).

The gray matter is a vital 2 to 6 mm layer of tissue known as the cortex (Latin word for *bark*). It is structured by nerve cells and fibers of different sizes and shapes, but it is specially rich in nerve cell bodies, that are gray to the naked eye. Most cognitive processing takes place in the cortex. This tissue is accommodated within the limits of cranium vault through intense folding, forming bulges and creases, named respectively *gyri* and *sulci* (plural of *gyrus* and *sulcus*). This convoluted surface enlarges the extension of the sheet of gray matter available for processing, resulting in the possibility for the human being to compute incredibly large quantities of information. Other animal species do not have the same amount of folding as we do and, therefore, have more limited processing space.

The white matter composes the deeper parts of the brain. Despite the fact that it is also formed by nerve cells, it is essentially different from the cortex, because it is richer in nerve cell axons, arranged in bundles or fibers. These fibers are sheathed with myelin, an insulating layer made up of glial cells, rich in protein and fatty substances that are white in color to the naked eye. The purpose of the myelin sheath is to allow rapid and efficient transmission of impulses along the nerve cells.

There are three types of fibers in the white matter: (i) association fibers, that interconnect structures within the same hemisphere; (ii) projection fibers, that unite the hemispheres to the lower parts of the brain; and (iii) commissural fibers, that connect the two hemispheres.

2.2.5 The action in microperspective

While the gray matter is mostly connected with information processing, the white matter takes over information flow. Resorting to a between-neuron transmission system, named synapse, the action potential can travel at high speed along an incredibly ample network of neurons.

There are two kinds of synapses: electrical and chemical. The electrical synapse is the direct connection between two neurons that are very close to each other, so there can be direct transfer of the electrical signal. In fact, in the narrow gap between the two cells there is a junction of protein that allows ions to go from the cytoplasm of one cell to that of the next. Although the electrical synapse is much less common in mammals, they can be found in human fetuses, in early embryonic neuronal development, and in adults, between non-neural cells and in cardiac, glandular and liver cells.

Chemical synapses are highly complex information systems about which whole books have been written and a lot is still unknown¹. But, in essence, the chemical mechanism that was first described in the early 1900's, by the Austrian neuroscientist Otto Loewi, consists of a contact zone between two neurons: a presynaptic neuron and a postsynaptic one (fig. 11).

In the presynaptic phase, a fleeting electrical signal of digital nature named action potential is fired and travels towards the end branches (buttons) of an axon. The arrival of the action potential to the axonal buttons triggers the release of molecules of neurotransmitters that are kept in synaptic vesicles.

These molecules reach the synaptic cleft, that is, the contact zone with the next neuron, and are absorbed by receptors placed at the postsynaptic membrane of the dendrites of the next cell.

Depending on the neurotransmitter released, the action potential that will flow to the next neuron may be intactly propagated, blocked or modified. This means that the electrical bits being transmitted can be modulated by such neurotransmitters.

Synapses are very productive processes. After one firing takes place, it takes one-thousandth of a second for the neuron to be ready to transmit another action potential. Since all this is a fleeting process that can be reiterated indefinitely, and since the number and frequency of firings varies in relation to external stimulation, the action potentials function as digital codes that can be deciphered by cortical processing. This is how contents that reach the brain, such as language, sensations, planning for movement, thoughts, memories and feelings, ultimately make sense to the mind. So this also is how the mind-brain relationship is established in microperspective.

3. Historical landmarks that came before neurolinguistics (Day 2)

3.1 When 10 is more than 100,000

Incredibly as it may seem to us now, the thought about thought and about other mental processes took a long time to be clearly posed as pertaining to the mind-brain relationship. In fact “a long time” is an understatement. From the dawn of mankind, 100,000 years ago, until 10,000 years from now, our ancestors barely managed to fend for themselves. Only after the basic survival advancements were established could mankind transcend to more complex questions about the nature of human existence and the role of the brain.

¹ For a comprehensive neuroscience view of synapse, (cf. Bear, Connors & Paradiso, 2001, p. 92)

Until the nineteenth century, the questions that specialized the history of neurology were updated very slowly and depended on a number of extraneous variables, such as the political regimen, dominant religion and philosophical tradition. These defined, among other things, the possibility to perform postmortem dissections and the ability to establish a scientific method of observation of the brain connecting it to its cognitive products, among which language was, perhaps, the most mysterious.

However, in the last ten years, this once slowly evolving history zoomed into the end of the twentieth century under the name of the neuroscience of language or basic neurolinguistics and attained light speed in the twenty first century, with the new technological breakthroughs – the neural chips, the functional imaging machines and the high density EEG and MEG – displaying breathtaking perspectives for the highly interdisciplinary teams of linguists, neurologists, biologists, electrical engineers, and psychologists that are devoted to the study of cognition of language.

3.2 The first concern: the brain or the heart?

Starting off on the slowly-evolving part of this history, the first principled tentative to understand the brain happened in ancient Greece, with Hippocrates (460-377 b.C.), the founder of Western medicine, who believed not only thought and reason, but also feelings and moods were produced in the brain: "Men ought to know that from the brain, and from the brain only, arise our pleasures, joys, laughter and jests, as well as our sorrows, pains, grievances, and tears" (Hippocrates, 400 b.C.). Hippocrates' theory is known as cerebrocentrism. With colleagues and students, he performed postmortem dissections and followed the neuronal fibers from the organs of sense (nose, eyes, mouth, ears) to the brain.

But the idea of cerebrocentrism was soon overridden. With an increased concern about man's spiritual connections, Plato (427-347 b.C.) and Aristotle (384-322 b.C.) formed a new current of thought, named cardiocentrism. Instead of the brain, the *heart* was viewed as the essence of man: "The seat of the soul and the control of voluntary movement - in fact, of nervous functions in general - are to be sought in the heart. The brain is an organ of minor importance."(Aristotle, 350 b.C.)

Influenced by the advent of the huge aqueducts that saved Greek cities from epidemics, by taking pure water from mountain streams to the population, Aristotle defended that all our health was guaranteed by the flow of four basic fluidic substances called humors: blood, phlegm, yellow bile, and black bile. These four humors were thought to be bodily representatives of the four elements of nature: earth, fire, water, and air. Cardiocentrists thought that when humors were balanced, the person was healthy. All diseases – mental and physical – derived from humor excess, deficit or overheating. The brain's minor role in this system was restricted to the cooling of the body, whenever the core cardiac function increased the blood temperature to dangerous levels.

Although the great anatomist Herophilus (335-280 b.C.) reinforced Hippocrates' cerebrocentric views and contributed to the description of the brain

by distinguishing between motor and sensory nerves, cardiocentrism and the Theory of Humor still prevailed well into the Roman Empire.

3.3 The focus on the magic nature of brain space

A definite reinforcement to cerebrocentrism came from an eminent Greek physician, Galen (130–200), who was very exposed to open brain injuries in gladiators and to the effects that these injuries provoked. Although post mortem dissections of the brain were not allowed in his time, Galen noticed that the cerebrum tissue in his patients was rather soft, while that of the cerebellum was firmer and coarser. These observations made him conjecture that the cerebrum was designated for thoughts, feelings and sensations, and that the cerebellum controlled the muscles. Although the firmness of the tissue has no direct bearing on specific brain function, his conjectures turned out true to a certain extent. Subtle cognitive functions are indeed performed in the cerebrum and the cerebellum is the center of muscular balance. But many other motor functions are processed within the cerebrum.

With the legal constraints on human dissections at that time, Galen focused his studies on dissecting the brain of many animals. From this endeavor, he invariably found that the brain had hollow spaces (ventricles) filled with fluid. Thus, despite reinforcing cerebrocentrism, Galen also came to sustain the Humor Theory, by proposing that the brain sensed cognitive contents and initiated muscular movement from the flow of intraventricular fluids.

Additions to the Humor Theory continued being made for another fifteen hundred years, and the traditional Galen's view of medicine was basically unaltered. Among these additions, the two most important ones were advanced by the Belgian anatomist Andreas Vesalius (1514-1564) and by the French mathematician and philosopher René Descartes (1596-1650).

Differently from Galen, Vesalius could make real dissections of the human body. His findings enabled him to discover that the Galen's system of medicine was based on fundamental anatomical errors. In his monumental *De Humani Corporis Fabrica*, published in 1543, Vesalius identified over two hundred galenic errors, including Galen's description of the *rete mirabile*, a coil of blood vessels at the base of the brain, a structure that does not exist, but that was fundamental to the proposed Humor Theory. Despite the clear demonstrations that the galenic model of the brain was wrong, Vesalius' work could not be fully appreciated until after his death.

In the seventeenth century, with the advent of machines that were powered by water flow, the galenic view of the brain was once again reinforced. Descartes presented a Hydraulic Model of the brain, which specialized the galenic idea that motor behavior was the result from the pumping of fluid through the ventricles. However, in his version, Descartes also believed that the Hydraulic Model did not account for man's soul and intellect. These were exclusively human, God-given treasures inserted in the mind, which, by the way, becomes a new and vital player in the history of neurology. For Descartes, the mind was able to influence the pineal gland to control the movement of the animal spirits through hollow nerves that pumped up the muscles.

3.4 And now, substance!

It was only during the eighteenth century that the galenic brain was completely abandoned. From this point on, the direct observation of the open brain turned scientists' attention from brain space to substance. Under this new point of view, the pattern of convolutions was studied and so were the brain tissues and lobes. This new concern aroused great speculation about brain localization, and this trend marked most scientific pursuits until the twentieth century.

In the nineteenth century, however, a series of unrelated facts will intertwine and bring hallmarks to the evolution of neurology and the neurosciences, and most especially to the neuroscience of language.

An accomplished Viennese neuroanatomist, one of the first to describe the brain circumvolutions with accuracy and the measure in which they differed from person to person, Franz Gall (1758-1828) developed the belief that each part of the brain was responsible for one instance of man's behavior, personality and aptitude. Making a comparison with muscular atrophy that depends on muscle mobilization, Gall believed that if a person exercised a specific behavioral instance or personality trait, the part of the brain that controlled it would be enlarged and consequently would deform the skull over it in a specific way. Thus, he purported that by examining the bumps and indentations on the skull, an easy and non-invasive practice named *cranioscopy*, one would be able to map a person's emotional and intellectual functions (fig. 12).

Gall and his student Johann Spurzheim (1776-1832) checked his theory by examining skulls of criminals and insane people. After analyzing a big number of *cranioscopies*, they came up with 27 distinct topographic *organs* of the brain, including a murder organ and a benevolence organ. This technique, known as *Phrenology*, drew incredible attention in Europe and in the US and was influential even to the view of the brain that we have today.

After the regions were mapped, *Phrenology* became much too popular and served fortune-telling purposes. It started being disseminated all over the world by *skull readers* practicing the *art of measuring the skull* in offices and parlors. Gall and Spurzheim were paid to *read* the skulls of politicians and famous people in general. Even marriages among the rich and famous depended on a phrenological analysis of the couple compatibility. In no time, *Phrenology* shook the reliability of neurology as a scientific field and the scientists in Europe and in States started to break away from it, despite the fact that the idea of the brain possessing several organs was already an integral part of the scientific thought about man.

In Vermont, USA, in the summer of 1848, a railroad construction foreman, named Phineas Gage, suffered a serious accident. While doing a routine procedure to blow up a rock internally, Gage got distracted and caused a metal rod, supposed to puncture the rock, to enter his face from under his left cheek. It traversed the base of the skull at high speed, the front of his brain, and exited through the top of his cranium. From his skull, the rod was launched in the air, still at high speed, and fell about 200 feet away from Gage. To everyone's surprise, Gage did not die from the accident. Helped by his

coworkers, he stood up and went to see a doctor, John Harlow, who followed him for years, until Gage's death in 1861.

Despite the fact that Gage did not exhibit the most common post-trauma complications – he was not paralyzed nor lost his speech or memory – he was not himself anymore. He could not focus on anything, developed a drinking problem, could not make plans about his future and did not tolerate socialization. Conclusion: the accident only injured brain tissue that was specialized in mental disposition and planning. This concrete display of selective brain function, plastered all over the papers, reinforced the phrenologist claims in America for some time longer.

3.5 At last, anatomy and cognition go together

In Paris, Pierre Broca, a celebrated physician, was facing an impasse. He intuitively believed in localization and practiced cranioscopy himself, but did not want to get involved in the pseudo-science stigma of phrenology. But on April 12, 1861, in the same year that Phineas Gage died in the US, Broca was called to see a patient that had been transported to the general infirmary of Bicêtre Hospital. It was 51-year-old hemiplegic man, named Leborgne.

Because of his chronic and progressive epilepsy, he had been an outpatient at Bicêtre for thirty years. But he was admitted at that particular time because of a serious gangrene in his entire right inferior limb. When Broca asked him about his condition, Leborgne seemed to understand language, but all he could speak was 'tan', repeatedly, probably with preserved prosodic contour. As an experienced physician, Broca had seen many patients with aphemia, that is, generalized loss of speech. But that kind of aphemia intrigued him.

Six days later, Leborgne died of infection and Broca obtained authorization to autopsy his brain. During the procedure, Broca discovered a softening on a particular region, the posterior part of the left frontal lobe, at the third circumvolution (fig. 14).

After Leborgne, Broca followed eight patients with hemiplegia of the right side. They could not speak, but, like Leborgne, could understand language. When autopsied, all of them showed some kind of tissue injury in the third circumvolution of the left frontal lobe.

These cases led Broca to conclude that the integrity of this area was necessary for the articulation of speech. (Broca, 1861) He was then the first researcher to present a scientific anatomical evidence of a particular brain function – in this case, language production – to have a specific localization. The affected area is the one known as Broca's area.

In summary, there are three major ideas advanced by Broca: (i) language articulation lies the third frontal convolution of the inferior frontal gyrus; (ii) there is left hemisphere dominance in language articulation; (iii) understanding language is a different cognitive task than producing it.

The dysfunction studied by Broca, now named Broca's aphasia or disfluent aphasia or agrammatic aphasia, has been characterized as patients' inability to produce grammatical utterances. Their speech lacks closed class words (prepositions, articles, conjunctions and pronouns) and tends to be very slow and repetitive.

In 1874, in Germany, a young neurologist, Carl Wernicke (1848-1905), assisted two patients also affected by hemiplegia of the right side. Differently from Broca's patients, they spoke profusely, but senselessly: their speech was a stream of grammatical markers, pronouns, prepositions, articles and auxiliaries. In addition, the most salient deficiency in Wernicke's patients was that they did not seem to understand what was said to them.

Eventually, when they passed away, and Wernicke autopsied their brains, he found that the affected region was located in the left temporal lobe, posterior to the primary auditory cortex (fig.15). Correlating lesion with the linguistic output of his patients, he concluded that that area was responsible for the storage of what he called *sound images* (*Klangbilder*). And he believed that these *images* were necessary for the understanding of spoken words after they were heard.

3.6 A cognitive model for the Language Faculty

Wernicke went further in his discoveries. Combining his findings to those of Broca's, he proposed a model – Connectionism – for the cognition of language, a simple model but with great predictive power, published in the historical paper *The Aphasic Symptom Complex* (1874).

In short, his model connected the Broca and Wernicke's Areas by means of a large fibrous pathway, named arcuate fasciculus. With this, Wernicke implied that language processing is distributed in the brain, which is the central idea of most current cognitive models. He also made the correct prediction that patients with damage in the arcuate fasciculus would not be able to repeat speech sounds, a dysfunction later named as Conduction aphasia.

In 1885, Wernicke's model was specialized by a language-processing scheme introduced by the German physician Ludwig Lichtheim (1845-1928). In his model, known as Lichtheim's House, the basic elements are: A – center for auditory images, identified by Wernicke; M – center for motor images, identified by Broca; and B – center for *concept storage* (B from the German word *Begriff*, that means concept), introduced by Lichtheim, who understood that, for language input to be processed and for output to be generated, there had to be a connection with the semantic representation in the brain (fig. 16).

In Lichtheim's words, the scheme previews the following processes:

"The reflex arc consists in an afferent branch aA, which transmits the acoustic impressions to A; and an efferent branch Mm, which conducts the impulses from M to the organs of speech; and is completed by the commissure (*arcuate fasciculus*) binding together A and M. When intelligence of the imitated sounds is superimposed, a connection is established between the auditory center A, and the part where concepts are elaborated, B." (Lichtheim, 1885)

To each of the lesions, represented in the scheme by the numbered red lines, corresponds a different type of aphasia:

- 1.lesion in the motor center causes motor aphasia;
- 2.lesion in the acoustic center causes sensory or Wernicke's aphasia;
- 3.lesion in the acoustic-motor link causes conduction aphasia;
- 4.lesion in the semantic-motor link causes transcortical motor aphasia;
- 5.lesion in the motor output pathway causes subcortical motor aphasia;

- 6.lesion in acoustic-semantic link causes transcortical sensory aphasia;
- 7.lesion in the acoustic pathway causes subcortical sensory aphasia;
- B.lesion in the conceptual center causes anomia or semantic aphasia.

Wernicke-Lichtheim's Connectionism survived almost one hundred years and reached the twentieth century without any alteration. Nevertheless, the fact that it was not disputed did not mean that it inspired further investigations in the same line, seeking for the understanding of language in the brain. In fact, Wernicke-Lichtheim's Connectionism fell on the wayside, while the trend was to look for explanations about human cognitions outside of the brain, just like other theories we have examined so far.

With the flourishing of the social sciences and the introduction of Behaviorism (Skinner, 1904–1990) and Constructivism (Piaget, 1896–1980), human behavior, including speech, was viewed as deriving either from reward and punishment conditioning or constructed from the interaction with the environment. For Skinner, the child's mind at birth was a *tabula rasa*. Learning took place through *operant conditioning*, that is the shaping behavior (the actions of animals) brought about by the consequences that follow upon the occurrence of the behavior. Like with other cognitions, language was understood as verbal behavior and it was conditioned in humans by the patterns of reinforcement that children receive from parents for their verbal behavior. (Skinner, 1957)

For Piaget, children constructed domain-general learning processes that apply equally to each and every area of development, be it linguistic or not. That means that, in his theory, any structural operation that is available for vision is also available for acoustic decoding or for language. The assortment of operations is, thus, global and is incremented with the child's age. The different maturational stages the child goes through are somehow constructed taking the earlier phase as the stepping-stone for the next. In the language domain, for instance, Piaget believed that syntax was structured and vocabulary was incremented from exploratory problem solving with toys and interactions with other children and adults. (Piaget, 1969) First, children would learn words to designate their possessions. Then, as they learned social rules of sharing, they would also, by imitating adults, learn the language that made sharing possible. That is how they would go from "Puppy" to "Let's play with the puppy."

3.7 A cognitive revolution on its way

However, an about-face on the understanding of language as a social construct came with Noam Chomsky's definite review of Skinner's book *Verbal Behavior* (Chomsky, 1959), defending that it is virtually impossible for children to acquire language from reward-reinforcement schemes.

Chomsky's most fundamental argument is the Poverty of Stimulus: The spoken language to which children are exposed during their first years of life is too poor and degenerate to account for the development of perfect native language in a few years. Any simple observation reveals that the child knows more than the experience provided. Besides, reward-reinforcement cannot account for the linguistic creativity that children exhibit. The primary language

data to which children are exposed are made up of a finite number of sentences. And yet children understand and produce an infinite number of sentences. Granted these evidences, Chomsky concludes that children would not be able to acquire language so fast and efficiently if they were not genetically endowed with biological structures that aided language acquisition.

Chomsky advances the notion that we are born with a hard-wired *Language Acquisition Device (LAD)*, that grants us with the major principles of language, or language universals, invariable to our species. He also states that, aside from principles, we have parameters available. These consist of a small number of options, mostly binary, that babies must choose from when fine-tuning, or better, hard-wiring, their internal grammars, to resemble the one that is serving them as primary input. Note that, according to this theory, linguistic variation is accounted for in terms of how the values of these parameters are fixed by the children.

For example, while all languages have subjects, and this is a principle, they also have a well known parameter relative to *the kind* of subject they might have. As it can be seen in (1), a language may be like Portuguese, that accepts null subjects (these are known as pro-drop languages), but can also fill the subjects, depending on the level of formality – more formal: null subject; less formal: subject. Other languages like English (2) will not accept null subjects. So, babies listen to the language around them and are able to sense the cues that will tell them which value to set for that parameter. Brazilian babies will set the value for pro-drop parameter, and American babies will set value for the filled subject.

- (1) a. Fui ao médico ontem.
 went (1st p.s.) to the doctor yesterday
- b. Eu fui ao médico ontem.
 I went (1st p.s.) to the doctor yesterday
- (2) a. *Went to the doctor yesterday.
- b. I went to the doctor yesterday.

Another example of parameter is the one concerning the *wh-phrase*. This is a constituent characterized as a question operator. It can be a word, like *what*, or an entire phrase, like *which books*. In the natural languages, this phrase can be pronounced *in situ*, that is, where it is semantically interpreted, adjacent to the verb, or it can be moved to the front of the sentence. In Chinese (3), the *wh-phrase* is *in situ* only; while in Italian it may be moved to the front or may stay *in situ*, depending on emphasis (no emphasis: moved; emphatic: *in situ*). So, when it is moved in Italian, and it is in the majority of times, the point of interpretation is next to the verb, but it is pronounced right in the front of the sentence (4). However, this parameter still has a third possible value. The last one is selected by the German language: *wh-phrase* must be moved (5).

- (3) a. Mario mai-le shen-ma?
Mario bought what
- b. *Shen-ma Mario mai-le?
what Mario bought
- (4) a. Cosa ha comprato Mario?
What has bought Mario
- b. Mario ha comprato cosa?
Mario has bought what
- (5) a. Was hat Mario gekauft ?
What has Mario bought
- b. *Mario hat was gekauft?
Mario has what bought

From the second half of the twentieth century on, these nativist ideas, which were contrary to the dominant thoughts about language, affected the prevailing understanding of the Faculty of Language and of cognition in general and have deeply influenced the work of many scientists in other areas of knowledge, for instance George Miller, Eric Lenneberg, Jerry Fodor, David Marr, and Niels Jerne, respectively in psycholinguistics, biolinguistics, philosophy, vision and immunology.

In 1965, with the scientific environment propitious to instill biolinguistic investigations, Norman Geschwind (1926-1984), a pioneer in behavioral neurology, reintroduced the Wernicke-Lichtheim's Connectionist Model with more anatomic detail and defended that any cognitive output had to be analyzed in terms of explicit hypotheses as to their underlying neural mechanisms (fig. 17). And in fact, this model previews the dynamic interaction among several inputs and outputs, in specific brain areas.

For example, it predicts task distribution among specific processing sites in order for someone to do something as simple as repeating a spoken word. The following areas would be recruited in sequence: (i) primary auditory cortex, (ii) Wernicke's area; (iii) arcuate fasciculus; (iv) Broca's area; and (v) motor cortex. This is the official language model adopted in most cognitive science textbooks until today.

3.8 Secrecies

Now we have arrived in the fast-evolving part of this history: in the sixties, a revival of the connectionist model; in the seventies and eighties, the introduction of non-invasive or minimally invasive assessment techniques in the study of the healthy Faculty of Language. In the nineties, the field already had a large amount of electromagnetic and imaging data about language processing, and the

number of neurolinguistics laboratories in the best universities in the world started to grow exponentially.

But there is a major challenge to cope with: being able to share neurolinguistics findings with those of other related fields, not only to communicate with them, but also to build a common language that may integrate the different areas of investigation in neurolinguistics. Integration is the key to bringing the field to a deeper level of understanding of the nature and functioning of the faculty of language.

As far as communication and data unification go, the areas that make up neurolinguistics have a very poor track record. For complex epistemologic reasons, linguistics for one has never been readily absorbed into other scientific fields. If you do not agree, take a neuroscience or even a medical textbook. In the chapters about vision, the contents deriving from the physics advancements in optics are smoothly unified with the medical ones. If you are studying vision you must know optics! Vision is one aspect of optics, and optics is one aspect of vision. The same goes for acoustics. But unfortunately, linguists have not managed yet to make morphology, syntax and phonology vocabulary items of neuroscience of language.

Thus, since our marketing department is not a quality one, things we have already known for long time in linguistics and in psycholinguistics may take a long time to be integrated in the field. On account of that, we have not yet been able to forget all about the Wernicke-Geschwind model, since in the last ten years much evidence has been supplied that it is not right. Here is just a small list of its discrepancies and inexactnesses:

1. Warburton *et al.* (1996) and Müller *et al.* (1997) report bilateral activation of the Broca's area before word and sentence generating tasks, although there is more activity on the left hemisphere. Then Broca's idea that language was completely lateralized to the left does not hold.
2. Dronkers (1996) shows that there is damage to Wernicke's area without Wernicke's aphasia and there is Wernicke's aphasia without damage to the Wernicke's area. The same goes with Broca's area and Broca's aphasia.
3. In fact, damage to the temporal cortex that includes subcortical damage is the kind of damage most likely to cause lasting aphasia (Dronkers, 1996). Even in the classical case of Leborgne, when the preserved brain was put in a *cat scan* (Signoret *et al.*, 1984), it was possible to see that there was subcortical injury to Leborgne's insula.
4. It is not true that Broca's aphasics have preserved language understanding. There are specific movement operations, just like the one in reversible passives (*The boy was kicked by the girl*), that present a problem for aphasics (Grodzinsky, 1990). Broca's aphasics have difficulty in assigning the correct theta role to the arguments in the passive construction. So, differently from what Broca supposed, they do have some degree of comprehension difficulty.

Thus, because of 1, 2, 3 and 4, it is not possible to consider Wernicke-Geschwind as the standard to predict the neurophysiology of language anymore.

While we struggle to break the secrecy about our discoveries in the scientific world, other models are being constructed, taking inspiration from the compiled evidence brought in by recent research in the imaging and electromagnetic research fields. The 'Asymmetric Sampling in Time' (Poeppel, 2001) is one of such models being recently developed for auditory perception (fig. 18).

Very differently from Wernicke-Geschwind, the *Asymmetric Sampling in Time's* focal point of analysis is much subtler and more primordial. It assesses the elementary steps in the transformation from the input signal (speech, writing, sign) into the mental representation (roughly, words in our heads). It is a timing-based model that assumes that there exist in the brain fundamental temporal *quanta* working asymmetrically on each one of the hemispheres. These are two temporal integration windows, one on the left to analyze subsegmental-sized units within the 25-50 ms window, and another one on the right to analyze syllable-sized units within the 200-300 ms window. Research exploring this and other aspects of language cognition will be the themes of our next classes.

4. The psycholinguistics' legacy (Day 3)

4.1 The competence-performance dichotomy

In yesterday's class we talked about the chomskian revolution since the 1950's (Chomsky, 1957, 1959) and about the biolinguistic course of investigation that has been revealing a great deal about the nature of the Faculty of Language since then. Generative Grammar introduced a dichotomy between competence and performance (Chomsky, 1965). Competence in this framework refers to the knowledge of our language syntax as it is represented in our minds. Performance refers to the way we access these language representations accompanying our use of language.

Psycholinguists have taken upon themselves to study performance. And since there is a gap between competence and performance, a processing engine is theoretically assumed – the parser – to guide the order in which elements of a sentence are processed and the on-line syntactic assumptions we make while we get language input or produce output. These assumptions are automatic, high-speed and relentless. Once you hear language, you cannot help, but try to process what you hear, the same way that when you open your eyes, you just scan your visual field and process vision.

But the case with language is even more complex, because not only can we process input, but we can also produce instantaneous output in parallel, be it in speech or in thought.

4.2 A powerful processor at work

But how is it that we strategically guess the content of on-line linguistic input and process language so perfectly? Let's simulate a situation of verbal interaction and imagine how the parser might work. Suppose you are in a conversation with a friend. You are listening to an utterance and you have to decode it. But differently from a normal conversation, now everything is in slow motion, so that all the operations that zoom in your mind, while you process language, this time will come by slowly enough for you to notice them.

So, let's begin. Your friend's lips start moving, and before you finish hearing *John*, you already know, at the *Joh* part, that it *will* probably be *John* and that it is the subject of the sentence you are processing. So that makes you immediately expect for the next word, that you suppose is a verb. In the sequence, if you hear *an* and you are expecting a verb, you then might guess it is the verb *answer*, a high frequency verb. But then, surprise: you hear *d Mary*. You were frustrated in your assumptions and had to reformulate your analysis: "OK, the subject is *John and Mary*, so *now* comes the verb", you think. Then you hear *walk*. No morphology of the past tense, so you know it is a habitual action: *John and Mary walk*. You may still expect for an adverb of manner to come next: *John and Mary walk fast*; or a prepositional phrase introducing a locative: *John and Mary walk to school*. Context could play a major role in defining what to expect in this case. But, remember: in reality the sentence is coming to your ears as sounds waves, and you must transform this medium into sound representations, so that your brain can work with them. And then you still think about the organization of the structure, analyze the frequency of words and attribute a semantic content to them. So, while you are hearing and processing, it is unlikely that you will access more data from context. You have limited processing resources and you must be very economical. Of course you will resort to context, but not within this time frame.

But then, let's imagine that you will be frustrated one more time. Instead of an adverb, you will hear the phrase *the dog*. Well, that is a costly frustration because you will have to back up and reinterpret the verb *walk*, that is more frequently used in its intransitive form (with no object) than as transitive. Only then, with its new interpretation, can you make any sense of *the dog*.

In this slow motion processing exercise we just did, we stipulated a number of things. One of them is the assumption that language processing is done bit by bit, serially, as we hear or read input. Another possibility might well be that we hear more input, store it and analyze later. This one would save us backtracking time, but would pose a burden on memory, which is another important player in processing. Besides, since we process so fast, we would probably have some processing slack, while we just listen and store. Again working with a transitorily idle processor is much too costly for a system with the performance excellence that we find in language.

Still another possibility is that we make a combination of the two systems, integrating each and every item we hear, making assumptions about what will come next, but sometimes refraining from these assumptions in ambiguous cases, like *walk* intransitive and *walk* transitive. Whenever there is

room for ambiguity, we would process the two or more options in parallel, store them and decide which one to choose after the disambiguating item appears and is integrated in the processing structure.

In our simplified description of processing, another aspect we did not take into account is the prosodic cues. If they are processed with all other sentence data, there would be no ambiguity between *walk* intransitive and *walk* transitive, because the prosodic contour would be different.

Despite the fact that there is a strong debate about which of these models accurately depicts the psychological reality of processing, psycholinguistics has taken major steps in the understanding of the Language Faculty through the principled observation of linguistic behavior like the one we simulated. Scientists in this field develop an assortment of clever experimental techniques that help understanding how the mind really works with language processing. And they do that mostly by interpreting two basic observable dependent variables, elicited by the experiments: *response time* to carefully delimited experimental tasks and the *error rate of response* in relation to these tasks.

4.3 What we can learn from response time and the error rate of response

Imagine you are doing an on-line psycholinguistics experiment right now, structured by protocol P. You are sitting in front of a computer monitor, and some words are flashing in the middle of the screen for 200 ms. After each second word, your job is to press the green button in a button box, if the second word in the pair is really a word; and the red button, if it is a non-word. As you press any of the two buttons a new pair of words will come. The computer is able to record your answer latency (how long it took for you to press the button) and also if you pressed the right button – green or red – in accordance with stimulus.

But how can scientists make theories about what is happening in the mind by analyzing the two variables, response time and error rate of response? The only way that this can be equated is if they think up the experiment protocol to conform to a theory and use the variables to prove or refute this theory.

Let's suppose that you are working with theory T. According to T, one's mental lexicon is a set of abstract items– roots and syntactic features. For a word to be formed, some abstract items are selected and combined in a special way, called merge, to form a structural unit. After the unit is formed, it is given phonological substance and is interpreted semantically. To T this is what is called *word*. It follows from this theory that if two words are made up by the same root, that is, are morphologically related, like *teach-teacher*, they share a special relationship of identity between each other because they have the same internal structure. But if the two words just share some phonological features, like *ban-banner*, they will be *related*, but will not maintain this special *identity* relationship. According to T, the mental lexicon uses computational resources and little memory. That is, it is more similar to a calculator than to an address-book.

But notice that T may not be the only ball in the game. T', for instance, might purport that one's mental lexicon is filled with ready-made words. So, we

can say that for T' the mental lexicon uses little computational resources and a lot of memory. That is, it is closer to an address book than to a calculator.

Now that we know T and what may be predicted from it, we can ask a question to be verified experimentally: Do the words in our mental lexicon have an internal structure and are computed from smaller bits? And we can also make a hypothesis according to T: Words do have internal structure.

4.4 Finding the right protocol

In order to test our hypothesis we need an experimental protocol that is able to distinguish between a mental lexicon that stores atomic items to be combined and another one that has ready-made complex items. Let's resort to a well known psycholinguistics protocol: priming.

Exposure to a word (prime) that bears some kind of relationship with the next (target) is known to influence lexical access of the next. The kinds of relationship usually studied are orthographic, phonological, semantic, and morphological (cf. Table 1). And *priming* is the influence that these relationships may play. More specifically, priming is characterized by an increase in the speed and in accuracy of lexical access as a consequence of a prior exposure to some of the information in the access context. If one word primes the other then it is because the prime produces an excitation bias in the direction of the target and so the target is recognized faster because of this bias.

Table 1

Relationship	Prime	Target
orthographic	itch	latch
phonological	bar	barbed
semantic	idea	notion
morphological	teacher	teach

Given T, a possible experiment to test the *psychological reality* of this special identity relationship could be implemented by protocol P described above, that is, a priming experiment. This sort of experiment consists of using an experimental *corpus* made up of sets of two words, prime and target, and asking a volunteer to recognize if the target of the set is a word or non-word. Granted that the volunteer is invested in the task, for her to be able to press the button to judge, she must have passed through the operations purported by T: merging abstract items, forming a unit that was given phonological substance by lexical insertion and attributing meaning to this unit. These stages culminate with what is known as *lexical access*. Only after lexical access is performed can the volunteer use the word for a number of cognitive operations, including that of word/non-word judging. The task required is useful because response latency can be a measure of lexical access. Nevertheless, we should notice that response latency is still an *indirect* measure, because it includes the time it takes for the volunteer to obtain lexical access *plus* the time that it takes for her to press the

button according to her judgment. Thus, this test measures behavior *related to* lexical access, not lexical access *per se*.

If you present exactly the same word twice, for example, *ocean* (prime) – *ocean* (target), the lexical access time to the target is shorter than to the prime. It is also shorter than the target in the pair *ocean-wood*. Therefore, we can conclude that identity provokes priming.

Thus, if T is correct in its assumption that morphologically related words are technically identical, we can then expect that a prime will facilitate a morphologically related target more than when the prime merely shares some partial formal features with the target, such as phonological segments.

For this experiment to be successful, the choice of the sets of prime/target and their mix are strategic. Let's look at some possibilities:

Series 1: usual / usually
Series 2: box / bought
Series 3: fail / *failly
Series 4: street / *stround
Series 5: same / *twrizrk

Series 1 is formed by pairs of words that have a morphological identity. Pairs like this are essential in your mix because they are the core of the question you asked. Pairs in Series 2 share some phonological features, but not the same internal morphological structure. This series makes an essential contrast with Series 1. If it takes longer for volunteers to access words in Series 2, then you may suppose that pairs in Series 1 do hold a special relationship between the words: the relationship of identity.

Series 3 is, first of all, a demand of the very experimental task. If the volunteer must choose between word/non-word, then you must include non-words in the experiment. However, to make extra profit from this demand, you can also manipulate with non-words and extract more knowledge about lexical access from them. So, in Series 3 you have a non-word target that bears an identity relationship with the prime, and in Series 4 we have one that does not. This is a way to make Series 3 and 4 analogous to Series 1 and 2.

Finally, in Series 5, we have pairs that include a non-pronounceable non-word target. Although Series 5 is not essential in the experiment, you might want to include it, in order to check the lexical access of items that cannot be handled by T.

A reasonable mix for the corpus in this experiment could be 75 tokens for Series 1; 75 for Series 2; 50 for Series 3; 50 for Series 4; and 50 for Series 5. This amounts to 300 tokens: 150 pairs of word-prime/word-target and 150 pairs of word-prime/non-word-target.

4.5 Interpreting the results

We have talked about a theory T, a question that we can ask based on T, and a hypothesis that derives from T. Then we investigated an experimental means to test the hypothesis with a priming experiment: a carefully structured *corpus* with 5 series of prime-target pairs and a specific mix for this *corpus*.

Imagining that we ran the experiment with about 20 volunteers, we now have the values for the two dependent variables: the mean response latency for each series of the pairs and the error rate of response.

If the difference among the mean response latency between Series 1 and 2 is statistically significant, then we can say that we have thought up a relevant question to ask with our experiment: the issue about the internal structure of words. And based on the findings we may evaluate if the assumptions drawn from T were correct or incorrect. Series 3, 4 and 5 will bring accessory arguments to our conclusions.

The error rate of response is primordially a reliability measure. With it, we can check if the volunteer was paying attention to the test. Based on an atypically high error rate of response of a volunteer, we may decide to eliminate her from the experiment. A high error rate in relation to one or several specific tokens may also indicate that there is a problem with the *corpus*.

When we reach the stage in which there are findings to be interpreted, it should be kept in mind that the conclusion that will be drawn lies within the limits of the theory that motivated the experiment.

4.6 The main areas in Psycholinguistics

The observation of behavioral variables in relation to language in an experimental setting has shown the most productive results in attempts to assess the following aspects of the language cognition: spoonerisms or slips of the tongue (Fromkin, 1999); tip-of-the-tongue effects (Gollan and Acenas, 2004); language acquisition (Pietroski and Crain, 2002); memory (Miller, 1956); reading strategies, lexical access (Marslen-Wilson, 1987, Luce et alii, 1990; Norris, 1994), and sentence processing (Frazier and Clifton Jr., 1996). Here we will restrain our comments to the last two fields.

4.6.1 Lexical Access

Lexical access or word recognition is the capacity to process a word that gets to us in spoken or written form. It is the most basic kind of linguistic processing because the word is out of any syntactic context. And yet, it still is a much debated cognitive capacity.

In our previous simulation, we went through the major steps of a lexical access priming experiment and followed a priming protocol to assess this operation. Now we will discuss about the most common theoretical assumptions behind this incredible cognition.

The field of lexical access has drawn so much attention because it displays, in a very simple and clear manner, some of the most awesome features of the language cognition. Any adult speaker has a mental lexicon of an average of 50,000 words. So, in fluent speech, whenever you want to produce or process a word, you have to start by choosing from a set this big, and you do that at the rate of 100 to 200 ms.

In order to glimpse into this capacity, let's do an exercise. Think about the very hypothetical situation that your brain is at rest. No words are activated in your mind. All 50,000 of them are at resting level. Then, you hear or read a

word and you must activate it in your mind immediately, since linguistic processing always works under incredible time pressure.

In order to have lexical access, we transform the auditory or written input that gets to us serially into a mental representation. Then, we pair this representation with one stored in the mind.

Most contemporary psycholinguists, adept to different models of lexical access, are convinced that when we listen to a spoken word or read a written word, in order to access it, we activate not only the target word but also other phonologically similar words. This is an effect known as *multiple activation*.

As you may anticipate, multiple activation becomes a problem for lexical access, because it makes it harder for the target word to be recognized among others that are also activated and compete for attention.

If that much is agreed upon, the exact characteristics of the competition mechanism are still controversial and will be summarized here as they appear in three main models. In the first one, named the Cohort Model (Marslen-Wilson, 1987), each phonological segment of the target word activates possible word matches. For instance, as we hear the word *banana*, we will initially activate all the words that start with [b], like *ban*, *boy*, *brown*, *banana*. This is a large word cohort (group with statistical similarities), because there are many words stored in the mind starting with this segment. Then it will activate all the words that start with [ba], like *ban*, *banana*, *back*. This time, the new cohort will be more reduced and will eliminate some of the words that were competing in the first cohort. The process is reiterated through all phonological segments, until *banana* is the only match left. In this system, the number of competitors slows down the decision stage of recognition.

In the second model, named TRACE (McClelland and Elman, 1986) the basic unit for activation is also the phonological segment that enters the competition. In TRACE all words have a different activation resting level, according to their frequency: a high frequency word has a high activation resting level; a low frequency word has a low activation resting level. When an input is presented, words that are partially consistent with that phonological segment start being activated. As the segments get into the system continuously, and words that have these segments are activated, they start inhibiting the words that do not bear these segments: an effect known as lateral inhibition or peer inhibition. Thus, competition is due to intralevel inhibition that operates between activated lexical entries, while activation of lexical items is due to excitatory input from the phoneme layer. Activation and inhibition take place cyclically affecting a network of words until the activation of a given word (target) rises above a threshold and is recognized. It follows from this model that high frequency words, that share dense neighborhoods, are very inhibited because they suffer the effects of competition among similar sounding words more intensely.

In the third model, named Continuous Activation (Pylkkänen *et al.*, 2001; Vitevitch and Luce, 1998), the basic activation units of auditory perception are the syllables and two different types of sound-matching may occur: one for phonemes at the beginning of the word, which selects *identical* matches, and one in the middle of the word, which selects *similar* matches.

So, upon hearing *teacher* [ti..], all the words with this exact syllable onset like *tea*, *teen*, *team*, *teach* are activated. But then, through exposure to the continuous input, the remaining syllable continues being scanned [it[...], and the first contrast with the target phoneme produces immediate elimination or suppression of those initially activated words. When the continuing input no longer supports them, these unmatched words are suppressed point-blank and achieve lower activation levels than their initial resting state. So, *teach* stays on, while *tea* and *teen* suffer a sudden deactivation. Thus, in this phase of the recognition process, words do not compete. Unmatched phonemes trigger word suppression.

In the next activation phase, as more input comes in from the middle of the target word [...it]δr], other words that share the phonemes start being activated for matching. So, for example, *each* and *reach* get activated, too. But with these words, immediate suppression of activation no longer happens, the reason being that the matching point in the middle of the word does not provide a retroactive view of the beginning part of the target word already scanned. So these words with non-initial matches, but with internal similarity, stay on at different levels of activation and compete for attention. The fact that they are not suppressed slows down the process until the winning match, *teacher*, gets maximally activated and recognition takes place.

A series of experiments based on this last model will be the theme explored in our fifth class. We will see some priming protocols to test this model, utilizing the magnetoencephalographer (MEG) that provides neurophysiological aspects of lexical activation.

By analyzing these three theories, it becomes clear that the field has attained a high level of specificity. It is also clear that the decision to adopt one or another model is critical for the structure of the experiment that will be put together and for the analysis of the findings.

4.6.2 Sentence Processing

Another classical field in psycholinguistics specializes in the analysis of words in context. We have already discussed some sentence processing strategies in the beginning of this class, when we simulated the processing of the sentence *John and Mary walk the dog*. In this overview, we talked about the time pressure imposed by the nature of the language cognition and about the necessity to make anticipatory assumptions in relation to input that must be processed in few milliseconds.

As stated before, there is theoretical dispute in this field as to the kind of strategy that is adopted to integrate input into the structure being formed. In summary, the field has three major approaches: (i) the serial approach that states that the sentence processing strategy is done bit by bit, sequentially, as input comes. All structural decisions are taken in view of a single structure, and mistakes caused by ambiguity are handled by backtracking; (ii) the parallel approach whose basic strategy is to maintain multiple interpretations in parallel until the disambiguating word comes up; and (iii) the interactive approach, that makes a combination of the other two. Here we will pursue the serial approach

in more detail, describing some of its natural consequences and processing strategies linked with the serial prerogative.

Adopting the serial model of processing means that on-line linguistic material is integrated immediately as it arrives. This model presents the most frontal response to the time pressures of language cognition, aside from offering solutions that conform to economy conditions also imposed by the limited processing resources available. On the other hand, a natural consequence of betting on one processing option only is the garden-path effect (Frazier and Fodor, 1978), a processing hesitation that arises whenever new input cannot be integrated in the structure being formed by the listener. We saw this hesitation in the processing of the sentence *John and Mary walk the dog*.

When the garden-path effect was identified by psycholinguists, many experiments were structured to analyze the nature of this hesitation. Data were ingeniously constructed to test the alternatives being pursued by listeners to anticipate incoming language and the time it took listeners to repair previously built structures.

The sentence provoking the strongest garden-path effect I have ever encountered was given to me by a friend, Marcus Maia, a psycholinguist at the Linguistics Department of the Federal University of Rio de Janeiro. When I give this sentence to Brazilian students, they usually get completely lost in the garden-path and are rarely able to make the necessary repair without help. The sentence is:

(6) Um navio brasileiro entrava na baía um navio português.
a ship Brazilian entered / blockades on the bay a ship Portuguese

When speakers of Portuguese hear the verb *entrava*, they think of the past imperfect tense of the high frequency verb *entrar* (to enter), which is a strictly intransitive verb. So, they get completely lost as to what to do with the NP *o navio português* (the Portuguese ship). When explained about the real meaning of the sentence, people usually say they did not even think about *entravar* (to blockade).

But if frequency was a strong intervening variable in (6), in other garden-pathed sentences frequency is not the primordial influence for the effect to happen. Betting on a given structure that turns out to be wrong is usually a stronger cause for the transitory halt in processing. This can be seen in the following example from Pinker (1994).

(7) The horse raced past the barn fell.

After the subject, one expects the verb, and so *raced* is interpreted as one. However, *raced* is part of past participle phrase that is delaying the appearance of the verb *fell*. The garden-path effect is experienced at the end of the sentence when *fell* is left dangling for a while. After the momentary impasse, structural repair is performed, and the sentence is reinterpreted.

An efficient way to measure the psycholinguistic reality of garden-pathed processing is by using a protocol named self-paced reading. In this protocol words flash in the middle of the screen, but they only flash out when the volunteer presses a button. This way, it is possible for psycholinguists to

measure the time the volunteer needed to read each one of the words in the sentence.

But how can psycholinguists characterize the psychological reality of the garden-path? The answer lies in time measurement. Granted that every task performed in processing takes some time, if the listener falls in a garden-path, then she will have to retrace and correct her previous decisions, in order to accommodate the incoming words that don't fit in.

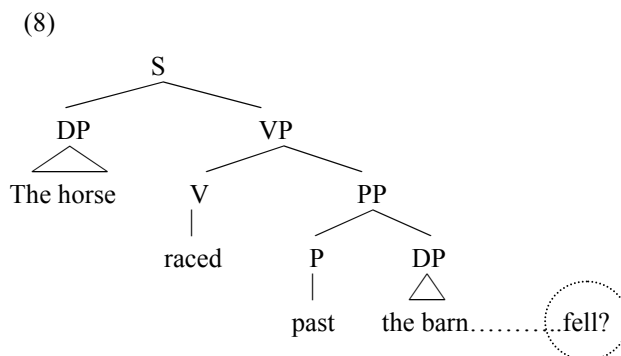
Thus, if (7) were being assessed by such a protocol, the time the volunteer would take to flash out *fell* would be equivalent to the time she spent in the garden-path plus the time that it took her to reanalyze and repair the structure. And this amounts to more time than it usually takes for a verb to be appended to a structure that previews it.

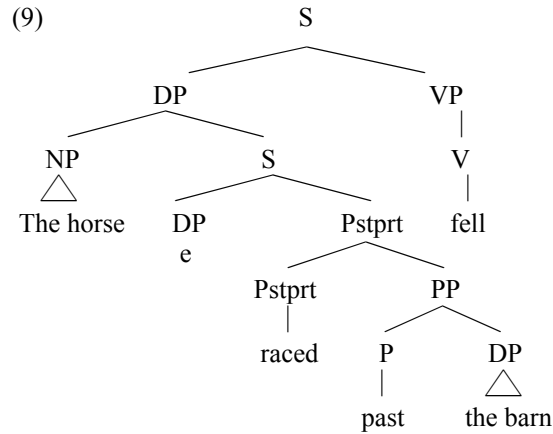
But still, the most interesting point to observe is that if some garden-path sentences, found at random or engineered by psycholinguists, provoke the same interpretation mistake in many listeners, that means that all these listeners are influenced by the same bias in processing. There is an underlying structure to language that is shared by the native speakers of this language. Understanding this bias means unveiling the cognitive strategies that guide sentence processing.

4.6.3 Cognitive strategies: Minimal Attachment and Late Closure

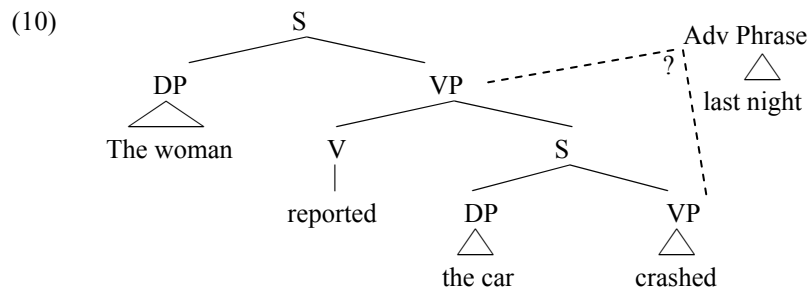
Psycholinguistics research focuses on unveiling the mental processes that control the development and the use of language. Within the serial model of sentence processing, a basic strategy and another one that derives from it were proposed as being language universals: Minimal Attachment and Late Closure (Frazier and Fodor, 1978).

The Minimal Attachment is a strategy guided by the economic requirements of computation of the Faculty of Language. The need to be minimal leads listeners to adopt the analysis that postulates the simplest structure or fewest structural nodes. Looking at (8) and (9), it is easy to see that listeners fall into the garden-path when listening to (7), because they first bet on the analysis that will derive the minimal number of attachments. Then, when (8) proves to be wrong, they hesitate, think and reformulate their analysis to the more complex structure in (9).





While the structure in (8) is clearly a more economical starting bet than that in (9), there are cases in which the number of nodes is the same and still there is room for ambiguity. Let's examine (10).



In (10) the adverbial phrase *last night* may be appended either to the lower or to the higher VP. The case was reported last night or the car crashed last night? The number of nodes will be the same in either case. So, the strategy of Minimal Attachment cannot hold here. To cases like this, Frazier and Fodor (1978) proposed a second-resort universal alternative named Late Closure or Right Association, also guided by a single, least-effort processing tendency. This strategy states that, preferentially, the listener will attach structures for incoming lexical items to structures that have been built more recently. Thus, in (10) *last night* should be appended to the lower clause. In fact, this preference has been verified in many languages.

However, when the ambiguity involves a complex NP, like in (11), the preference for the structure that has been built more recently does not seem to hold universally.

(11) They shot the servant of the actress who was on the balcony.

Who was on the balcony, the servant or the actress? While the late closure strategy seems to be preferred in English (Frazier and Fodor, 1978), there are reports of a preference for appending adjuncts to the higher clause in languages like Spanish, Portuguese, Japanese, and Dutch (Cuetos and Mitchell, 1988, 1996; Maia and Maia, 2004; Kamide and Mitchell, 1997; Mitchell and Brysbaert, 1998). So, for these languages, the interpretation seems to be that *the servant was on the balcony*.

However, those in favor of the universal power of the Late Closure strategy argue that prosodic rules that are processed in parallel with syntactic processing may neutralize the otherwise universal structural preference for low attachment (Fodor, 1998b, 2002).

4.6.4 Getting closer to the real thing

Building on recent developments in linguistics theory, psycholinguistics research has been a crucial tool in the quest to unveil the architecture of the human language. Although psycholinguistics assessments are restricted to the behavioral counterpart of the language faculty and mainly draw on variables such as response latency and error rate of response, the cleverness of the protocols and the fundamental connection with theoretical linguistics leave a mark in several areas of linguistics studies.

Moreover, with the development of the eye-tracking research, a technique capable monitoring eye movement in reading to make inferences about linguistic cognitive processes, psycholinguistics has gained even more ground in the understanding of the mind-brain relationship.

When reading, our eyes stop briefly (from 250 to 500 ms) at individual letters, words or pairs of words. Then they move fast either forward or backward. This jerking eye motion is known as saccades. Studying eye saccades, we may understand the strategies readers use to process text regions that present difficulties. For instance, longer fixation times and repeated visual backtracking may indicate that a reader is having difficulty with a particular word or structure. From this measure, one can infer about the kinds of processing mechanisms and the interpretations readers pursue when they have to resolve ambiguities. So with the development of the eye tracking field, the sentence processing theoretical dispute among serial, parallel and interactive models, seems to be on the way to a resolution, with the direct observation of eye motion while reading.

The eye-tracker and the many traditional psycholinguistics protocols assessing the time course and accuracy of processing have laid out the fundamental leads for modern neurolinguistics research which, on its turn, has propitiated a more direct level of observation of how a person's linguistic capacity is represented in the mind and how these representations can be connected with the genetic endowment. These will be the themes of the following classes.

5. The neurolinguistics research: part 1 (Day 4)

As mentioned in the first class, basic neurolinguistics, that is, the study of how language is represented in the healthy brain, zoomed into the twenty first century after a slow-paced history that used to focus solely on language impairment.

The reason for that staggering progress since the late 1980's was the introduction of non-invasive or minimally invasive neurophysiological assessment techniques that have allowed scientists to take a more direct look at normal brain activation for language processing. When this happened it was possible to start testing specific linguistic phenomena that had been clearly described by linguistics theory and language behavior that had been accurately depicted by psycholinguistics experiments².

Two types of neurophysiological techniques are applied to language assessment: hemodynamic and electromagnetic. As shown in Table (2), the hemodynamic techniques – PET (positron emission tomography) and fMRI (functional magnetic resonance imaging) – have excellent spatial resolution (~1-2mm). Contrastingly, they offer a poor temporal resolution (~1sec), which does not conform to the temporal window of language cognition. Both PET and fMRI are efficient methods in localizing specific brain functions.

The electromagnetic techniques – EEG (electroencephalography) and MEG (magnetoencephalography) – provide excellent temporal resolution, which lies on the order of milliseconds. This is an accountable strength when assessing the brain in the execution of linguistic tasks whose time window is also on the order of milliseconds. However, the EEG does not offer good spatial resolution (~1cm). Even the MEG, which offers improved localization in relation to the EEG, sometimes reaching the resolution of a few millimeters, cannot match the spatial precision of PET and fMRI.

Table (2) – Non-invasive research techniques

Non-invasive recording from human brain (Functional brain imaging)	Hemodynamic c	{	Positron emission tomography (PET)	<u>Excellent spatial resolution (~1-2mm)</u> <u>Poor temporal resolution (~1sec)</u>
			Functional magnetic resonance imaging (fMRI)	
	Electromagnetic c	{	Electro- encephalography (EEG)	<u>Poor spatial resolution (~1cm)</u> <u>Excellent temporal resolution (<1msec)</u>
			Magneto- encephalography (MEG)	

² See Poeppel and Embick (2004) for an important word of caution in the establishment of direct analogies “to validate concepts and categories introduced to the experimental research program by linguistic theory”.

5.1 Hemodynamic techniques

The hemodynamic assessment techniques take advantage from the fact that for each cognitive task to be performed, there is an underlying increase in neural activity. To sustain this increase, there is an extra metabolic demand for glucose and oxygen in the active cerebral region. Since glucose and oxygen are in the blood, there is a rush in the regional cerebral blood flow (rCBF) to the recruited areas to supply the demand for these elements. Both PET and fMRI are able to monitor cognitive activity by tracing such rCBF alterations. After the hemodynamic alteration is traced by these machines, it is reconstructed into appealing images of lit regions in the brain, which are related to the task being performed.

But notice that these rCBF alterations are just hemodynamic *reflections* of the underlying neural activity. So, we can say that PET and fMRI are indirect assessment methods of neural activity (Villringer, 2000). More than that, when these cognitive assessment techniques are applied to language, researchers have to deal with hemodynamic lag, that is, the time it takes for the blood to reach the active areas, in the order of seconds up to one minute. That is not an efficient time window for language tasks that are performed in the order of milliseconds (Poeppel, 1996a, 1996b). Thus, poor temporal resolution presents the most serious challenge to the effective use of hemodynamic techniques to assess language cognition.

The second most serious drawback of hemodynamic techniques is the subtraction-based method. These tests typically produce rCBF maps of contrasting experimental conditions. For instance, one map may be related to a listening task contrasting with another one, used as control, that is produced by a brain scan when no task is requested. The two maps are then subtracted and the result produces a third map, which should depict only the areas involved in the listening task. Among several technical problems with this approach, the most striking one is the assumption that what is taken as the control scan is really depicting a resting baseline of the brain (Hickok, 2003).

5.1.1 Positron Emission Tomography

Despite the fact that both PET and fMRI are sensitive to hemodynamic alterations, they do it in their own specific way. Although the use of PET is allowed in healthy individuals for research purposes, it is still a minimally invasive technique, because it traces isotopic material inserted in the blood

³ This table was transcribed from David Poeppel's slides presented by him during his Neurolinguistics course at the Federal University of Rio de Janeiro in 2001.

stream. Its method consists of injecting the volunteer with a radioactive tracer, usually oxygen-15, while she performs a linguistic task. The brain regions that receive more blood flow also get a greater quantity of this tracer, to which the PET machine is sensitive.

Technically, when the unstable molecule of oxygen-15 decays, its nucleus liberates kinetic energy that expels an anti-matter particle, named positron. The positron travels at light speed away from the nucleus until it collides with one of the billions of electrons inside the body. At the point of collision, named annihilation site, two photons (gamma rays) are emitted and travel in opposite directions until they leave the head. The energy of these photons can be captured by the PET scanner, which has a donut ring lined with photon sensors. From the coordinates and speed of the photons, the machine calculates the exact location of the annihilation site in the brain. An image plotter transforms the localization data into a three-dimensional brain graph, with the involved region lit. PETs are designed to deliver a minimum amount of radiation to the volunteer, often less than would normally be received in an X-Ray exposure.

The most frequent criticisms to PET can be summarized in three points: (i) the decomposition of linguistic tasks into their complexities is rather difficult; (ii) the findings of PET studies may rarely be related with the established description of language phenomena by linguistics theory; and (iii) as pointed out before, the subtraction method is not reliable enough. The same linguistic test performed at different laboratories yield activity in fairly different brain areas (Poeppl, 1996a).

5.1.2 Functional Magnetic Resonance Imaging

The second most common hemodynamic research technique is the fMRI (functional magnetic resonance imaging). It measures changes in blood flow using magnetic resonance. It is a non-isotopic assessment method, and, therefore, it is not invasive to the research volunteer. In one of its most used designs, images are acquired by the blood oxygenation level-dependent (BOLD) protocol, which uses deoxyhemoglobin as a natural isotope.

The fMRI machine takes clever advantage of the brain metabolic changes. When cognitive tasks are performed, oxygen is needed, and it is rushed to the scene by hemoglobin. This oxygenated type of hemoglobin, named oxyhemoglobin, produces a different magnetic resonance form than that of deoxyhemoglobin, which is the hemoglobin after it passes oxygen to the needy tissue of the recruited area. Regions of increased neuronal activity receive more blood, and this blood donates more of its oxygen.

The fMRI machine has a very powerful magnet that surrounds the volunteer with an intense magnetic field. When the volunteer is inside this magnetic field, all of the billions of hydrogen brain water atoms align with it. Inside the fMRI machine, radio frequency pulses are applied to excite some of these hydrogen atoms. When the magnetic field is interrupted, the relaxed atom emits a photon, which can be traced by the fMRI machine. Images are created based on how these excited atoms lose their energy.

The fMRI is capable of tracing atoms in multiple brain slices. It measures ten axial slices of 5 mm in 1.4 seconds. Measurements are time-locked to the presentation of linguistic stimuli. These images connected to linguistic tasks are fed into a computer to generate an incredibly detailed look of the brain in activity (Kim, S.-G. *et al.*, 1999).

The most serious drawback of this technique is that the volunteer has to be completely still during the task, because even the slightest head movement may define a bias to the results.

Another hard challenge that has to be met by researchers using this technique is that despite the fact that metabolic changes can be depicted in the multiple brain slices that the machine can scan, the more slices scanned, the worst resolution they get. So the researcher has to find a reasonable setting between the number of slices and the resolution.

5.2 An example of PET study

Ever since imaging techniques became more amply used, the old field of lesion studies, focusing primarily on research of the different types of aphasia, was revitalized by the more accurate description of several language disorders. But this enhancement to the area was also made possible due to a new research design: testing a given language task in normal volunteers and using results as a term of comparison with those obtained from brain lesion patients who have trouble performing this same task.

Many previously tested language deficits are being tested again using this protocol. Among them is anomia, the loss of access to certain items of the mental lexicon. Everyone of us has experienced some degree of anomia temporarily. Not knowing how to name an object or person often happens when we undergo stress, a sleepless night or a period of exhaustion. Anomia is also more common in old age. However, there is an acute form of anomia, called pure anomia, that is connected with brain injury to the temporal lobe. Patients with this kind of lesion are not able to name specific things, although they continue accessing the concept behind that missing word. For instance, the researcher shows a picture of a skunk to a patient with pure anomia. The patient might not be able to say "It's a skunk." But she may say that it is a furry wild animal, that sprays secretion with bad smell as self-defense.

Hanna Damasio and colleagues (1996) noticed that some patients with this condition had pretty selective losses: some had trouble naming tools, others could not name animals, and still others could not name celebrities. So, she set up a PET experiment with a group of 29 temporal lesion patients and a control group with 9 healthy individuals. All subjects were stimulated by 327 pictures: 109 illustrations of celebrities, 109 of animals and 109 of tools. In case the subjects could not name the illustration, they were asked to come up with a definition. In case they could not, that token was excluded from the study. During stimulation, subjects' brains were monitored by a pet scanner that drew rCBF maps of subjects' neural activity.

By the design of the experiment, we can infer that the authors' hypothesis was that there is cortical specificity correlated with type of information. In fact, the rCBF maps of the neural activity of the control group (healthy individuals) showed that the area recruited for celebrity naming was the

anterior part of the left temporal lobe, named the temporal pole. The area recruited for naming animals was the inferotemporal region, and the one for tools was more posterior: the posterior part of the inferotemporal lobe, extending to the anterior part of the left occipital region. This preliminary result from the naming task performed with the control group guided the experiment done with the brain lesion patients. The results are summarized in Table (3).

Table (3): Percentage of right answers per group of illustration and lesion localization

Picture groups	Temporal pole		Inferotemporal region		Posterior part of the inferotemporal lobe + anterior part of the left occipital region	
	Healthy subjects	Lesion patients	Healthy subjects	Lesion patients	Healthy subjects	Lesion patients
celebrities	X	59.8		75.5		91.7
animals		93.3	X	80.1		88.3
tools		96.0		84.5	X	78.5

Patients who had an anterior temporal pole lesion had more difficulty in naming celebrities (only 59.8% of right answers). They did very well in naming animals and tools (close to 100%). They were also the subjects with greater difficulty in naming celebrities when compared with the other subjects with more posterior lesions.

Notice that patients with inferotemporal region also had more difficulty in naming celebrities. So the correlation with the control group, in this case, could not be made: normal subjects had the inferotemporal region activated when they named animals, while patients with lesion in inferotemporal region still scored worse in the naming of celebrities. Besides, the naming difference in the three types of information groups is not much different from one another.

Patients with lesions in the posterior part of the inferotemporal lobe extending to the anterior part of the left occipital region have more problems in naming tools. This group presents a match with the activated area in normal individuals.

The study's conclusion is that there is a tendency for stored lexical items to be organized in the temporal lobe according to the type of information. Results are also indicative that persons, tools and animals are relevant parameters of type of information.

6. The neurolinguistics research: part 2 (Day 5)

6.1 Electromagnetic techniques

Each and every cognitive task the human being performs – seeing an object, recognizing a sound, articulating a phoneme, sensing a smell – has an electrochemical basis. Cognition takes place through the cortical processing of electromagnetic bits of information that flow along neurons in the brain, a process that is triggered and maintained by biochemical reactions (cf. 2.2.4).

The electromagnetic assessment techniques take advantage from the fact that this electromagnetic activity can be sensed at the scalp by the EEG (electroencephalographer), already used since 1920's, and by the MEG (magnetoencephalographer), a very sophisticated machine made available since the 1990's. Both EEG and MEG are non-invasive brain activity assessment tools that can provide incredible temporal precision, on the order of the millisecond time-scale resolution. Thus, these techniques are considered as a direct measure of cognitive activity.

6.1.1 Electroencephalogram (EEG) and event-related brain potentials (ERPs)

EEG signals are obtained through electrodes placed on the scalp over multiple areas of the brain. Due to the electrode's dimension and its location over the scalp, each one detects the bioelectricity that stems from a neuronal population, which can reach millions of neurons. In fact, the EEG performs the recording of the temporal and spatial summation of all excitatory and inhibitory post-synaptic potentials in the brain along the time (Lopes da Silva, 1999).

However, since the bioelectricity in the cortex is propagated not only through neurons but also throughout the brain, it encounters propagation barriers, such as non-neuronal tissues (meninges and bone), which absorb electricity and diminish the strength and deviate the signal. Thus, when the wave reaches the scalp, the electrodes acquire an attenuated signal that has to be amplified in order to be studied.

A negative point about this technique is that it has little spatial precision, on the order of centimeters, so it is impossible to correlate the position of the electrode on the scalp to the precise spot in the brain where the signal was originated.

Throughout its history, the EEG has been more commonly used in the diagnosis of diseases such as seizure disorders, mental confusion, and in the evaluation of head injuries, tumors, infections, degenerative diseases, and metabolic disturbances that affect the brain. It is also used to evaluate sleep disorders and to investigate periods of unconsciousness or to confirm brain death in a comatose patient. In fact, conditions such as epilepsy, attention deficit disorder, dyslexia, sleep disorders and many others can be accurately characterized in terms of the bioelectric wave patterns displayed by the EEG (Bear, Connors, Paradiso, 2001).

Since the last decade, however, the field of neurolinguistics has also been making wide use of the EEG to assess the electrical cortical effects caused by different sorts of linguistic phenomena in normal people. Such studies apply a

signal amplification and averaging technique to extract electrical brain components known as ERPs (Event Related Brain Potentials).

The averaging protocol has accountable advantages over the subtraction-based method of PET and fMRI, because it does not have to consider a resting baseline of the brain. On the contrary, it only works with brain activity. It adds multiple occurrences of the signal (electrical data) in relation to a target phenomenon. For instance, let's suppose the volunteer is stimulated with 80 sentences with the same syntactic structure. Each sentence has four words, each appearing on the computer screen for 200 ms. In 50% of the tokens these sentences will yield an incongruous reading provoked by the transition between the verb and the object (verb-complement merge). The remaining 50% of sentences will yield congruous readings, as can be seen below in Table (4).

Table (4)

Congruous

1. He crashed the car.

2. She emptied the box.

....

40. They started the work.

Incongruous

41. She ate the socks.

42. I opened the sneeze.

....

80. We rubbed the air.

All 40 waves, picked by each electrode placed on the scalp, resulting from the verb-complement merge of the congruous sentences will be averaged, and so will those of the incongruous sentences. If the two resulting averaged wave forms, one for congruous and one for incongruous sentences, are different from each other, then we can say that verb-complement incongruity causes a recognizable electric effect.

Another advantage of this method is that it can eliminate an unavoidable experimental problem. If each and every cognitive task the human being performs produces an electric response, then if the volunteer scratches the hand, adjusts to the chair, yawns or blinks during the EEG recording, exactly at the same time that she is processing the verb-complement merge, there will be an electric component resulting from this action that will be contaminating the recording of the electrical activity of the target stimulation. But since this intervening component, known as artifact, appears at random, not repeatedly associated with all the tokens of the target event, it is cancelled by the linear summation and averaging of the multiple signals. So the averaged ERP wave is free from artifact and can be analyzed in terms of its main parameters: a specific latency (the onset time of the peak), amplitude (how high the peak is) and topographic distribution (regions on the scalp where it is elicited).

For instance, taking the experiment we described above, once a volunteer is stimulated with a sentence containing a semantic violation between the verb and its complement, a special electrical event takes place in her brain (fig 19). This event is a wave of negative polarity which peaks at around 400 ms post-stimulus (Kutas, Hillyard, 1980). This electrical component, well known in the literature as the N400 (N standing for negative polarity and 400 for the time it takes for the wave to raise since its stimulus onset), is widely described as a

consequence of increased difficulty in morpho-syntactic integration due to the semantic anomaly (Osterhout, Holcolmb, 1993; Osterhout, Holcolmb, 1995).

In the N400 context, three aspects of the resulting ERP are usually analyzed: latency, amplitude and topographic distribution (Kotz and Friederici, 2003). The latency of the cortical response is referenced to a specific instant in time – for example, the target stimuli presentation – and usually ranges from 300 to 500 ms. Amplitude is commonly related to the level of facility to perform morpho-syntactic integration (Kutas and Hillyard, 1984; Fonteneau *et al.*, 1998).

Thus, it can also be seen as the inverse function of context, i.e., the least supporting context for semantic satisfaction results in the largest amplitude of the waveform, as a direct result of the integration challenge (Holcomb and Neville, 1991).

Note that the N400 is robustly related with verb-complement merge effect. It does not relate to language-irrelevant attribute of words, such as mistakes in font type, size or color, and it does not relate with grammatical incongruities. Nevertheless, the verb-complement merge is not an indivisible operation. With the ERP extraction and the unveiling of the neurophysiology of this operation it is possible to observe subtle differences among different types of merge.

6.1.2 An example of an event-related brain potential (ERP) study

In order to test the ERP protocol capacity to discriminate among a repertoire of sub-tasks involved in the different kinds of merge operations, França et al. (2004) presents a study in which the standard N400 experiment is diversified by testing less commonly explored contexts of verb-complement integration. Assuming that different types of syntactic contexts surrounding the merge operation trigger specific neuronal activation modes, this study compares the following three series of sentential stimuli that are derived by verb-complement merges embedded in different syntactic configurations.

Series 1 gathered sentences involving the most studied type of verb-argument combination: the local one. The verb finds its argument right beside it. Incongruence is established by the local incompatibility between the selectional requirements of the verb and the semantic properties of its complement.

Series 2 stimuli establish a local relationship between the verb and a pronominal complement, so that, in this case, the semantic properties of the complement are inherited from a distant nominal antecedent (Binding Theory Principle B. Cf. Chomsky, 1981).

In Series 3, instances of verb selection involve syntactic structures with WH-movement. In such constructions, the complement (WH-phrase) appears displaced from its interpretation site. As a complement, it is interpreted right beside the verb, but it is realized phonetically in a sentence initial position, formed at a later merge, in spite of being pronounced first.

Examples of the three series follow:

Table (5) The three series

Series	Congruous Sentences	Incongruous Sentences
--------	---------------------	-----------------------

1.	Ela leu o livro <i>she read the book</i>	Ele comeu os sapatos <i>he ate the shoes</i>
2.	Lúcia ia lavar a louça e depois secá-la <i>Lucia was going to wash the dishes and then dry them.</i>	João está segurando a pizza e quer lê-la <i>John is holding the pizza and wants to read it</i>
3.	Quais músicas Vera compôs? <i>what songs Vera composed</i>	Que histórias Denise temperou? <i>what stories Denise seasoned</i>

The averaged ERPs resulting from the stimuli from series 1-3 are shown in figure (20).

The simple visual inspection of the averaged waves presented above reveals so many morphological wave differences that we are led to suppose that the verb-complement merge operation is not an indivisible neurophysiological process. And, in fact, linguistic theory predicts decomposition of the verb-complement merge operation into different sub-modules, when deriving the sentences in each one of the three series.

For instance, while in Series 1 the tasks are: (i) combining the verb semantic selectional criteria with the semantic properties of the complement; and (ii) assigning a theta role to the complement; in Series 2, after (i) we still have to (ii) establish co-reference; (iii) retrieve semantic properties from the previous sentence; and finally (iv) transfer semantic properties from the antecedent to the pronoun. In Series 3, the non correspondence between the interpretation site filled with an empty category and the pronunciation site of the Wh-expression requires reconstructing semantic features of the empty category from the distant Wh-phrase.

Thus, the different wave morphologies led us to launch the possibility that the length of the slope prior to the peak, the amplitude of the peak, its latency, its shape and the slope after it are important parameters related to urgency of the merge, integration facility, activation time, complexity of the cognitive tasks and readiness to move on to other tasks, respectively.

Taking these parameters into consideration our findings can be better visualized in Table (6)

Table (6) – Table of findings in França *et al.* (2004)

Parameters	Series		
	Série 1	Série 2	Série 3
Latency: speed of peak formation	●●	●	●●●
Motor response speed (behavioral parameter: finger pressing)	●	●●	●●●
Amplitude: height of the peak	●●	●	●●●
Merging urgency: Duration and inclination	●●●	●●	●

of the line prior to the peak			
Topographic diffusion of the electric activity	Distribution:	Cortical	• •• •••
Complexity of the Openness of the peak angle	cognitive tasks:		• ••• ••
Readiness to move on to other tasks: post peak descending speed.			••• • ••
Error response rate (behavioral parameter)			••• •• •
Table key			
•	••	•••	
lowest	medium	greatest	

Our conclusion was that the data extracted from this experiment indicate that there are different N400 morphologies related to the different types of merge and their associated morpho-syntactic components. This finding is consistent with premises from current neurophysiology of language models that view language processing as resulting from the interaction of task-specific computations, plausibly involving different neurological subsystems and neuronal firing modes (Hickok, Poeppel, 2000, 2004; Poeppel & Marantz, 2000; Hickok, 2001).

One can also conclude that the experimental protocol was successful to reveal electrophysiological data that mapped linguistic specificities onto different wave morphologies derived at different cortical regions.

Besides the N400, there are other well known linguistic ERPs: (i) ELAN (early left anterior negativity), that relates to the early (between 125 to 180 ms) perception that there is something wrong with word class choice, for instance *John bought smiled*; (ii) P600, that is connected with bad syntactic formation (for example, *They rans*); and (iii) LPC (late positive component), a late positivity (between 500 and 800 ms) that relates to the irregular word formation of morphologically irregular words (for instance, *fastly, cutted*).

6.1.3 Magnetoencephalography (MEG)

MEG is an ultra sensitive cognitive assessment device that captures small electrical currents arising inside the neurons of the brain. These currents produce minute magnetic fields that pass unaffected through brain tissue and the skull, so they can be recorded outside the head.

It is a completely noninvasive technology for functional brain assessment, localizing and characterizing the electrical activity of the central nervous system by measuring the associated magnetic fields emanating from the brain. MEG provides an excellent temporal resolution on the order of 1 ms and still good spatial discrimination that may reach a few millimeters.

This technique takes advantage from the fact that every electric current generates a magnetic field according to the *right-hand rule of physics*⁴. This same principle is applied in the nervous system whereby the longitudinal neuronal current flow generates an associated magnetic field. MEG measures these intercellular currents of the neurons in the brain giving a direct information on the brains activity related to a given stimulus.

Despite the fact that the magnetic fields are very small, they can still be detected by a sophisticated technology based on superconducting detectors and amplifiers known as SQUID (superconducting quantum interference device). To put it simply, a SQUID can be thought of as a very low noise device for transforming magnetic fields into electrical voltage. A SQUID acts as a low-noise, high-gain, current-to-voltage converter that provides the system with sufficient sensitivity to detect neuromagnetic signals. SQUID extracts data that can be used to calculate the location and strength of the activity within the brain.

To some degree, MEG is similar to EEG, except that the skull and the tissue surrounding the brain affect magnetic fields much less than they affect the electrical signals measured by EEG. The advantage of MEG over EEG is therefore greater accuracy due to its minimal signal distortion. This allows for more usable and reliable localization of brain function. MEG is being used in neurolinguistics for ERP analysis and the resulting wave is named by an introductory M (from magnetic) followed by the latency of the magnetic component. For instance, exposure to a word time-locked to an MEG recording yields three magnetic components: M170, M250 and M350, the latter usually viewed as the magnetic analogous to the N400. Since MEG is still a fairly new research implementation, the exact functional role of these components is still being studied.

The greater disadvantage of MEG is that it is still too costly a machine and it requires an extremely complex installation (among other things, it has to function inside a magnetically shielded room) to be widely used in linguistics research. Most laboratories that have an MEG facility devote it to medical purposes.

6.1.4 The Neurolinguistic Investigation

Perhaps the main point of this Introductory Course to Neurolinguistics is to enforce the fact that these incredibly exciting non-invasive assessment methods applied to the cognition of language will only be used to improve our

⁴ The direction of the magnetic field can be discovered by making a comparison with the direction of the torque vector that is known by the application of the right-hand rule. Imagine you are pushing a door to open it. The force of your push (F) causes the door to rotate about its hinges (the pivot point). How hard you need to push depends on the distance you are from the hinges (r). The closer you are to the hinges the harder it is to push. Using the *right hand rule*, we can find the direction of the torque vector. If we put our fingers in the direction of r, and curl them to the direction of F, then the thumb points in the direction of the torque vector. In electromagnetic terms, if we put our fingers in the direction of the electrical signal, and curl them, then the thumb will point to the direction of the magnetic field.

knowledge of the Faculty of Language in humans if we can connect protocols and findings to the results of linguistics theory. More than that, the ideal situation to which we should be directed is that the interdisciplinary efforts, among medicine, engineering, linguistics and psycholinguistics, find a common denomination to the related facts in each of these fields. Although ideal is still far from reality, a close connection to linguistics theory to guide all neurolinguistic pursuits is minimally necessary.

In this last chapter⁵, we will reanalyze a PET research already introduced in 5.2 (Damasio et alii, 1996), performed within the medical/neuroscience framework. Then we will check out a different perspective to deal with the problems presented, more in tune with linguistics/psycholinguistics literature.

In Damasio et alii (1996) a naming task with words from three information categories was applied to patients diagnosed with a language dysfunction – pure anomia - and to a small control group of healthy individuals. Although the results (rCBF maps of the neural activity of the two groups) are interesting and showed a percentage correlation between lesion and localization of selective information stored items (cf.5.2), they shed no light on the architecture of the Language Faculty. There are no findings coming from the linguistics and psycholinguistics literature that relate a special lexical organization of types of information discriminated into the groups proposed by the authors.

If we were to understand pure anomia within the linguistics framework, the problem would be handled by a much studied facet: lexical activation. Not being able to retrieve the name for something must be related to one or more than one sub-tasks involved in lexical activation. Thus, if we were to glimpse into the neurophysiology of lexical activation through the eyes of neurolinguistics, the starting point would be to identify its biomeasure using a non-invasive assessment method. This has, in fact, been attempted by a series of clever experiments using the MEG facility performed by Alec Marantz and colleagues at the KIT/MIT MEG Joint Research Lab.

Based on the fact that whenever a volunteer is stimulated by a word, three MEG components appear –M170, M250 and M350 - the intuition was that one of them could be related to lexical activation. Logically, it should be the one with the longest latency, the M350, due to the fact that lexical access is the final product of a sequence of cognitive tasks that go from transforming the auditory or written input that gets to us serially into a mental representation, activating multiple representations that somehow relate to it, filtering the matches that are not identical and finally pairing the winning representation with content stored in the mind. In reality, the M350 was also a good candidate because the other two earlier components had been respectively related to the perception of the visual form of the word and to phonological processing. The third component – M350 – was then hypothesized as being related to lexical access.

To find out if the intuition was right, a connection with robust findings from linguistics/psycholinguistics theory was considered essential. Thus, for the M350 to be an index of lexical activation it should:

⁵ The format of this comparison and the papers it encompasses was proposed by David Poeppel during his 2001 lectures in the Federal University of Rio de Janeiro (UFRJ).

- Vary with stimulus properties that are known to affect lexical access (word frequency, priming, recency and phonotactics) and be the earliest response component affected by these properties
- Vary independently of reaction time in tasks where lexical activation contrasts with lexical decision
- Be affected by both phonological and morphologic factors
(List extracted from David Poeppel's 2001 lecture at UFRJ)

In order to check each one of these items, a series of studies were performed by the mentioned research group. In the first one (Embick et alii, 2001), lexical frequency was varied in a lexical decision task (Is it a word or a non-word?). Volunteers were visually stimulated by words from six groups with systematic variation of frequency (from the 64th to 69th frequency group as classification by the Cobuild frequency corpus). The stimuli appeared randomly and were mixed with two types of non-words (pronounceable and non-pronounceable).

The results showed that the M350 was sensitive to frequency: its latency increases as frequency decreases. The other two MEG components (M170 and M250) were unaffected by frequency. So the M350 proved to vary with one of the stimulus properties that are known to affect lexical access: frequency.

The second experiment (Pylkkänen, L, 2002a) made use of a lexical activation theory presented in 4.6.1: the Continuous Activation Model (Pylkkänen et alii, 2001; Vitevitch and Luce, 1998). In it, the basic activation units of auditory perception are the syllables and two different types of sound-matching may occur: one for phonemes at the beginning of the word, which selects *identical* matches, and one in middle of the word, which selects *similar* matches. Here again the intuition was that the simple lexical activation task is not indivisible. Taking that framework presented in 4.6.1., there is the moment that the linguistic input reaches us, the moment to transform visual or acoustic input into a mental representation to be paired. This tentative of pairing activates multiple lexical candidates among the 50,000 representations of linguistic words that compete for attention. Then there is the gradual elimination of imperfect matches and, finally, recognition of the right match. In this experiment, Liina Pylkkänen and colleagues aimed at finding neurophysiological discrimination between the moment of continuous activation and the final moment of recognition. To test this dissociation, they strategically controlled a special variable: phonotactics. Phonotactics is the set of allowed arrangements or sequences of speech sounds or written letters in a given language. A word beginning with the consonant cluster (zv), for example, violates the phonotactics of English and Portuguese, but not that of Russian (example, zv'er'- *beast*). Words with an uncommon phonotactics will make the multiple activation task slower, since it is harder to find the few words, among the 50,000 words in one's mind, that will have that rare phonotactics. However, after these words are activated and start competing for attention, the fact that they are few accelerates lexical recognition.

In this study, two groups of non-words were compared to words in a lexical decision task. A group of non-words had phonotactically common words

like *mide* and the other presented phonotactically uncommon forms like *yush*. The finding suggests that phonotactic probability really affects activation and selection in opposite ways. While phonotactically uncommon forms took longer to be activated, once the possible matches were found and started competing for attention, decision was faster because the selected group was potentially smaller than the group that had phonotactically common words. Whereas when the non-word was phonotactically frequent, like *mide*, it activated faster a much larger pool of candidates and creates a dense neighborhood that delays decision. These findings made it possible to associate M350 with the initial word activation and to dissociate it from reaction times. Thus, the M350 latencies were found to be decreased, rather than increased, for high probability stimuli. This shows that while the M350 is sensitive to lexical factors, it is not affected by competition.

Finally, the third experiment (Pylkkänen, L, 2001) - to test if the M350 was affected by both phonological and morphologic factors, was a multi-modal priming experiment in which the prime was auditory and the target was written. There were four tested conditions:

- Phonological correspondence at word onset: *spinach*-SPIN
- Phonological correspondence in the middle: *teacher*-REACH
- Semantic Prime: *idea*-NOTION
- Morphological correspondence: *teacher*-TEACH

The results show that the M350 had the fastest latency when the stimuli were morphologically related. It was also affected positively by semantic relatedness. Another important finding was that two distinct phases of word recognition were unveiled: recognition by the onset and recognition by the middle (cf. 4.3)

Thus, after these three experiments the group concluded that the M350 can be a reliable index of lexical activation, which is the first subtask involved in lexical access.

7. Conclusion: The neurolinguistic advantage

Data from the medical literature about the direct correlation between the hemodynamic effect of pure anomia and the organization of the lexicon (Damasio et alii, 1996) could not reveal all the cognitive details characterizing the impaired cognition of patients who suffer from it.

However, a change in methodological perspective, invoking a neurolinguistics protocol, including a specific linguistic theory of the organization of the lexicon, brought light into the basic cognitive tasks of lexical access, which should now lay the fundamental ground for the neurophysiological study of pure anomia.

Neurolinguistics' scope is much more specifically defined than that of general cognitive sciences because it is based on linguistics theory. Only after thorough studies of linguistic phenomena, inspired by theoretical findings in linguistics, are performed can one make sense of linguistic dysfunctions.

As a matter of fact, after these discoveries about the multiple tasks involved in lexical access, and about the segregation between activation and suppression tasks, an important contribution has been given to the field of

aphasia: while Broca aphasia is connected with an activation problem, Wernicke's aphasia seems to be connected with suppression (Pylkkänen, L, 2002). This is a sample of the neurolinguistic advantage.

References:

- Andersen, Birgitte B.; Lise Korbo and Bente Pakkenberg (1992). A quantitative study of the human cerebellum with unbiased stereological techniques. *Journal of Comparative Neurology* 326(4): 549–60.
- Aristotle (350 b.C). *On the motion of animals*. Translated by A. S. L. Farquharson. The Internet Classic Archive, MIT. Available at http://classics.mit.edu/Aristotle/motion_animals.html
- Bear, Mark; Barry Connors and Michael Paradiso (2001). *Neuroscience: exploring the brain*. Baltimore: Lippencott Williams & Wilkins. 666 pp.
- Broca, Paul P. (1861) *Remarks on the seat of the faculty of articulate language followed by an observation of aphemia*. Reprinted in: *Some Papers on the Cerebral Cortex*. Translated by G. von Bonin. Springfield: Thomas, 1960, pp. 49-72.
- Chomsky, Noam (1957). *Syntactic structures*. The Hague: Mouton. 117 pp.
- Chomsky, Noam (1959). A review of BF Skinner's verbal behavior. *Language* 35(1): 26–58.
- Chomsky, Noam (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press. 261 pp.
- Chomsky, Noam (1993). A minimalist program for linguistic theory. In Hale, Kenneth and Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA.: MIT Press, pp. 1–52.
- Cuetos, Fernando and Don C. Mitchell (1988). Cross-linguistic differences in parsing: restrictions on the use of the late closure strategy in Spanish. *Cognition* 30: 73–103.
- Cuetos, Fernando; Don C. Mitchell and Martin Corley (1996). Parsing in different languages. In Carreiras, Manuel; Jose E. García-Albea and Nuria Sebastián-Gellés (eds.), *Language processing in Spanish*. Mahwah, New Jersey: Erlbaum, pp 145–87.

- Damasio, Hanna; Thomas J. Grabowski, Daniel Tranel, Richard Hichwa, and Antonio R. Damasio (1996). A neural basis for lexical retrieval. *Nature*, 380: 499-505.
- Dronkers, Nina F. (1996) A new brain region for coordinating speech articulation. *Nature*, 384: 159–161.
- Embick, David; Martin Hackl, Jeannette Schaeffer, Melltem Kelepir, and Alec Marantz (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research*, 10(3): 345–8.
- Fodor, Janet D. (1998a) Learning to parse? *Journal of Psycholinguistic Research* 27: 285–319.
- Fodor, Janet D. (1998b) *Parsing and prosody: the Late Closure crisis*. City University of New York (ms). Available at http://www.ias.or.jp/research/res_gengo/gengo98rpt/Foder.html
- Fodor, Janet D. (2002) Prosodic disambiguation in silent reading. In Hirotani, Masako (ed.), *Proceedings of NELS 32*, Amherst, MA: GLSA. University of Massachusetts, pp. 113–32.
- Fonteneau, Elisabeth; Uli H. Frauenfelder and Luigi Rizzi (1998). On the contribution of ERPs to study of language comprehension. *Bulletin Suisse de Linguistique Appliquée: Etudes en neurolinguistique*, Neuchâtel, 68: 111–124.
- França, Aniela I.; Miriam Lemle, Mauricio Cagy, Pedro Constant, and Antonio F. C. Infantesi (2004). Discriminating among different types of verb-complement merge in Brazilian Portuguese: an ERP study of morpho-syntactic sub-processes. *Journal of Neurolinguistics*, (forthcoming).
- Frazier, Lyn and Charles Clifton Jr. (1996). *Construal*. Cambridge, Mass.: MIT Press, A Bradford Book. 230 p.
- Frazier, Lyn and Charles Clifton Jr. (1998). Sentence reanalysis and visibility. In Janet Dean Fodor and Fernanda Ferreira (eds.), *Reanalysis in Sentence Processing*. Dordrecht: Kluwer, pp. 143–76.
- Frazier, Lyn and Janet D. Fodor (1978). The sausage machine: a new two-stage parsing model. *Cognition* 6: 291–325.
- Friederici, Angela D. (2002) Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6(2): 78–84.
- Fromkin, Victoria A. (1999). *Slips of the tongue: windows to the mind*. University of California, Los Angeles. Available at <http://psikoloji.fisek.com.tr/psycholinguistics/Fromkin.html>
- Gazzaniga, Michael S.; Richard B. Ivry and George R. Mangun (1998). Language and the brain. In *Cognitive neuroscience: the biology of the mind*, New York: W. W. Norton, pp. 289–321.
- Gollan, Tamar and Lori-Ann Acenas (2004). What is a TOT?: cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory & Cognition* 30: 246–69.
- Gonnerman, Laura (1999). *Morphology and the lexicon: exploring the semantics-phonology interface*. Doctoral thesis, University of Southern California.
- Grodzinsky, Yosef (1990). *Theoretical perspectives on language deficits*. Cambridge, MA: MIT Press, A Bradford Book. 192 p.
- Hickok, Gregory (2003). Functional brain imaging. In Kent, Raymond D. (ed.), *MIT Encyclopedia of Communication Disorders*. Cambridge, Mass.: MIT Press, A Bradford Book, pp. 305–6.
- Hippocrates (400 b.C). *On the sacred disease*. Translated by Francis Adams. The Internet Classic Archive, MIT. Available at <http://classics.mit.edu/Hippocrates/sacred.html>
- Kamide, Yuki and Don C. Mitchell (1997) Relative clause attachment: Non-determinism in Japanese parsing. *Psycholinguistic Research*, 26(2): 247-54.

- Kiel, Christianson; Andrew Hollingworth, John Halliwell, and Fernanda Ferreira (2000). *Reanalysis of garden path sentences: thematic roles assigned along the garden-path linger*. Michigan State University. Available at <http://eyelab.msu.edu/people/ferreira/linger.pdf>
- Kim, Seong-Gi; Sang-Pil Lee, Bradley G. Goodyear, and Afonso C. Silva (1999) Spatial resolution of BOLD and other fMRI techniques. In Moonen, Chrit T. W. and Peter A. Bandettini (eds.), *Functional MRI*, Medical Radiology: Diagnostic Imaging and Radiation Oncology, New York: Springer, pp. 195–203.
- Kutas, Marta and Steven A. Hillyard (1980). Reading senseless: brain potentials reflect semantic incongruity. *Science*, 207: 203–5.
- Kutas, Marta and Steven A. Hillyard (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307 (5947): 161–163.
- Lichtheim, Ludwig (1885). On aphasia. *Brain*, 7(4): 433–84.
- Lopes da Silva, Fernando H. and Van Rotterdam A. (1999) Biophysical Aspects of EEG and MEG Generation. In Niedermeyer, Ernst and Fernando H. Lopes da Silva (eds.) *Electroencephalography, basic principles, clinical applications, and related fields*. Baltimore-Munich: Urban & Schwarzenberg, 4ed., pp. 15-26.
- Luce, Paul A.; David B. Pisoni and Steven D. Goldinger (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (ed.) *Cognitive models of speech processing: psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press, pp. 122–47.
- Maia, Marcus and Juliana Maia (2004). A compreensão de orações relativas por falantes monolíngües e bilingües de português e de inglês. *Revista Letra*, UFRJ, 5 (1): 68–80.
- Marslen-Wilson, William D. (1987). Functional parallelism in spoken word-recognition. *Cognition* 25: 71–102.
- Marslen-Wilson, William D. (1989). Access and integration: projecting sound onto meaning. In Marslen-Wilson, William D. (ed.), *Lexical representation and process*, Cambridge, Mass.: MIT Press, pp. 3–24.
- Marslen-Wilson, William D. (1993). Issues of process and representation in lexical access. In Altmann, Gerry and Richard Shillcock (eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting*, Hillsdale: Lawrence Erlbaum, pp. 187–210.
- McClelland, James L. and Jeffrey L. Elman (1986) The trace model of speech perception. *Cognitive Psychology*, 18(1):1-86.
- Miller, George (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information, *The Psychological Review* 63: 81–97.
- Mitchell, Don C. and Marc Brysbaert (1998) Challenges to recent theories of cross-linguistic variation in parsing: evidence from Dutch. In Hillert, Dieter (ed.), *Sentence processing: a cross-linguistic perspective*. New York: Academic Press, Syntax and Semantics, 31, pp. 313–35.
- Müller, Ralph-Axel; Robert D. Rothermel, Michael E. Behen, Otto Muzik, Thomas J. Mangner, and Harry T. Chugani (1997) Receptive and expressive language activations for sentences: a PET study. *NeuroReport*, 8: 3767-70.
- Neville, Helen; Marc Pratarelli and Ken Forster (1989). Distinct neural systems for lexical and episodic representations of words. *Society of Neuroscience Abstracts* 15: 101.11.
- Norris, Dennis (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, 52(3):189–234.
- Lee Osterhout and Phillip J. Holcomb (1993). Event-related potentials and syntactic anomaly: evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8(4): 413–37.

- Lee Osterhout and Phillip J. Holcomb (1995). Event-related potentials and language comprehension. In: Rugg, Michael D. and Michael G. H. Coles (eds.) *Electrophysiology of mind: event-related brain potentials and cognition*. New York: Oxford University Press, p. 171–215. (Oxford Psychology Series, 25)
- Piaget, Jean. (1969). *The Mechanisms of Perception*. London: Rutledge & Kegan Paul, New York.
- Piattelli-Palmarini, Massimo (ed.) (1980). *Language and learning: the debate between Jean Piaget and Noam Chomsky*. Cambridge: Harvard University Press. 409 pp.
- Pietroski, Paul and Stephen Crain (2002). Why language acquisition is a snap. *The Linguistic Review* 19: 163-83.
- Pinker, Steven (1994). *The language instinct: how the mind creates language*. New York: William Morrow. 494 p.
- Poeppel, David (1996a). A critical review of PET studies of phonological processing, *Brain and Language* 55(3): 317–51.
- Poeppel, David (1996b). Some remaining questions about studying phonological processing with PET: a response to Demonet, Fiez, Paulesu, Petersen, and Zatorre, *Brain and Language* 55(3): 380–5.
- Poeppel, David (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, 25(5), p. 679–93.
- Poeppel, David and David Embick (2004). *Defining the relation between linguistics and neuroscience*, (ms.), 16 p.
- Pylkkänen, Liina; Andrew Stringfellow, Elissa Flagg, and Alec Marantz (2000). A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. *Proceedings of Biomag 2000*. 12th International Conference on Biomagnetism. Helsinki University of Technology, Espoo, Finland, 363–7.
- Pylkkänen, Liina; Andrew Stringfellow, Laura Gonnerman, and Alec Marantz (2001). *Are inhibitory effects in lexical decision due to suppression of activation or competition: an MEG study*. Cognitive Neuroscience Society 2001 Annual Meeting. New York, NY. Available at <http://web.mit.edu/liina/Public/CNS01/CNS01-Pylkkanen.pdf>
- Pylkkänen, Liina; Andrew Stringfellow and Alec Marantz (2002a). Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language* 81(1-3): 666–78.
- Pylkkänen, Liina; Andrew Stringfellow and Alec Marantz (2002b). *Neural mechanisms of spoken word recognition: MEG evidence for distinct sources of inhibitory effects*. A report presented to the Department of Linguistics and Philosophy, MIT/MEG Laboratory. Available at <http://www.cogsci.msu.edu/DSS/2002-2003/Marantz/MEG&InhibitoryProcesses.pdf>
- Sartori, Giuseppe and Carlo Umiltà (2000). How to avoid the fallacies of cognitive subtraction in brain imaging. *Brain and Language*, 74(2): 191–212.
- Signoret, Jean-Louis; Paul Castaigne, François Lehermitte, René Abelanet, and Pierre Lavorel (1984). Rediscovery of Leborgne's brain: anatomical description with CT scan. *Brain and Language*, 22(2): 303-19.
- Skinner, Burrhus Frederic (1957). *Verbal Behavior* Englewood Cliffs, NJ: Prentice-Hall, p. 66.
- Van Petten, Cyma and Marta Kutas. (1991). Influences of semantic and syntactic context in open- and closed class words. *Memory & Cognition*, 19(1): 95–112.
- Vesalius, Andreas (1543) *De humani corporis fabrica* – The Warnock Library. First CD-ROM edition (1998), Octavo Corporation, p. 363, Palo Alto, CA, USA.

- Villringer, Arno (2000). Physiological changes during brain activation. In Moonen, Christ T. W. and Peter A. Bandettini (eds.), *Functional MRI*, Medical Radiology: Diagnostic Imaging and Radiation Oncology, New York: Springer, pp. 3-13.
- Vitevitch, Michael S. and Paul A. Luce (1998). When words compete: levels of processing in spoken word recognition. *Psychological Science*, 9: 325-9.
- Vroomen, Jean and Beatrice de Gelder (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 23: 710-20.
- Warburton, Elizabeth; Richard J. S. Wise and Cathy J. Price (1996). Noun and verb retrieval by normal subjects: studies with PET. *Brain*, 119(1): 159-79.
- Warren, Paul and William D. Marslen-Wilson (1987). Continuous uptake of acoustic cues in spoken word-recognition. *Perception and Psychophysics* 41: 262-75.
- Wernicke, Carl (1874) *The Aphasic Symptom Complex*. Reprinted in: Wernicke's works on aphasia: a source book and review. Translated by G.E. Eggert. The Hague, Netherlands: Mouton, 1977, pp. 91-144.