Updating and Unification of Uncoupled Cohesive Elements

PhD Thesis

Hajnal Jolsvai

Faculty of Psychology and Education, ELTE
Psychology PhD School,
Cognitive Psychology Programme

2009

Mentor: Prof. Valéria Csépe

Co-mentor: Dr. Elyse Sussman

Readers: Dr. Miklós Győri

Dr. László Balázs

Committee: Prof Dr. Márk Molnár (chair)

Prof. István Czigler

Dr. István Winkler

Dr. Katalin Szentkúti-Kiss

Dr. Ildikó Király

Faculty of Psychology and Education, ELTE
Psychology PhD School,
Led by Prof. György Hunyadi

Cognitive Psychology Program
Led by Prof. Valéria Csépe

i. Contents

ii. Acknowledgements

First, I thank my advisor Professor Valéria Csépe, for her continuous support during the last six years. Valéria was always there to listen and to give advice. She involved me in Neurophysiology in the first place. She also encouraged me and made it possible that I persuaded the topic I was interested in the most. She is a lot more than an advisor, she is a parental person to all of her several students. She thought me how to work and be organized in an electrophysiological lab, as well as always had time to read and comment my papers, thesis besides her numerous Academic engagements. Without Valeria, this thesis would have never been made.

Special thanks go to my co-advisor, Dr Elyse Sussman, who is most responsible for helping me complete the writing of this dissertation. She taught me how to write academic papers, and broaden my scientific horizon. Without her encouragement and constant guidance I could not have finished my thesis. She was always there to meet and talk about my ideas, to proofread and mark up my papers, and to ask me good questions to help me think through my problems (whether theoretical, analytical or computational). She also made the Sussman Lab a wonderful workplace and home for the past three years.

Besides my advisors, I would like to thank Professor Istvan Czigler and Dr Istvan Winkler and Viktor Tron for their encouragement and hard questions and to read my papers and provide very useful comments on them. I would like to say thank you to Dr Tamas Makany who gave insightful comments and reviewed my thesis on a very short notice.

Let me also thank to the following people: Bertalan Danko who constantly makes my work easier by introducing new technical solutions to me and Dr Kata Egyed for her support all along these years. I would like to thank to Csépe Lab as well as to Sussman Lab to help my work. Thank you: Gabriella Baliga, Roland Csuhaj, Wei Wei Lee, Elysabeth Dinces, Latha Swamy, Grace Cunanan, Jean DeMarco and Julia Wang.

Last, but not least, I thank Dr Yonathan Fishman and Ted Tucker for their help in programming my last control study with Tucker Davis Program and Kevin Fisher for his valuable technical help for the Sussman Lab.

iii. Publication list

Jolsvai, H., Sussman, E. Unification of associative learned contingencies. (In preparation)

Jolsvai, H., Csépe, V., Csuhaj, R., Sussman, E. Direct comparison of Electrophysiological correlates of morphosyntactic violations and working memory load (Submitted)

Rimmele, J., Jolsvai, H., Sussman, E. An ERP investigation of Temporal and Spatial expectation of auditory attention (Journal of Cognitive Neuroscience (Accepted pending revision)

Jolsvai, H., Sussman, E. (2009) ERP correlates of memory load and congruency in processing familiar songs. Frontiers in Human Neuroscience. Conference Abstract: MMN 09 Fifth Conference on Mismatch Negativity and its Clinical and Scientific Applications

Rimmele, J., Jolsvai, H., Sussman, E. (2009) Auditory Attention: Effects of Temporal and Spatial Expectation. 32nd MidWinter Meeting, Baltimore, Maryland, USA, February 14-19, 2009 (Poster)

Jolsvai, H., Csépe, V. (2008) Electrophysiological correlates of Language processing In: Csépe, V., Gyori, M., Rago, A. (eds): General Psychology III. Language, Osiris kiado, Budapest (Book chapter)

Jolsvai, H., Csépe, V., Csuhaj, R., Sussman., E. (2008) Direct comparison of Electrophysiological correlates of morphosyntactic violations and working memory operations. Neuroscience Meeting, Washington DC, November 15-19, 2008 (Poster)

Jolsvai, H., Csépe, V., Mészáros, É., Csuhaj, R. (2006) Number-person agreement in Hungarian and the role of time delay in lexical access. An ERP study. Pszichologia, 26, (3), 227-246.

Jolsvai, H; Csépe, V. (2006) SLI a case study of set of twins. XVII. Conference of the Hungarian Psychological Association (Presentation), Budapest, Hungary, May 25-27 2006

Jolsvai, H., Csépe, V., Mészáros, É. (2006) The role of time in language processing. An ERP study. International IBRO Workshop Budapest, Hungary, January 26-28 2006 (Poster)

Jolsvai, H., Csépe, V., Mészáros, É. (2005) The role of the time in sentence comprehension. An ERP study. 42. Meeting of The Hungarian EEG and Clinical Neuropsychology Society. Gyula (Poster)

Jolsvai, H.: Williams Syndrome, a model for Cognitive Neuroscience (Book review) Hungarian Psychological Review (2003/2)

Jolsvai, H. Pathways to language (Book review) Hungarian Psychological Review (2002/3)

iv. Preface

This thesis presents a set of neurophysiological investigations of the processing characteristics of uncoupled dependencies that can account for the attention among theorists of both memory and language. Both an artificial and a human neural network are equally required to update and integrate uncoupled dependencies[1], which are quite common in natural language. In behavioral studies on sentence comprehension, much evidence indicates that shorter dependencies are preferred over longer dependencies, and that longer dependencies show a relationship with a greater processing cost. However, it is undecided which processes are involved in the processing of long-distance dependencies is responsible for the increased cost of longer dependencies (Philips et al, 2005). The processing problem of sentences has long been a focus of the psycholinguistics and neurolingustics communities (Gibson, 2000; Kluender & Kutas, 1993). However, from a biological perspective, there are a number of open issues including the neural underpinnings of building an uncoupled e.g. subject-verb structure in relation to the memory and integration processes.

Another intriguing question is whether the integration of two uncoupled contingent elements relies on a specific rule based procedure or whether there is instead an informationally rich knowledge base memory network and a more indulgent unification process that fosters building of complex representations.

Main questions of the thesis

Studies 1 and 2 examine the neurophysiological characteristics of updating and integration processes of uncoupled dependent elements with the focus on whether these two are separate processes.

The other set of studies (3, 4 and 5) investigate the question whether the processing of uncoupled dependencies relies on a specific rule based procedural network, or is there rather a knowledge base that embeds information about the possible

[1] They can arise in situations where elements are structurally close to one another and in their canonical positions, but are separated by additional material, as in subject–verb agreement relations.

probabilistic environment, and, therefore, instead of a traditional rule based procedural system, there is an indulgent unification process that fosters the building of multi element representations?

By focusing on these questions, I endeavor to link neurophysiological data to the relevant psycholinguistic and connectionist assumptions.

1. Introduction

Comprehending a sentence in any language requires a continuous and online organization of the words in use to create a representation of the meaning of the sentence. Many theorists claim that working memory (WM) subserves language processing (Baddeley, 1986; Just & Carpenter, 1992; Waters & Caplan, 1999). As Fuster (1995) puts it:

"*language, especially when it is new, complex and extended in time, makes constant use of those functions of memory and set. As I speak, I need to keep track of what I just said a few moments ago, and, at the same time, prepare for saying what is in accord with that. The predicate is dependent on the subject, the verb on the subject and the predicate, the dependent clause on the larger sentence, and so on. All these are, in essence, cross temporal contingencies, and in speech there is a running reconciliation of these contingencies that constantly change, interleaved and embedded within another*." *(pp. 280-281).*

Shortly after the seminal work of Chomsky (1965), a strong need arose to describe language beyond linguistic formalisms. According to Chomsky, grammar had been described predominantly in terms of a set of generative phrase structure rules, often coupled with rules or principles for further transformation of phrase structures. The Derivational Theory of Complexity (Miller & Chomsky, 1963) suggested that the application of a given rule or transformation could be measured directly in terms of the time it takes for a listener to comprehend a sentence (Miller & Chomsky, 1963). This direct correspondence between syntactic rules and reaction times was soon found to be incorrect (e.g., Bresnan & Kaplan, 1982; Townsend & Bever, 2001). However, the concept of explicit rules has remained almost impregnable. Since language processing requires the combination of abstract categories, such as subject, verb or object, there was a long period in cognitive science celebrating innate abstract syntactic categories in the human brain (Chomsky, 1965). This was especially true after the 1950s as a rebound from the relatively sudden liberation from the behavioral doctrine. However, from the late 1980s connectionism revealed an alternative approach (Rumelhart, 1989) suggesting that categories can emerge through statistical learning. Subsequently, instead of predestined syntactic modules (or empty categories, Pinker, 1994), it has once again become apparent that one of the marvelous features of the human brain is its ability to generate abstract categories instantaneously from individual tokens and handle them in an extremely flexible way. The once frightening Conditioning doctrine (Watson, 1924) got renewed support from some recent language acquisition theories that also emphasize the role of social support in learning (Vigotskij, 1967) the category of verb from individual cases (Tomasello, 2003). Since language is the most over learned human ability, it is really hard to tell if its real characteristics can be described by abstract rules governed symbols or rather statistical learning (Charter & Christiansen, 2001). Since there are new imaging methods in the toolkit of Cognitive Science, these questions have attracted even more attention. At first it seemed simply that the human brain is a “good generativist” following syntax and semantic phases with distinct ERP components (Friederici, 2002). However, acquiring more and more about the brain and the imaging techniques themselves, some theorists have started to reveal that the picture might be more colorful (e.g. Hagoort, 2007; Kutas et al, 2006).

The question to date is not only whether the linguistic descriptions (e.g. Chomsky& Miller 1965) and psycholinguistic data (Gibson, 2000) could be implemented in neural network models (e.g. Berwick &Weinberg, 1984) but also whether they can be verified by brain imaging data (Friederici, 2002).

Event-related brain potentials (ERPs) provide high temporal resolution, on the order of milliseconds, and they have the advantage of making information available on the timing of cognitive processes associated with stimulus events (e.g. Coulson, 2007). ERPs have been used extensively to index cognitive processes initiated by language processing. There have been huge expectations for describing language comprehension from a neurophysiological perspective since Kutas and Hillyard (1980) first reported a semantic related electrophysiological effect and named it N400 effect. Conversely, there have only been limited attempts to articulate how language processing is implemented in the brain due to the few available paradigms. Another problematic issue in the language ERP field is that there is limited research combining methodological and theoretical considerations. In fact, there are many identical studies repeating similar studies (e.g. Friederici & Mecklinger, 1996; Hahne and Friederici, 1999). Similar paradigmatic issue surrounded the well studied MMN component that was long thought to be engendered by an odd-ball or deviant sound among a stream of standard sounds, like AAABAAB, while repeating the same "random" oddball paradigm (Näätänen et al, 1978). On the contrary, it turned out that if the standards and deviants are presented in a predictable context such as AAABAAAB with short SOA between the individual elements then our brains do not perceive B as a deviant anymore but treats the whole sequence AAAB as a standard (Sussman, 2007). Introducing re-evaluated paradigms to the language field would help theorizing operations involved in comprehension. Indeed, theorizing about how the brain implements the process by which a string of words is mapped into the meaning of the conveyed message is in short in supply in the neurocognitive language field. Since ERPs is limited in topographical information, many laboratories use both ERPs and fMRI methods in the hope of extending their assumptions (Snijders et al, 2009; Opitz and Friederici, 2007).

In spite of the overwhelming number of imaging studies used as the ultimate verification of cognitive theories nowadays, concerns have arisen around considering the biological facts of the brain. One of these facts brought up by Uttal (2001) and emphasized by others (e.g. Dobbs, 2005), is that it is a serious mistake to try to describe the highly interconnected human brain displaying nonlinear properties according to independent functional operations or cognitive processes. fMRI can scan a brain cross-section in less than two seconds, enabling it to model most of the brain in one to two minutes. It can work at spatial resolutions as small as two to three cubic millimeters, although in practice it usually collects information in voxels (a term that merges "volume" with "pixel") about two millimeters square and four to five millimeters long (Dobbs, 2005). In reality, neuronal action takes milliseconds, whereas the blood flow follows by two to six seconds; a detected increase in blood flow therefore might be "feeding" more than one operation. In addition, because each voxel encompasses thousands of neurons, thousands or even millions may have to fire to significantly highlight a region (Dobbs, 2005).

Cabeza and Nyberg (2000) reported one of the most comprehensive meta-analyses of cognitive localization studies. Cabeza and Nyberg compared the results for higher cognitive processes such as “working memory” and found that the data from different laboratories were broadly dispersed; the brain activity associated with working memory was spread over half of the brain (Cabeza and Nyberg, 2000). The variability of the reports from different laboratories suggests that narrow localization is, to a substantial degree, mythical and depicts only the “figments of our experimental design” (Uttal, 2001; p.6). Dobbs (2005) also suggests that fMRI neglects the networked or distributed nature of the brain's workings, when it is the communication among regions that is most critical to mental function.

The recording of ERPs is the oldest and cheapest method in the toolkit of cognitive neuroscience. ERPs reflect the sum of simultaneous postsynaptic activity of a large population of mostly pyramidal neurons recorded at the scalp as small voltage fluctuations in the electroencephalogram (EEG), time locked to sensory, motor, or cognitive processes (Kutas et al, 2006). An averaged waveform consists of positive or negative amplitude with a certain latency range and topography, like the centro-parietal N400 peaking around 400 ms post stimulus (Kutas & Hillyard, 1980). This method faces similar methodological problems, as fMRI research does, however instead of a certain brain area, in ERPs research a particular time window reflects a cognitive process by a certain ERPs wave. This approach implies a commitment to the belief that the processing of a complex cognitive event (a sentence for example) is linear or, even if not, it can be captured by comparing brain activities related to one time point; e.g. a word in a sentence. Uttal (2001) draws our attention to the question to what extent our brain is engaged by the demands of the psychological “subdivisions of thoughts” a problem fMRI and ERP research share. It is an open question whether the well described linguistic: semantic and syntactic processes exist on the level of the brain and if so, how they are manifested in a biological sense. Another methodological question how these linguistically different processes can be separated by scalp recorded brain activity. There is an agreement among neuroscientist interested in language comprehension that the human brain can make a distinction between semantic and syntax processing (Hagoort, 2003; Friederici, 2002). The semantics and syntax dissociation concept arose from experiments in which semantic violation correlates with an N400 (Kutas & Hillyard, 1980), while violating the syntax of the sentence elicits a LAN-P600 complex (Friederici & Mecklinger, 1996. Violation paradigm has popularity in the language ERP field since these neurophysiological effects are small and by violation these responses can be made larger (Vos et al, 2001). Confusingly enough, syntax-semantics interplay studies using combined (semantic and syntactic violations) paradigms reported highly inconsistent results with almost all possible combinations of increase/decrease/presence/absence of all the three components, namely LAN, N400 and P600 (see the review of Martin-Loeches et al, 2006). Moreover, unlike any other ERP research field, language studies mostly lack of any control situation in regards to the language specificity of these brain responses. Nevertheless, LAN has also been elicited by correct and more demanding sentences (Kutas et al, 2006), and it is not exceptional that without a LAN only P600 is found to correlate with syntactic violations (Hagoort, 2003). The frontal LAN and centro-parietal N400 occur in the same time window and are both negativities. Furthermore, it is a well-known phenomenon in perception research that ERP elicited by successive stimuli can overlap in time, which can result in distortion of the ERP averages (Woldorff, 1993). The same was suggested for complex cognitive processing (Kutas et al, 2006). Even if these processes indexed by N400 and LAN are originated from different generators, differentiate between them by scalp recorded potentials is quite challenging considering that every sentence incorporates syntactic and semantic features. It goes without saying that manipulating the syntactic characteristics of a sentence changes the semantics of the sentence as well.

Furthermore, both LAN and N400 have a role in memory processes; namely N400 thought to reflect retrieval from long term knowledge (Kutas & Federmeier, 2000) while LAN indexes WM processes (Kutas et al, 2006). Shortly, I will discuss theories that deny the distinction of long term memory and WM stating that representations in WM refer to the same long term representations, but in a different activation sate (e.g. Cowan, 1999; MacDonald & Christiansen, 2002). Following this line of though, it is plausible to predict that these two are not separate components but if so, since they occur in the same time window at neighboring areas, the measurable responses might overlap during sentence processing even if their generators are dissimilar. It is also a logical possibility that these two component index different processes, but only defined by improper tasks after considering assumptions of inseparable WM and knowledge (e.g. Cowan, 1999). One might consider that the violation paradigm most often used in syntax and semantics has been criticized as being inappropriate for use in ERP and fMRI language studies (Thierry, 2003). Thierry argues that instead of the intended psycholinguistic processes responses elicited by the violation paradigm, they only reflect general error detection, attentional shift or repair processes. For this reason a group of researchers chose to manipulate syntactic complexity (e.g. comparing the processing of Subject-Object (SO) sentences to Subject-Subject (SS)[2] sentences, King & Kutas, 1995) instead of grammatical violations. By comparing the processes of SS and SO sentences, one might also consider that the meanings of these sentences are different as well, since the thematic role changes. In a serial model the changed semantics of the SO sentences would not be significant since syntax processing is not influenced by other modules (Friederici, 2002). Considering the interactive models of sentence processing, however, semantic processes can and does influence syntactic processes (e.g. Just & Carpenter, 1992). Furthermore, it is well-known that SO sentences, similar to passive structures, are less frequent than SS and active structures. Therefore the processing differences between them do not necessarily indicate low verbal working memory capacity (Just & Carpenter, 1992) or differences in the domain specific grammatical processor (Caplan & Waters, 1999) but rather less experience with these sentences (MacDonald & Christiansen, 2002).

Moreover, subjects are usually asked in grammatical violation studies to make grammaticality judgments about the presented sentences while their neurophysiological responses are recorded in parallel. Syntactic positive Shift, known also as P600, has been elicited by semantic violations (van Herten et al, 2004). It has also been argued simply as being a P300 component that has nothing to do with

[2] e.g. “The reporter who attacked the senator admitted the error” (SS). “The *reporter* who the senator attacked admitted the error”(SO)

syntax, but is elicited in cognitive tasks when subjects are forced to respond (Picton, 1992; Polich, 1998; Pritchard, 1981). It has been suggested that the P300 reflects some aspects of stimulus categorization (e.g., Johnson & Donchin, 1980; Mecklinger & Ullsperger, 1993) while others claim that P300 indexes the neurophysiological activity required the ongoing model of the environment is revised or updated in working memory (Donchin, 1981; Donchin & Coles, 1988).

In fMRI studies using the subtractive method, brain activity measured during the "experimental" condition is subtracted from the "control" condition. The subtractive method ignores places where activity may be constant (Uttal, 2001). It would be necessary to assume that the mechanisms were activated in serial order; that the mechanisms were steady enough to remain unchanged when the task changes; and that the data used to select between alternative processes were sufficiently clear-cut to distinguish between alternative explanations for the subtraction of one image from another (Uttal, 2001). Most importantly, even if done with great accuracy, fMRI method disregards the fact that many regions other than those highlighted may be active and involved in the cognitive process under study. Even though the cumulative activity in these other regions may appear to be unchanged, this does not necessarily mean that the underlying patterns of neuronal activity are identical in the two conditions (Uttal, 2001). Thus, there is a logical flaw built into the subtractive method; it ignores places where activity may be constant in integrated quantity (which is what the fMRI measures) but quite different in microscopic detail (Uttal, 2001).

Another methodological problem is that experiments intended to measure brain activity during a particular cognitive operation are identical only in how the cognitive processes are named (e.g., working memory), but differ in the actual tasks or experimental design (Cabeza & Nyberg, 2000; Uttal, 2001). It is very difficult issue because it is also disputed what exactly the well-established and generally used linguistic WM tests (e.g. Danneman & Carpenter, 1980) measures (MacDonald & Christiansen, 2002) not to mention the difficulties in comparing different imaging studies used altering WM tasks (e.g. Just et al, 1991).

Thus, recent neurophysiological models of sentence comprehension (Friederici, 2002) are far from successful in integrating psycholinguistic models with neuroimaging data or in attempting to resolve any of the controversies that have arisen. Moreover, as long as the paradigms are not well established in the neuroimaging field one must be cautious interpreting brain responses correlating with complex cognitive behavior (Thierry, 2003).

It is important to bear in mind that imaging techniques available to date have several methodological and technical difficulties. However, one should be optimistic since there are new studies that hope to reevaluate the paradigmatic issues (e.g. Lau et al, 2006; Dicker et al, 2009), and the techniques and processing methods have been developing constantly (Kutas et al, 2006).

This dissertation is making a novel contribution in fusing neurophysiological data with new psycholinguistic theories (Gibson, 2000) along with artificial neural network models (Vosse & Kempen, 2000). The focus of this thesis is twofold. The first two studies concentrate on the updating and integrating operations of processing uncoupled dependencies, i.e., subject and verb separated by other intervening sentence elements or time. To resolve dependencies between nonadjacent constituents (e.g., long distance dependencies) 'the listener' must refresh or hold online memory representations until encountering the second element of the dependency. The role of working memory in processing uncoupled dependencies is in fact an intriguing topic. Evolutionary theories of syntax suggest a direct link between increased capacity of working memory and the capability of combining words into more complex structures that eventually led to the appearance of syntax in our ancestors' language usage (Alboitz et al, 2006).

I will compare the most prominent psychological language-related working memory models (Baddeley, 1986; Caplan & Waters, 1999; Just and Carpenter, 1992; Cowan, 1999) as well as psycholinguistic models (Gibson, 2000) with a focus on processing resources. This is followed by a discussion on connectionism as a new alternative of symbolic, serial language comprehension models (Frazier, 1987). This approach proposes a model (Vosse & Kempen, 2000) that gives new insights into the known ERP components (Hagoort, 2003). Based on empirical evidence, I will bring a neurophysiological model of language comprehension (Hagoort, 2005, 2007, 2009) together with a psycholinguistic model (Gibson, 2000) in an effort to suggest an overarching solution to the paradox of a syntax ERP component called Left Anterior Negativity (LAN).

The focus of the last three experiments is on the issue of the validity of distinction between traditional grammar (rule) and the declarative memory (e.g. Ullman, 2001). The thesis ends with three studies by questioning whether syntactic ERP components are tied to rule based processing and provides evidence to support the assumption that the traditional dissociation of rule and declarative memory is dematerialized (e.g. Hagoort, 2007).

2. Working memory in language comprehension

The role of working memory has long been in the spotlight of cognitive research (Baddeley & Hitch 1974), yet there is more divergence surrounding it than agreement (e.g. Cowan, 1999). Even the operationalization of the notion of working memory is not an easy task. It has had a key role in scientific thinking, however (Cowan, 1999). Working memory is not a separate domain of cognition, but has a role in the processing of different representations such as comprehending language or understanding the mental states of others. Modeling working memory means committing to a certain type of thinking about the whole cognitive system (MacDonald & Christiansen, 2002). Below, I present a series of five issues to illustrate the models of neurophysiological language processing and to prepare the ground for the experiments described later in this thesis.

The four issues are: (i) whether or not working memory is separated from long term knowledge; (ii) unitary nature of working memory (iii) type of representations (iv) nature of processing (v) evidence type each model relies on.

The question of separated memories (short and long term memory) is relevant here as words depend on long-term knowledge in order to be activated when comprehending an actual sentence. It is also intriguing because the last three experiments investigate the relevance of the long standing tradition of separating rule (e.g. procedural) and knowledge (declarative) systems.

There is no agreement among theorists as to whether working memory is separated from Long Term Memory[3]. Interestingly, most of the psychological theories claim that working memory is functionally and physically separated from Knowledge (Baddeley, 1986; Just & Carpenter, 1992; Waters & Caplan, 1996). As a consequence of the long dominant interpretation that human cognition has a serial processing nature, with separated processing boxes like short and long term memory, working memory is a

[3] Long term memory and Knowledge and knowledge base will be used interchangeably.

separated system usually localized in the frontal lobe (e.g., Kimberg & Farah, 1993). In fact the frontal lobe evolution is the most recent and its activation is typically registered by fMRI during all complex cognitive tasks (Semendeferi et al, 1997). Instead of remaining content with the traditional “one area, one function” view of working memory, a group of scientists broke with this heritage by considering the underlying principles of the human neural system. In fact, there are suggestions that attention and working memory load correlate with a low gamma event-related synchronization between frontal and parietal areas (Philips & Takeda, 2009). Neural net models also propose that neither is working memory separate from knowledge base nor does the locus of language processing stand apart from knowledge and working memory (MacDonald & Christiansen, 2002), which will be discussed shortly.

There is a long tradition since Williams James (1890) of viewing human memory as a dual system of short and long term experience, a concept that has continued to dominate in subsequent and current memory research (e.g. Atkinson & Shiffrin 1968, Baddeley, 1986). However, questions about its legitimacy appear again (Craik & Lockhart, 1972) and again (MacDonald & Christiansen, 2002) in the history of remembering or experience. The “Levels of Processing Theory” (Craik & Lockhart, 1972) was welcomed as an early verification of the unitary nature of memory (e.g. Postman, 1975). Nevertheless, Craik and Lockhart (1972) did not deny the concept of a dual memory system; their model focused on the levels of encoding in long term memory and ignored short term memory. Starting in the mid seventies many models “zoomed in” from the opposite direction and working memory theories have started to have an effect on the scientific thinking of cognitive architecture (Cowan, 1999). Even the distinction between short and long term memory was built on the hypothesis - among others- of the limited nature of short term memory vs. the unlimited capacity of long term memory. From the demands of the World War II scientists became obsessed with identifying the limits of human short term capacity (Miller, 1956) since there was a critical need to know how long a person could pay attention to a monitor showing monotone radar movements.

Since the expression, “short term memory” was associated with rigid boxes, there was a desire to replace it in the newer models (e.g. Baddeley & Hitch, 1974). The term "working" indicates that mental work requires the use of information and that short term memory has a dynamic nature. In the eighties with this new inspiring and modern name, more and more working memory research appeared. Nevertheless, over the next twenty years, working memory still remains functionally separated from long term memory in most of the theories (e.g. Just & Carpenter, 1992 and see an opposite approach: Cowan, 1999). Different working memory theories have been proposed conveying diverse views about the whole cognitive architecture. Most importantly, the concept of limited working memory resources or capacity for temporary storage and manipulation of information has played an important role in many theories of language processing (e.g., Baddeley, 1986; Just & Carpenter, 1992). Most of these linguistic working memory models share the common feature of having functionally and physically separated knowledge and "working "areas. However, newer psychological models (Cowan, 1999) and neurocognitive data (Ruchkin et al, 2003) and neural network models (MacDonald & Christiansen, 2002) suggest that it might be plausible to reevaluate this traditional idea and consider that the representations of long term and working memory only dissimilar in their activation level or state (Nairne, 2002; Ruchkin et al, 2003). This concept can be easily correlated with the proposition that working memory is a synchronization state between the frontal and parietal areas (Philips & Takeda, 2009). Not surprisingly, many neuroscientist being occupied with the topic of consciousness, suggest a similar synchronization theory (Singer, 2001).

In the last 30 years, Baddeley's (1986) model has grown into the most prevalent working memory theory, especially in Europe. Nevertheless, it has limitations and critics particularly in relation to its key component; the executive function. Baddeley's model consists of a central executive control and two "slave systems" specialized respectively for temporal storage of phonological and spatial information: the phonological loop and the visuo-spatial sketchpad. The characteristics of these systems have been identified largely from behavioral investigations. The phonological loop is further fractioned into a passive phonological store and an active rehearsal system.

Baddeley's (1986) theory has been attacked in most case for his positing a modern homunculus embodying the central executive (Hazy, Frank & O'Reilly, 2007). This mechanism is responsible for the control and regulation of working memory, including the coordination of the subsidiary memory systems, the control of encoding and retrieval strategies, the switching of attention and the manipulation of the material held in the slave systems (Baddeley, 1996). The central executive however lacks supplementary storage capacity. This model has an inherently non-unitary WM nature with specialized subcomponents; even the executive function may not be a unitary construct. But this question will be resolved based on empirical exploration currently underway (Baddeley & Loggie, 1999). Each component has constraints in line with the specific function that each is responsible for and these constraints may arise from the capacity for activation, the capacity of rehearsal, the capacity for complexity of material, or from the extent to which they are supported by acquired strategies and prior knowledge (Baddeley, 1986). According to Baddeley, working memory and long term memory are two functionally separable cognitive systems and the symbolic representations are processed in a serial manner (Baddeley & Loggie, 1999).

Two associated models should be mentioned here that are similar in spirit: the theories of Waters and Caplan (1999) and Just and Carpenter (1992). The difference between them is that the latter denies the reality of specific working memories for different linguistic processes, while the former permits it. The other distinctive feature between these two models is that while Waters and Caplan's model follows the tradition of serial processing, Just and Carpenter suggest a parallel view.

Serial models (e.g. Levelt, 1999) propose that information flows from processing level by level, while parallel models assume the co-occurrence of processing information at different levels that can interact. Serial models go hand in hand with the concept of modular localism. The principle characteristic of localistic models is the establishing of a direct connection between the structure and functioning of a particular brain region with behavior (Lurija, 1966). A challenge to these theories comes from the consideration of brain damaged patients with dysfunctional behavior but without damage to the responsible brain region and from cases with preserved function in spite of damage to corresponding regions (Mildner, 2008). On the bases of the neuronal properties of Hebbian synapse, Pulvermüller (1992) calls functional neuronal groups “cell assemblies.” However this is not correspondent of localistic theories since his approach relies on neural networks. Contrary to the modular assumptions, interactive models (e.g. Marslen-Wilson & Tyler, 1980; Hagoort, 2003) propose that there is interaction occurring in parallel between different levels of information and different processes (e.g. MacDonald & Christiansen, 2002).

Waters and Caplan (1999) also differentiate between “interpretive” processes, referring to subtracting the meaning of a sentence through puzzling out its structure and “post-interpretive” linked to long term semantic memory. The unconscious interpretive processes are unlike those involved in other conscious verbally mediated tasks such as learning a poem or a list of words.

Waters and Caplan (1996) argue that there is evidence against the Single processing Resource (SR) model (Just & Carpenter, 1992; King & Just, 1991) that claims that verbal memory resources can be dedicated to all verbal tasks (Just & Carpenter, 1992). The evidence comes from patients with broken verbal working memory who have been shown to do well on tests of syntactic comprehension. The Separate Sentence-Interpretation–resource (SSI) theory (Waters & Caplan, 1996; 1999) suggests that there is a specialized part of the verbal working memory system devoted to the interpretive processes such as assigning syntactic structure and through this structural information determining the meaning of the sentence. Their model is mostly based on neuropsychological evidence of patients’ behavioral performance in demanding concurrent tasks. They reported patients with poor performance in standard verbal WM tasks[4] while their interpretive “from structure to meaning” transferring capacity remained intact (Waters & Caplan, 1999). It goes without saying that Waters and Caplan’s model is closely related to generative linguistic theory (Chomsky, 1965). It has been suggested that the left perisylvian cortex – the pars opercularis and traingularis of the third frontal convolution, Brodmann’s areas 44 and 45 (Broca area) - can be associated with interpretive processes (Waters & Caplan, 1999). Patients with

[4] E.g. digit-span.

Broca's aphasia who have lesions in this region have difficulties in processing sentences with relative clauses. However, fluent aphasia patients with lesions outside of this area could process relative clause sentences (Swinney & Zurif 1995; Swinney et al. 1996). Two studies have reported a localized increase in regional cerebral blood flow (rCBF) in the pars opercularis when subjects made acceptability judgments about sentences with object vs. subject-relative clauses (Caplan et al., 2000; Stromswold et al. 1996). These results support a narrower degree of localization (Broca area) of syntactic comprehension (Caplan et al, 2000). Not all available data support this narrow localization, however. Just et al. (1996) have reported increased rCBF in both the left frontal and the left temporal lobes (and, to a lesser degree, in the homologous contralateral cortical regions) in a question answering task using sentences types that were similar to those used by Stromswold et al. (1996).

In contrast, the WM in Just and Caprenter's model is not segregated into different linguistic processes, but propose one single resource mechanism within the language domain (Just & Carpenter, 1992). This limited natured resource is shared for temporal storage and processing (Just & Carpenter, 1992).

Danneman and Carpenter originated a WM dual task test requiring participants to memorize target words while comprehending sentences by e.g. reading them loud. The most number of target words recalled immediately after reading through the sentences is taken as a measure of reading span. It has been suggested that reading span reflects the processing efficiency of reading (Danneman & Carpenter, 1980, 1983) or working memory for language (Just & Carpenter, 1992). Differences in linguistic working memory capacity are liable for individual differences in language comprehension (Just & Carpenter, 1992). Simulating individual differences Just and Carpenter chose different values for the maximum activation parameter in their computational model, called CC READER, developed within the 3 CAPS production system frameworks (Haarmann, Just, & Carpenter, 1992; Just & Thibadeau, 1984; Thibadeau, Just, & Carpenter, 1982). In this production system, declarative knowledge, such as lexicon and rules, and the set of condition-action rules (productions) are stored in long term memory. A separate working memory area is used to process and store current input and partial products of ongoing computations. WM and knowledge can be damaged separately:

"*Stated from the perspective of the CC READER model, the theory assumes that, while the lexicon and the production rules in the system are still intact, the maximum amount of activation available for the storage and processing of linguistic information is far more severely limited in the aphasic system than in the normal system". (Carpenter et al., 1994, p. 1093).*

In Just and Carpenter (1992) model several productions can occur in parallel within a given processing cycle. This allows the different pieces of information to simultaneously coexist in working memory. Information in working memory is assigned by a numerical activation. If a representation is lack of sufficient activation, it starts to decay from working memory. During the processing of a sentence the model uses the items in working memory to build explicit syntactic and semantic representations for the purpose of interpretation. This information may be set by several kinds of productions (lexical, syntactic, etc.) allowing interactions across processing levels. In Just and Carpenter's (1992) model, the existence of a single working memory for language comprehension reflects interactive processing. The most important part of their concept is the emphasis on individual differences. Even though the interpretation of their data might be questionable (Waters & Caplan, 1996), their work had a great impact on cognitive science by emphasizing individual differences as a refreshing change in cognitive psychology. Most of the criticisms of Just and Carpenter's (1992) model came from Waters and Caplan (1996) who questioned their data in support of a single linguistic working memory capacity[5]. Other critics emphasized that individual differences can arise by practice and experience (MacDonald & Christiansen, 2002). Cowan (1999) proposed an alternative to the above-mentioned models. While his model operates with symbolic

[5] The fact that Just and Carpenter's 'one resource pool' only focused on language comprehension led to an interesting misunderstanding in the literature that they assumed a domain general working memory (Miyake & Shah, 1999).

representations similar to the above-discussed theories (Just & Carpenter, 1992; Baddeley, 1986; Waters & Caplan, 1996), he proposed the unique concept of unitary assumption which does not separate two memory systems. Cowan (1995; 1999) proposed that short-term memory stores are constituted by an activated subset of long-term memory. Cowan argues for the construct that short-term memory involves all information accessed by a task, including activated memory in the focus of attention, activated memory not in the focus of attention, and inactive memory accessible by activated retrieval cues. Short-term auditory sensory memory processes in experiments involving the presentation of multiple streams of stimuli are examples of the latter two types of activation (Cowan, 1984; Cowan, 1995). In Cowan's model the capacity limitation of short-term memory is due to the limited capacity of the focus of attention (Cowan, 1999). Baddeley (2001) argued that construing short-term memory as activated long-term memory is conflicting with neuropsychological data, since there are individuals with long-term memory deficits but not short-term memory deficits, and individuals with short-term memory deficits but not long-term memory deficits. Cowan (1999) relying on Rickard & Grafman (1998) amnesic data argues, that such deficits can be explained by impaired hippocampal-neocortex binding processes.

Besides the existence of psychological models depicting working memory in relation to language comprehension (e.g. Baddeley, 1986; Just & Carpenter, 1992), there are other attempts to describe sentence processing by computational resources. For example, Gibson (2000) proposed a new psycholinguistic model (Dependency Locality Theory, DLT) to describe the relationship between sentence processing and available computational resources. DLT is relevant to this thesis for two reasons. First, the separation of integration and memory processes (Gibson, 2002) will be tested in the first two experiments. Secondly, the DLT (Gibson, 2000) theory will be joined in later chapters with a new neurophysiological language processing model (Hagoort, 2003, 2005, 2007) that has a close association to a connectionist model (Vosse and Kempen's model, 2000). By taking account of these assumptions, I aim to resolve the paradox of the existence of two functionally different language ERP components named LAN (which will be discussed shortly).

3. Dependency Locality Theory

Gibson (1998; 2000) proposes that the complexity of a sentence is greatly affected by the distance between two uncoupled constituent elements in a sentence. According to the Dependency Locality Theory (DLT) (Gibson, 1998), the sentence comprehension process comprises two different components, each applying neural resources. The “memory cost” component or structural storage is associated with the online tracking mechanism of the syntactic requirements as the online comprehension process takes place. In other words, this memory component is activated when, for example, a noun is presented in the sentence; an encounter with a verb is then predicted in order to form a complete clause. The memory cost is high if following the first part of the dependency there are discourse elements needing completion in a given sentence before encountering the second element of the uncoupled dependency. The second component is “integration” which connects each incoming word to a prior word on which it depends in the sentence. DLT suggests that both memory cost and integration cost are influenced by locality: the longer a predicted category must be kept in memory before the prediction is fulfilled, the greater the cost for maintaining that prediction. Additionally, the greater the distance between an incoming word and the dependent to which it is attached, the greater the integration cost. The DLT theory is an activation-based approach with a strong association with Just & Carpenter, (1992) model. Each discourse structure is associated with an activation level (i.e., a number between 0 and 1) representing the probability of the structure element, according to the permutations of the constraints. There is a limited pool of computational resource units available to activate representations, so if there are more available resources, the activations take place more rapidly. In addition, the more resources a particular sentence element requires, the slower that activation will be. High frequency lexical items require fewer resources to become activated than low frequency lexical items. Structure building consists of looking up the current word in the lexicon and matching the categories of these lexical entries to the predictions in the structures built thus far. Integration has a syntactic component, responsible for linking structures together, such as matching a syntactic category prediction or connecting elements in a dependency string. In the DLT theory (Gibson, 2000) a fixed amount of computational resources is required to perform the integration as well as additional resources relative to the distance between the elements being integrated. Consequently, longer distance integrations require more resources. Each lexical item in a structure has an activation level independent of the activation level for the whole structure. The lexical activations decay as additional words are integrated. Performing an integration requires, first of all, matching the category of the current word with a syntactic prediction. This match then reactivates the lexical head/dependent associated with the syntactic prediction so that the plausibility of the head-dependent relationship can be evaluated within the discourse context. Linguistic integration processes and storage rely on the same pool of working memory resources (Just & Carpenter, 1992 and the comprehension of a sentence performed by parallel processes.

In the first two studies of this thesis, the neurophysiological correlations of memory and integration (Gibson, 2000) of uncoupled dependencies will be tested. These two processes are separable but according to Gibson they share the same resources. Grodner and Gibson (2005) offer behavioral evidence for their theory, however, their study was unable to demonstrate that these two processes are two cooperative, yet different processes. It is possible that these two cooperating processes are share the same neural recourses (Just & Carpenter, 1992) or they are functionally inseparable and represent activation pattern differences of the same connectionist network (MacDonald & Christiansen, 2002) Studies 1 and 2 attempt to investigate this interesting issue further.

4. Connectionism as an alternative concept of cognition

The symbolist predominance of the Chomskian cognitive psychology (1965) led a group of scientists to create a new approach to cognition as an alternative to the conception of mind (Rummelhart et al 1986; Churchland & Sejnovski, 1994). Connectionist models describe cognition as a neural net of interconnected units. Simulating human cognition has long been at the center of scientific interest. The first attempts (von Neumann, 1951; Turing, 1950) involved manipulating symbols in accordance with the mainstream scientific thinking of the time. Until the 1980s, connectionism was mistreated on the one hand because of the supremacy of symbolic theories in cognitive sciences (Fodor, 1983; Chomsky, 1965), and on the other hand due to the limited performance and capability of the early models. These early models mostly focused on modeling perception and neglected 'higher order cognition'. Some of the limitations have since been overcome (Rummelhart et al., 1986) and connectionism became a viable alternative to the symbolic models of cognition.

A salient fact about connectionism is that it relies on the mathematics of dynamic systems as opposed to the discrete mathematics used by the formulae of formal logic (syllogisms) in symbolic theories (Ramsey et al, 1991). Learning in the connectionist framework refers to the changes of the node weights reflecting the influence of Hebb's (1949) concept of learning of synapses. While in traditional cognitive science theories cognition is a rule-governed symbol manipulation, according to the connectionist approach it is seen as neural network activity of distributed, not necessarily symbolic representations. Distributed representations are expressed by connections between multi-layered nodes; therefore they are more resistant to damage than local representations (Rummelhart, 1989). An appealing feature of connectionism is its biological plausibility (but see a contradictory proposal: Dror & Gallogly, 1999).

Connectionist networks are deliberately analogous to neural processes in the brain, where units represent neurons, connections embody synapses, and activations symbolize neural signals. "Roughly speaking the human brain can be seen as comprising of a very huge number of straightforward processors, neurons which are heavily interconnected into a compound network" (Christiansen & Charter, 2001, p.21). Even Dennett (1991) makes reference to "connectionist architectures of neuron-like elements" (p. 239). As Christiansen and Charter describes connectionist models mimic information processing occurring by large numbers of neuron operating cooperatively and simultaneously by communicating numerical values encoded by firing rate"(p.21). Neural network models also consist of large numbers of simple processor units or nodes interconnected to a complex network to simultaneously and cooperatively process information. In line with the assumption that biological neurons are numerical processors nodes transmit numerical values and the output of a unit is usually assumed to be a numerical function of its inputs (Christiansen & Charter, 2001). Another attractive feature of connectionist models is their fast parallel processing. Biological neurons change state very slowly, operating in a time scale of milliseconds, in comparison to computer computations that operate in nanoseconds (Rumelhart, 1989). Nevertheless, the brain can solve many complex problems in less than a second despite the fact that individual neurons operate slowly compared to the speed of an artificial neuron. The operation of a connectionist network simulates the performance of the human brain mainly through its capability to recognize patterns (Bedini & Tonazzini, 1992). Neural networks can learn through examples as well generalize connections between characteristics or properties (Chen et al, 1996). They are robust as well, as they are resistant to some degree to incorrect, noisy input or partial damage (Rumelhart et al 1986). Connectionist networks may also provide a naturalistic mechanism for creating concepts as they embody a theory where there is no need to operate with inborn concepts. There have been numerous criticisms against connectionism, (e.g. Fodor & Pylyshyn, 1988). Having discussed all the promising features of neural networks, Table 2 summarizes the prominent criticisms, though without discussing them in full.

Criticisms
Lack of reverse connections in the brain that is necessary for backward propagation used by many of the networks (e.g. Kukla & Walmsley, 2006).
Connectionist units are too fast in comparison to the biological neurons (Rummelhart et al, 1986).
Connectionist models mostly ignore the role of neurotransmitters and hormones (e.g. Thagard, 2002)
Network learning is slow; it requires a large amount of explicit feedback to learn. Although some networks, for example self-organizing maps, may use unsupervised learning mechanisms, the brain often seems to learn a new concept or pattern after a single presentation.
Difficulty in implementing systematic and productive operation in connectionist architecture, despite the fact that one might think our cognition is based on these features.
There are many different types of neurons in the brain and the connectionist units are meant to represent them all.
Neurons follow the 'all or nothing' principle as they either fire or they do not fire, but cannot be both excitatory and inhibitory at the same time (Hebb, 1949).

Table 1. Summary of the most prevalent critics against the Connectionists assumption.

MacDonald and Christiansen (2002) have proposed a connectionist model for language processing. They disagree with Just and Carpenter (1992) that there is a functionally separated working space to process and store current input and partial products of ongoing computations. According to MacDonald and Christiansen, there is only one network; the processing takes place within the activation patterns of the units, which are distributed throughout an entire network organized in a multilayered fashion rather than being localized in functionally discrete states. They suggest that working memory is not separate from other cognitive systems but that there is one unitary system, the "language processing capacity" where different activation patterns can be responsible for the different processes of language comprehension.

Another significant neural network is the Unification model developed by Vosse and Kempen (2000).

"The dynamics of the Unification Space model were inspired by the metaphor of bio chemical synthesis: molecules floating around in a test-tube and entering into chemical bonds with other molecules (unification of nodes). The resulting larger structure may be insufficiently stable and fall apart again. After that, the segments continue their search for suitable unification partners until a stable configuration — that is, the final parse tree – has been reached" (p.4.).

According to a neural network model introduced by Vosse and Kempen (2000), words in the mental lexicon are held as lexical frames, which are four-tiered unordered trees which are “mobiles‘‘, i.e., there is no ordering among branches. They have variables, such as NP, DP, PP and S at certain places that can be linked to form structure descriptions for sentences. Each variable has an associated set of features, for example, person—{first, second, third}, number—{singular, plural}, case—{nominative, accusative}. This model does not construct and manipulate syntactic trees directly. Instead, it creates and operates on a network of connections between nodes of “lexical frames” (i.e. lexically anchored elementary syntactic trees) that are retrieved from the mental lexicon as “chunks”. Every connection in the network represents a potential attachment alternative. In the online comprehension process, structural frames associated with the individual word forms incrementally enter the Unification workspace. Constituent structures spanning the whole utterance are formed by a unification operation in this workspace. This operation consists of linking up lexical frames and checking agreement features between them. The unification process used in this theory differs from head-driven phrase structure grammar and lexical-functional grammar in being non-recursive and involving only feature unification. Two lexical frames are combined by unifying a variable from one frame with a variable from the other frame. Unification is an agreement check between two nodes that become unified. At each moment, the Unification space consists of a set of lexical frames. Lexical frames are linked by unification-links that represent unifications of variables present in these frames. Each unification-link has an instantaneous strength and each lexical frame has an activation value, which are real numbers. Alternative unification candidates are usually available at any point of parsing because of the inherent ambiguity of language, but ultimately one phrasal configuration results. This is reached through a lateral inhibition process implemented in the model (Vosse & Kempen, 2000; Kempen & Vosse, 1989). Recency also plays an important role, since foot (bottom layer) nodes that enter the unification space more recently have a higher activation level than those entered earlier and therefore find linkable other nodes easier than nodes entered earlier. In addition, strength levels of the unification links can vary as a function of plausibility and semantic effects.

5. Neurophysiological models of sentence processing

With roots in the generative theories, Angela Friederici (2002) was the first to introduce a model of sentence comprehension that is motivated by the time course of different language-relevant ERP effects. She proposed that the first ERP effect reflects an autonomous, purely syntactic phase (at around 150 ms) called Early Left Anterior Negativity (ELAN) and it is associated with the identification of word category information. The empirical evidence for this autonomous first phase in sentence processing is derived from a series of German language studies in which Friederici and her colleagues (1993) elicited ELAN by auditory words presented with word category violation coded in the prefix. Participants were listening to sentences such as "Die Birne wurde im gepflückt" ("The pear was being in-the plucked") or "Die Freund wurde im besucht" ("The friend was being in-the visited"), in which the prefixes "ge-" and "be-" in combination with the preceding auxiliary "wurde" indicate a past participle, while the preposition "im" requires a noun. In this case a very early ELAN was observed. In the Friederici (2002) serial model, LAN and N400 components accompanying the second processing phase range between 300 and 500 ms and they are linked to the identification of lemma and morphological information. This processing phase results in the integration of semantic and morphosyntactic information and thematic role assignments. The third phase (from 600 ms), reflected by the P600 component, is interactive yet; in this phase the processes of reanalysis and repairs occur. According to the model, the first syntactic process (ELAN) is obligatory, automatic and encapsulated.

However, it is disputed whether ELAN is a separate component from LAN. Recent studies suggest that the earliness of these ERP correlates depends on other factors, providing evidence that the early latency of ELAN can be explained in a more parsimonious way. Lau et al. (2006) showed that the preceding sentence context builds up expectancies for future upcoming sentence elements and elicited a greater early negativity for violations of strongly predicted word categories compared to weakly predicted ones. Dikker et al. (2009) showed that the earliness of this effect is limited to words where the category is marked by frequent and salient closed class word category morphemes, generating an enhanced sensory response when presented in unexpected syntactic context. Additionally, Hagoort (2003) elicited later latency for word category violation in Dutch where word category information was coded in the suffix and suggested that the earliness of ELAN's latency is only an artifact of word category coded in the prefix.

LAN is a more established syntactic component around 400 ms. It has been elicited by different morphosyntactic mismatches like number, gender, and tense violations (e.g., Friederici, 2002; Friederici & Meyer, 2004; Hagoort et al., 2003; Kutas & Hillyard, 1983; Münte et al., 1993, 1994). It has also been engendered by word category violations (Friederici & Mecklinger, 1996; Lau et al, 2006). In some studies, LAN is accompanied by a posterior positivity (P600) for syntactic violations (Hagoort, 2003; Friederici, 2002). However, similar LANs were elicited in other studies that manipulated memory load instead of grammaticality (King & Kutas, 1995; Streb et al, 1999, Jolsvai et al, submitted). Interestingly, the interpretation of the functional nature of LAN[6] is convoluted. The complications originate from different group of studies that have reported similar LANs elicited by either syntactic violations (e.g. Münte et al, 1994) or by grammatical but demanding sentences (Kluender & Kutas, 1993). In contrast to the syntactic view (Friederici, 2002), other studies proposed that LAN reflects working memory operations (Kluender & Kutas, 1993; King & Kutas, 1995; Streb, Rösler, Henninghausen, 1999; Anderson & Holcomb, 2005). According to this view, (left) anterior negativities (LAN)[7] elicited by syntactic violations represent a special case of more general working memory processes (Kutas et al, 2006). This is compatible with the finding that both lexical and referential ambiguities seem to elicit very similar negativities (Van Berkum, Brown, & Hagoort, 1999; Hagoort & Brown, 1994). The

[6] A number of studies have either failed to elicit LAN or were unable to reach statistical reliability using violation paradigms (Hagoort, Brown, & Groothusen, 1993; Hagoort & Brown, 2000; Günter & Friederici, 1999; Günter, Stowe, & Mulder, 1997).

[7] Neither the anterior distribution nor the left lateralization of these negativities was consistently confirmed (e.g., Coulson, King & Kutas, 1998; Günter & Friederici, 1999; Hagoort, Wassenaar & Brown, 2003; Münte & Heinze, 1994).

ambiguities are not syntactic in nature; however, they tax working memory more than sentences in which lexical and referential ambiguities are absent. This view denies a special relation between the LAN effects and syntactic processing, and instead relates them to the general memory requirement for language processing. Additional support for the working memory model comes from studies showing that the amplitude of LAN correlates with working memory span (Vos et al., 2001).

Two critical aspects need to be considered here. First, language related ERP components also carry the traditional "one brain area/ ERP component, one function" modular postulation (e.g. Lurija, 1966). This is already implying at least a serial order of sentence comprehension. However, this method only relies on postsynaptic activity and one cannot determine when a certain process began based on the timing of the ERP components. In fact, knowing that P600 is later than N400 or LAN only means that neuronal activity that correlates with these ERP components became synchronized later, even though they may have started earlier (Patel, 2008). Second, there is a lack of efforts to prove that these components are in fact specific to language.

6. Hagoort's new insight of the known language related ERP components

In spite of the large quantity of neurophysiological language data reported, the neuroscience language processing approach has not yet merged with linguistic and psycholinguistic theories (Brown & Hagoort, 1999). The fact that grammatical sentences with more processing requirements correlate with a very similar LAN to sentences with morpho-syntactic violations has not been entirely discussed in the literature. The way individual words are combined into an interpretation of a sentence is intensely debated in language comprehension literature (Friederici, 2002; Delong et al, 2005; Kutas et al, 2006; Hagoort, 2009).

Jackendoff (1999, 2002) proposes that the only remaining rule of grammar is unifying pieces at different levels of knowledge, syntax and phonology, as well as across levels. A similar approach is the Unification computational model (Vosse & Kempen, 2000) where in the online comprehension process, structural frames associated with the individual word forms incrementally enter the Unification work space. In this work space constituent structures spanning the whole utterance are formed by a unification operation. This operation consists of linking up lexical frames and checking agreement features between them. The unification process used in this theory differs from head-driven phrase structure grammar and lexical-functional grammar in being non-recursive and involving only feature unification based on activation level that given by recency and semantic effects (Vosse & Kempen, 2000).

Hagoort's (2003, 2005, 2007) language processing account based on Vosse and Kempen's model, advocates a new understanding of the known syntax related ERP components. He proposed that LAN emerges when the root node[8] of a syntactic building block (e.g., NP) cannot bind to another syntactic building block with an identical foot node (i.e., NP). Another possibility for a LAN to emerge is when the agreement check finds a serious mismatch in the grammatical feature specifications of

[8] Root nodes are the top layer nodes of the network (Hagoort, 2003).

the root and foot nodes of the lexical trees. LAN results from a failure to bind as a result of a negative outcome of the agreement check to find a matching node. P600 on the other hand is related to the time and effort needed to establish the unification links of sufficient strength for the whole utterance. The time to build up the required strength of unification links is affected by ongoing competition between alternative unification options) caused by syntactic ambiguity or by syntactic complexity and semantic influences. The amplitude of the P600 is modulated by the amount of competition. Competition is reduced when the number of alternative binding options is smaller or when lexical, semantic or discourse context biases the strengths of the unification links in a particular direction, thereby shortening the duration of the competition (Hagoort, 2003). Violations result in a P600 as long as unification attempts are made. Hagoort (2005) recently proposed the neurobiologically inspired MUC language processing model (Memory, Unification, Control). The Memory component refers to the mental lexicon, containing information about the syntactic frames of words. The Unification component binds lexically retrieved information into a higher level representation. The Control component has an executive control for turn-taking, for example the selection of a language in the case of bilingualism. In the MUC model, the left superior temporal cortex is responsible for storage and retrieval of syntactic frames. The left prefrontal cortex has a significant role in holding online and binding frames together, unifying the whole utterance (Hagoort, 2007).

Neither Hagoort's (2003) approach nor the Unification model (Vosse & Kempen, 2000) discusses explicitly the relationship between holding online and checking agreement feature processes. However, both imply that in terms of fostering the unification of the sentence, these are functionally inseparable operations. One of the crucial functions of the prefrontal cortex is to provide the computational stability to hold online and bind syntactic frames together that are stored in the left temporal cortex (Hagoort, 2003). This is a relevant point since LANs that were found on the main verbs for sentences with intra-sentential dependencies (Kluender & Kutas, 1993), when the first syntactic element (e.g., subject of the sentence) must be held online or activated when encountering the second element (e.g., verb). Very similar LANs were elicited by morphosyntactic violations (Kutas & Hillyard, 1983; Münte, 1993).

In line with different models implemented activation (Gibson, 2000; Just & Carpenter, 1992; MacDonald & Christiansen, 2002; Vosse & Kempen, 2000) the newly incoming words have a higher activation level than words appearing earlier in the utterance. The first element of the intra-sentential dependency needs to remain active until or get reactivated when the second element is encountered. Only then can they be linked together by matching agreement features (Vosse & Kempen, 2002). This thesis will argue that it is also possible however that the linking of constituent elements can occur without any abstract feature matching (e.g. agreement features) but only based on statistical learning.

Two different neurophysiological models of language comprehension have been introduced in the literature (Friederici, 2002; Hagoort, 2003). The first one, presented earlier in this thesis, is a strongly generativist approach distinguishing between solid comprehending stages (Friederici, 2002). This model can be criticized based on how difficult it is to specify a temporal sequence of processing stages based on the ERP components. One can, on the other hand, describe the synchronization state of neurons contributing to the surface activity of a component. However, neurons correlating with later e.g. P600 component might have started to fire earlier than LAN or N400 (Patel, 2008). In contrast, the second approach (Hagoort, 2005) relies on a computational model (Vosse & Kempen, 2000) without explicit syntactic rules or discrete processing stages. This model emphasizes the significance of the parallel processing mode and gives new insight into the known ERP components. According to Hagoort (2003), LAN can arise when a node cannot find another appropriate node to bind to or when there is a mismatch between the linkable nodes. P600 on the other hand expresses the effort and time needed to reach the sufficient strength of unification connections (Hagoort, 2003).

7. Aims

Processing uncoupled constituent elements occurs in a single distributed network, where constituent relationship of elements emerges through encoding statistical associations that exist in the world. There are no separate mechanisms like WM and rule system or WM and knowledge but all of these characteristics are implemented in one network.

The first two studies investigate the link between WM and integration during sentence processing with the hypothesis that WM and integration are inseparable and implemented in one distributed network. Local Dependency Theory (Gibson, 2000) proposed that both integration and memory processes are two different processes relying on shared resources. This assumption was supported by behavioral data where longer reaction times were registered by complex sentences in comparison to simple sentences (Grodner & Gibson, 2005). However, it is not clear whether integration and memory processes are separable in a neurophysiological level. It is a valid question, since some theorists advocate that the frontal lobe function is to “hold online” and bind elements together (Mesulam, 2002; Hagoort, 2003, Hagoort, 2007). The aim of the first two experiments is to investigate the neurophysiological characteristics of the co-effect of integration and memory processes of constituent but uncoupled sentential elements in a novel mixed paradigm (Jolsvai et al, 2006). This design combines the experimental designs used in the syntactic ERP experiments (Friederici, 2002) with the paradigm used in ERP experiments manipulating working memory load of sentence comprehension (e.g. King & Kutas, 1995). The first two studies test the neurophysiological characteristics of integration and holding online processes with the hypothesis that these are not functionally separate processes representing different activation pattern of the same network (MacDonald & Christiansen, 2002) and therefore inseparable processes of sentence comprehension (Jolsvai et al, submitted).

The second set of the studies (Studies 3, 4, 5) test the relevance of making a sharp functional distinction between declarative memory and combinatory rule mechanism (Ullman, 2001). There is a consensus between theorists (Ullman, 2004; Williams & Hagoort, 2009) to permit domain general processes in the procedural (Ullman, 2001) or unification system (Hagoort, 2009). In the Declarative-Procedural (DP) model (Ullman, 2001) a general procedural system supports syntactic processing, and in the MUC model (Hagoort, 2005) a unification process can bind together representations other than language as well. Furthermore, a group of models assign different mechanisms to declarative memory and procedural system where WM is strongly linked to the procedural system. According to Ullman and Pierpont (2005):

"Working memory is strongly linked to brain structures that underlie the procedural system: Broca's area and premotor regions, including both lateral premotor cortex and the SMA (Ivry and Fiez, 2000; Smith and Jonides, 1999); the basal ganglia, (Menon et al., 2000); the cerebellum (Desmond and Fiez, 1998;) and parietal cortex, including the supramarginal gyrus (Ivry and Fiez, 2000).Working memory appears to involve the dorsal stream system (Hickok and Poeppel, 2000), and seems to be functionally related both to the retrieval of declarative knowledge and to procedural memory itself" (p.404).

Other neural network models advocate that the functional separation of memory and rule is dematerialized (Vosse & Kempen, 2000) and categories emerge through statistical learning (e.g. Rummelhart, 1999)

The experimental work presented in this thesis will contribute to this debate with providing evidence of permeable mechanisms of a single network as an opposing assumption of separated mechanisms of rule and declarative memory systems (Ullman, 2001). Studies 3, 4 and 5 investigate constituent musical phrases with the hypothesis that the processing of associative tokens with learned contingent[9] relationship correlate with similar neurophysiological responses than abstract rule processing; indicating that cohesivity of elements can be an emergent charactersitics of statistical learning (Rummelhart, 1989).

[9] The terms 'contingent' and 'learned contingency' will be used here according to Pavlov's (1927), terminology, therefore referring a learned relationship of two items, based on their co-occurrence.

8. Studies 1 and 2 – Updating and Unification

The influential work of Gibson (2000) proposes that sentence comprehension involves two different components each of which consumes neural resources. The first component is integration, which connects each incoming word to a prior word that it depends on in the sentence. The second is a memory cost or structural storage associated with keeping track of obligatory syntactic requirements as the online comprehension process takes place. This memory component is activated when a noun is presented in the sentence; then a verb encounter is predicted in order to form a complete clause. Structure building consists of looking up the current word in the lexicon and matching the categories (integration) of these lexical entries to the predictions (memory) in the structures built thus far that is very similar to Vosse and Kempen (2000) unification process. Both memory cost and integration cost are influenced by locality: the longer a predicted category must be kept in memory before the prediction is satisfied, the greater the cost for maintaining that prediction. Additionally, the greater the distance between an incoming word and the dependent to which it is attached, the greater the integration cost. According to Gibson (1998, 2000), the distance between constituent structure elements makes the processing harder. Gibson (1998) and Warren and Gibson (2002) defined distance in terms of the complexity of the intervening discourse structure. These studies calculate distance as the number of new discourse referents that are introduced between the endpoints of an integration, where a discourse referent is defined as a new discourse object (an NP) or a tensed verbal element. In a behavioral study, Grodner and Gibson (2005) contrasted simple sentences ("The nurse from the clinic supervised") versus complex sentences ("The nurse who was from the clinic supervised") and found longer reading times on the main verb, "supervised, " for the relative clause condition.

Gibson's assumption is potentially important in solving the paradox of two seemingly separated LANs. There is plenty of evidence showing that LAN is enlarged by subject-verb integration (e.g. morphosyntactic violations) (e.g., Palohati et al., 2005).

However, others such as Kutas and colleagues (2006) suggest that LAN reflects working memory processes with both 'look forward' and 'look back' functions. The former happens when a displaced constituent is seeking a subsequent coherent element (Kluender & Kutas, 1993; Kluender & Münte, 1998). The 'look back' function is triggered when either the current unexpected syntactic information must be reconciled and aligned with preceding information in the sentence (Kluender & Kutas, 1993; King & Kutas, 1995, Ueno & Kluender, 2003; Vos et al., 2001), or an anaphora precedes the information (van Berkum &, 1999; van Berkum et al., 2003). According to this assumption, LAN elicited by syntactic or morphosyntactic violation is not specific to syntactic processes but indexes a general WM process (Kluender & Kutas, 1993).

The DLT theory (Gibson, 2000) proposes that integration and memory costs are two different processes sharing the limited resources available. They provided behavioral evidence of longer RTs for complex sentences in comparison to the simple counter pairs (Grodner & Gibson, 2005). However, this behavioral study could not specify the characteristics of the contributing processes. This is relevant, because of the disputed functional nature of LAN. Variants of the original integration concept by Gibson (1998, 2000), "surprisal" (Halle, 1991) and expectation (Levy, 2001), were suggested as factors that influence the processing because the "comprehender" activates the next elements in the sentence based on probabilistic factors that are the result of previous experiences with the occurring language elements (MacDonald & Christiansen, 2002). This is an important point because all the 'syntactic experiments' manipulate violation. Lau (2006) suggests that instead of the intended syntactic processes, LAN reflects the degree of violation in expectancy based on the previous sentence context. In his study, greater expectancy elicited correspondingly larger LAN amplitude. In contrast, if syntax is an encapsulated operation and it is impermeable to meaning, the previous context could not affect it. This can be associated with the model (Vosse & Kempen, 2000) where the lexicon contains the words with their structural frames permitting other linkable structures based on semantic and probabilistic factors. Hagoort (2003, 2005) suggests that sentence comprehension occurs through the unification of the constituent sentence elements retrieved or activated from the memory/lexicon. The integration component (Gibson, 1998, 2000) can be paired with the unification process (Hagoort, 2003, 2009) as both are responsible for binding, or integrating, the constituent elements in a sentence.

Previous studies were inconclusive with respect the interpretations of LAN due to mainly the different experimental paradigms used. The first two studies in the thesis will try to settle this issue by manipulating both unification or integration (agreement violation) and memory cost or holding online variables.

The aim of Study 1 is to test whether integration and memory processes in DLT theory (Gibson, 2000) comprise two different neurophysiological processes relying on the same resources or rather they more likely represent activation pattern differences of one network (Martin-Loeches et al, 2006). Although the exact function of LAN is still disputed, it was suggested that it indexes syntactic processes (Friederici, 2002). Moreover, it was argued that LAN indicates the expectancy of a certain syntactic category (Lau et al, 2006). Others claimed that LAN represents WM processes (Kutas, 2006). Moreover, Gibson proposed that structure building consists of looking up the current word in the lexicon and matching the categories (integration) of these lexical entries to the predictions (memory) in the structures built thus far. Since the two assumptions (Friederici, 2002; King & Kutas, 1995) of the functional significance of LAN use very different paradigms it is hard to compare them. Therefore a novel a mixed-design paradigm (Jolsvai et al, 2006) was developed combining the two experimental designs.

9. Study 1 – Complexity

Hypothesis

Comprehending sentences containing uncoupled dependencies consists of looking up the current word in the lexicon and matching the categories (integration) of these lexical entries by the predictions (memory) in the structures built thus far. Memory and integration are inseparable processes indicated by two overlapping fronto-central negativities.

Participants

Fourteen native-Hungarian participants (4 males, 19-27 years old, M= 21, SD=2.09) took part in the study and they gave informed consent. All were paid subjects with normal or corrected-to-normal vision and they were right handed, according to Marion Annett's Handedness Questionnaire. None of the subjects had difficulty reading, or a history of either neurological or psychiatric disorders.

Materials

The grammatical violation in this study involved the mismatch of verbal agreement of number between the sentence's initial noun and the sentence's final intransitive verb. As the ERP responses were compared on the sentences' final verbs, the verbs were kept in singular in all sentences. Therefore, half of the sentences had plural nouns (violated sentences) and half of them had singular nouns (correct sentences). The choice of number- person violation was motivated by the fact that we wanted to keep the verb constants singular, manipulating the correctness of the sentences with the nouns only. Thus, syntactic and semantic information of the head noun had to be held in memory until the verb is encountered to unify them into a clause.

Correct sentence: A katona a sötét *üres szob*ában *ül*. / The soldier in the dark empty room sits.

Violated sentence: A katonák a sötét üres szobában ül. / The soldiers in the dark empty rooms sits.

The experiment also included a condition that manipulated the distance of subject-verb intra-sentential dependency. A set of Hungarian sentences were created similar to sentences Godner and Gibson (2005) used in their study.

The subject-verb dependency (The soldier sits) was decoupled by an intervening prepositional phrase: "sötét üres szobában/ in the dark empty room" or with an intervening clause: "akié (who[AN-POSS]) a fekete pöttyös sál/ to whom the black dotted scarf belongs".

Examples of sentences for each condition (four in total) are presented in Table 1. In Hungarian the relative pronoun, (3 singular) anaphoric possessive nominative case (akié) in the relative clause, must agree with the head noun; in the complex ungrammatical condition where the head noun was plural, the pronoun agreed with it, as such: A katonák/ The soldiers akiké (3 plural)/ to whom the fekete pöttyös sál/ a black dotted scarf belongs. Filler sentences were used to counterbalance all possible variations of sentences.

Simple Grammatical:
"A katona [noun, singular, 3rd person] a sötét üres szobában ül [verb, singular, 3rd person]. "The soldier in the dark empty room sits."
Simple Ungrammatical:
"A katonák [noun, plural, 3rd person] a sötét üres szobában ül. [verb, singular, 3rd person]". "The soldiers in the dark empty room sits."
Complex Grammatical:
"A katona [noun, singular, 3rd person] akié (who[AN-POSS]) a fekete pöttyös sál/ ül[verb, singular, 3rd person]. "The soldier to whom the black dotted scarf belongs sits".
Complex Ungrammatical:
"A katonák [noun, plural, 3rd person] akiké(who[AN-POSS]) a fekete sál ül." [verb, singular, 3rd person]". "The soldiers to whom the black dotted scarf belongs sits".

Table 2. Example of each sentence in each condition.

The stimuli consisted of a list of 400 visually presented Hungarian sentences. Of these sentences 320 were critical for the study. The remaining sentences were fillers with similar sentences with all possible combinations (40 sentences were with plural head noun and embedded relative pronoun and main verb; in another 40 sentences the main verbs were plural and the head subject was singular) and were used to distract subjects from the critical manipulation in order to minimize conscious strategies. The filler sentences were not included in the further analyses.

All of the sentences contained five words, where the first word was always a noun, and the last one was always an intransitive verb. One half of the sentences had simple syntactic structure, i.e., there were two adjectives and a locative adverb between the subject and the verb. The remaining sentences were complex, containing an embedded clause in between the subject and the verb. In addition to complexity, we also manipulated grammaticality. Half of these sentences were Grammatical (subject-verb agreed in number) and the other half Ungrammatical. 80 of the critical sentences were complex grammatical; 80 were simple grammatical; 80 were complex ungrammatical; and 80 were simple ungrammatical. Table 2. shows each sentence type.

Procedure

All subjects were presented with a list of all the sentences, mixing all four conditions (plus the filler sentences), in random order in a total of 10 blocks (40 sentences each). Sentences were presented in a word by word manner. Subjects were presented with a practice block of an additional 20 sentences in the beginning of the study. In 30% of all the sentences two types of test questions were asked, in a random manner, to keep the subject focused. The first type of test question was related to grammaticality (15% of the trials) ("Was this sentence correct?"). Since the main difference between complex and simple sentences were coded in the intervening material between the subject-verb dependency, and the grammaticality question could be executed without reading the intervening sentence material, subjects were presented with an additional test question (15% of the trials). This second type of test question pertained to whether one word (always a noun) occurred or not in the intervening prepositional phrase or embedded clause (e.g., "the room"). If this word did or did not occur in the sentence, subjects answered 'Yes' or 'No' respectively. After each block, feedback was given to the subjects about their hit rates.

Subjects were tested individually in a dimmed, sound-attenuating booth. They were seated in a comfortable reclining chair, and instructed to move as little as possible.

Participants were told that they would be presented with a series of sentences. They were asked to read each sentence. All sentences were self-paced, when the subjects were ready to start they pushed a button and the sentence presentation started. At the beginning of each trial a fixation cross was displayed for 400 ms. Sentences were presented in the center of the computer screen, word by word in white lower case letters (Arial, font size 20) against a dark background. Viewing distance was approximately 100 cm and the stimuli subtended a visual angle of 3° horizontally and 0.5° vertically. Each word was presented for 400 ms, followed by a blank screen for 400 ms before the next word appeared. The testing session began with a short practice block. The experimental trials were presented in 10 blocks of approximately 8 minutes each. Subjects were given short breaks between the blocks.

Electrophysiological Recordings

The acquisition of bioelectrical signals was performed by means of a 32-channel electroencephalogram (EEG) recording system (BrainAmp amplifier and BrainVision Recorder Software, BrainProducts GmbH). Recording electrodes were mounted in an elastic electrode cap according to the 10-20 electrode placement system at the following positions Fp1, Fp2, F9, F7, F3, Fz, F4, F8, F10, FC5, FC1, FC2, FC6, T9, T7, C3, Cz, C4, T8, T10, CP5, CP1, CP2, CP6, P7, P3, P4, P8, O1, O2, P9, P10.The reference electrode was at Pz, whereas the ground was placed between Fz and Fpz, on the midline. Eye movements were monitored on sites Fp1 and Fp2 and on an additional vertical electrode placed under the right eye. The contact impedance was kept below 5 kΩ at each electrode site. The recording bandwidth was 0.01-70 Hz and the sampling rate was 500 Hz.

Data Analysis

The EEG was re-referenced to the common average reference of all electrodes (Picton et al., 1998). Offline analysis of the data was done with Neuroscan 4.3 Software using the following steps: after applying a low-pass filter (15 Hz, 24 dB/oct), the EEG data were segmented into epochs of 950 ms (-150 ms before onset and 800 ms after onset); next the data were baseline-corrected on the entire epoch. Artifact correction criteria were set to ± 75 μV. The segmented epochs were then averaged by each condition and baseline corrected to the 150 ms pre-stimulus. The ERPs time-locked to the sentences' final main verbs were analyzed and compared. The main verbs were always in singular, as we controlled the number-person correctness by varying the number person suffixes on the sentence's first nouns. Global Field Power (GFP) was used to define the latency of component peaks. Analyses of the amplitudes of LAN between the 389-419 ms time window, over F3, Fz, F4, FC1, FC2, C3, Cz, C4 electrode sites and P600 between 550-650 ms, over P3, Pz, P4 were used to compare the different conditions. To compare component amplitudes, three-way repeated measures analysis of variance (ANOVA) with Statistica 7 software was conducted with factors of Complexity, (Simple vs. Complex), Grammaticality (Grammatical vs. Ungrammatical), and Electrode (F3, Fz, F4, Fc1, Fc2, C3, Cz, C4 or P3, Pz, P4). To compare latency differences, Fz and Cz (for LAN) and P3, Pz, P4 (for P600) were analyzed as the latencies of the components were defined by GFP. The Greenhouse-Geisser adjusted univariate test was used to avoid possible violation of compound symmetry and sphericity assumption. Tukey HSD post-hoc analyses were performed to analyze further the significant effects.

A reference-free measure of the current scalp density (CSD) was also used. Maps showing CSD were computed from the mean amplitude waveforms, corresponding to the peak latency of the component for each group, in each condition, as defined by the Global Field Power. The CSD analysis, an estimate of the second spatial derivative of the voltage potential using the spherical spline surface Laplacian algorithm (Perrin et al., 1989), was performed with BESA 2000 FOCUS software (Grafelfing, Germany) to sharpen the differences in the scalp fields. This method uses information from all electrodes on the head, providing a more accurate estimation for showing the scalp areas where the current either emerges (sources) from the brain into the scalp or enters (sinks) from the scalp into the brain.

Results

Behavioral data

Accuracy was very high; 94% (SD=1.0) of the sentences were correctly classified in the well-formedness judgment task. Overall accuracy was also high; 93% (SD=0.6) of the intervening test questions were answered correctly, suggesting that subjects were reading and comprehending the visually presented sentences.

ERPs Results

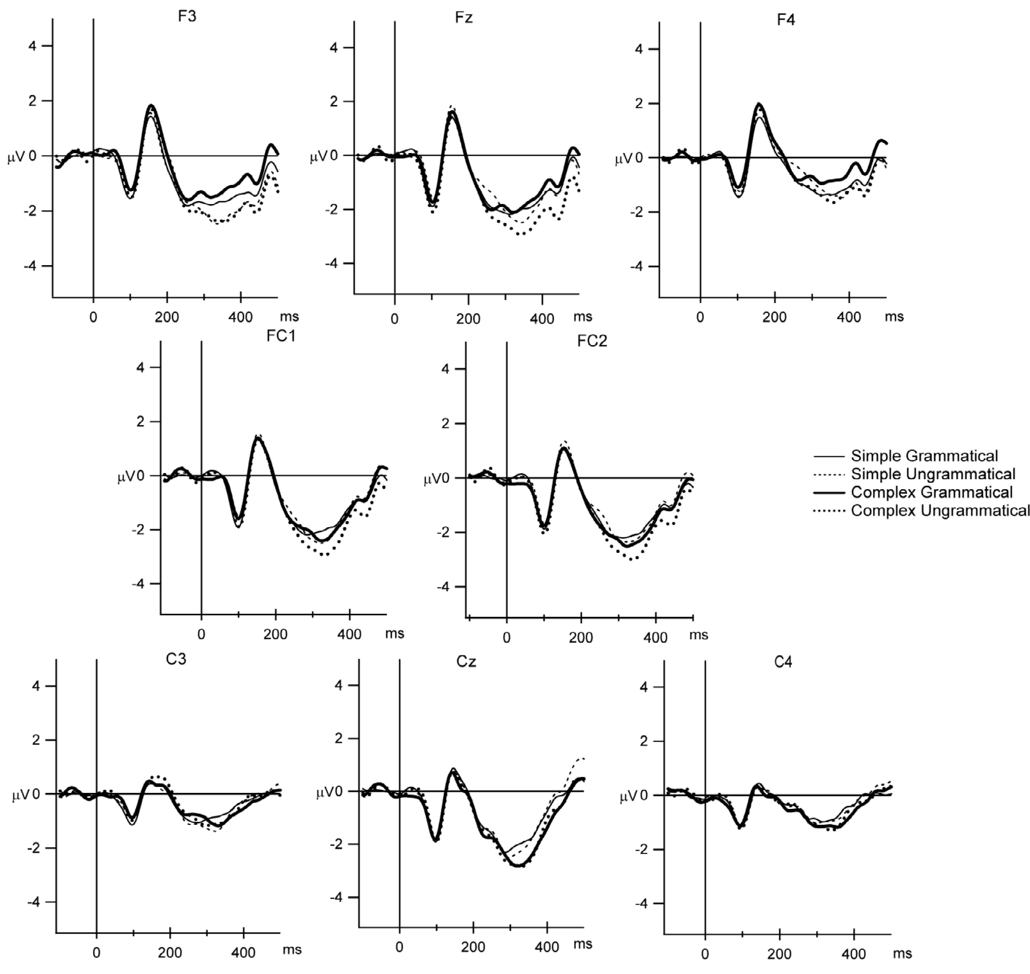


Figure 1. Averaged ERPs from the onset of the critical item (verb onset marked by vertical line) up to 600 ms for the frontal and central thereafter in all four critical conditions for a selection of eight electrodes. Negativity is plotted downwards.

Anterior Negativity (AN) between 300-500 ms

Figure 1 displays the grand average ERPs in all four critical conditions. The ANOVA revealed a significant main effect of Electrode [F(2.6, 34)=7.896, p<0.001]; LAN was larger at F3, Fz, Fc1, Fc2 Cz than at F4, C3 and C4. The Grammaticality violation elicited a significant main effect, [F(1, 13)=15.417, p<0.001]. Moreover, Ungrammatical sentences elicited a larger LAN at F3, Fz, and Fc1 and Fc2 than grammatical sentences as revealed by a Grammaticality x Electrode interaction, [F(3.09, 40.29)=3.599, p<0.001]; The Complexity condition failed to reach a significant main effect, [F(1,13)=1.112, p=0.31]. However, Complex sentences elicited a larger negativity at Fc1, Fc2 and Cz than Simple sentences, revealed by a Complexity x Electrode interaction F(2.54, 33.08)=4.40, p<0.001. LAN latency was longer for the complex sentences revealed by a Complexity main effect [F(1, 13)=4.72, p=0.04]. Complex sentences elicited a longer LAN latency, main effect of Complexity [F(1, 13)=4.72, p=0.048]

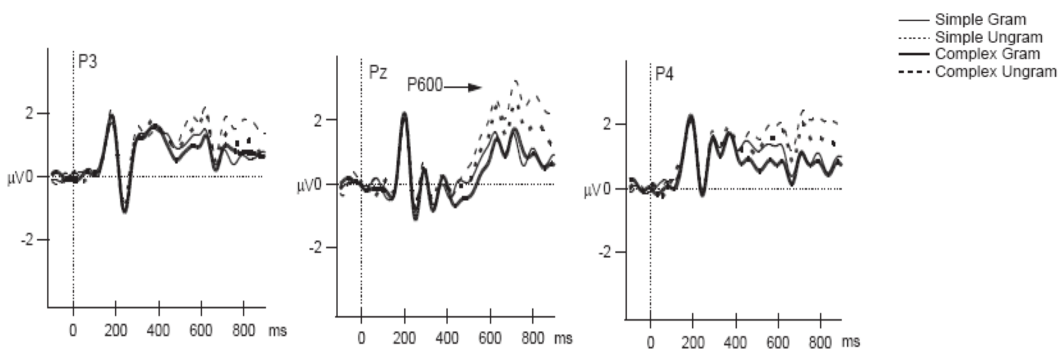


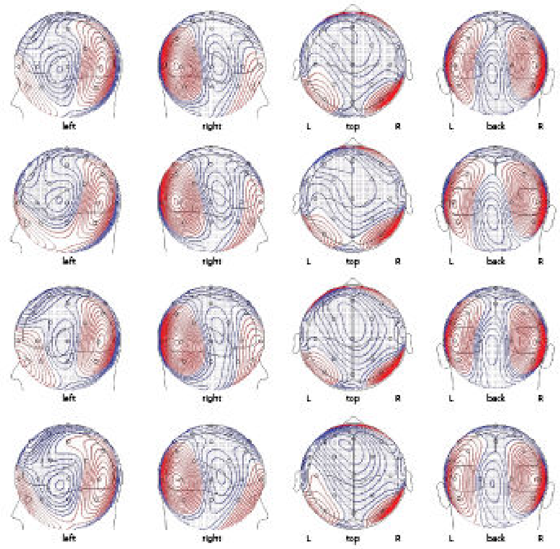
Figure 2. P600 elicited by Ungrammatical sentences. Negativity is downwards.

P600

Ungrammatical sentences elicited a larger P600 amplitude than grammatical sentences, as revealed by the Grammaticality main effect F(1, 13)=30.653, p<0.001. P600 was the most pronounced at Pz (Grammaticality x Electrode interaction F[1.63, 21.26]=7.608, p<0.004.

Simple sentences elicited a longer P600 latency at P1 than complex sentences, revealed by a Complexity x Electrode interaction [F(2, 26)=3.591, p=0.041].

Scalp Topography



Simple Grammatical
385 ms

Simple Ungrammatical
393 ms

Complex Grammatical
385 ms

Complex Ungrammatical
399 ms

Figure 3. The CSD maps of the AN in all four conditions CSD [0.03 µV/ cm2]. Negative voltage and current sinks are shown in blue and positive voltage and currents sources are shown in red (expressed in µV /cm2). Maps showing CSD were computed from the mean amplitude waveforms, corresponding to the peak latency of the component in each condition as defined by the GFP.

Figure 3 displays CSD maps for each condition. This method uses information from all electrodes on the head, providing a more accurate estimation for showing the scalp areas where the current either emerges (sources) from the brain into the scalp or enters (sinks) from the scalp into the brain. Positive and negative current densities in these maps are dependent on the orientations of the neural generators; however their meaning cannot be clearly defined (e.g. Picton et al., 1995,). Nevertheless, there is no need to interpret such information because one is only concerned with when different conditions are characterized by different patterns of positive and negative current densities (Murrey et al, 2008). Thus, the CSD maps show clearly that the different conditions did not elicit different patterns of positive and negative current densities.

Summary of Study 1 Results

Integration of agreement features between the subject-verb structure elicited a frontal- fronto-central negativity followed by a parietal positivity. Manipulating the distance as defined by Gibson (2000) engendered a fronto-central negativity. Although the Grammaticality elicited a more frontal-fronto-central negativity than Complexity that enlarged a fronto- central negativity, they both consistently correlated with a negativity over the fronto-central electrode sites (Fc1 and Fc2). CSD maps showed that the different conditions did not elicit different patterns of positive and negative current densities. It might indicate one network fostering the integration and memory process with different activation patterns. It is also possible that these results indicate shared resources of two parallel processes with a theoretical starting point namely that (L)AN[10]s exclusively reflect WM processes (Kluender & Kutas, 1983) or with a hypothesis that the fronto-central memory negativity at least partially reflects processes indexed by N400 in other experiments (Kutas et al, 2006) or the measurable scalp activity overlap with processes reflected by LAN and N400 (Günter et al, 2000).

The aim of the second study is to investigate further the processing characteristics of the co-occurrence of the integration process and the fostering updating/holding online memory processes. Due to their interdependence, Studies 1 and 2 will be discussed jointly.

[10] The negativity between 300-500 ms is usually called LAN. However, its left lateralization is disputed and some authors refer it as Anterior Negativity (AN)(Hagoort, 2003). I will call it LAN when I refer other studies which name it LAN and theories of LAN, otherwise I will refer it as AN since my data indicate this label.

10. Study 2 – Time Delay

Memory and integration processes are influenced by the distance of the constituent elements defined by the complexity of the intervening sentence elements (Gibson, 1998; 2000; Grodner and Gibson, 2005). In the previous study (1), two partially overlapping ANs were elicited. A fronto-central negativity was engendered for the memory load (complexity) and a frontal- fronto- central negativity was elicited by the agreement feature integration. Gibson (2000) suggests that distance is defined by the complexity of the intervening elements of the uncoupled dependency referring to the number of new discourse references introduced before an integration point. As an alternative explanation, based on the connectionist models (Vosse & Kempen, 2002; MacDonald and Christiansen, 2002), it is plausible to predict that pure time delay between the separated dependants affects sentence comprehension. According to Gibson, distance is calculated as the number of new discourse referents that are introduced between the endpoints of an integration, where a discourse referent is defined to be a new discourse object (an NP) or a tensed verbal element. In a behavioral study, Grodner and Gibson (2005) contrasted simple ("The nurse from the clinic supervised") vs. complex ("The nurse who was from the clinic supervised") sentences and found longer reading times on the main verb, "supervised" for the relative clause condition. However, it is a relevant question whether distance as defined by Gibson could be replaced by only time delay. The simple sentence in Grodner and Gibson's study contained three words in between the subject-verb dependency; however, in the complex condition there were five words. Would longer reading times on the main verb be the consequence of the longer time and fading of the activation of the first constituent or as Gibson states as a consequence of intervening uncompleted discourse elements?

Some theorists suggested that time plays a crucial part in the fading of representations (e.g. Cowan, 1999, Baddeley & Lodge, 1989; Just & Carpenter, 1992). Hebb (1949) and current connectionist models of language processing (MacDonald & Christiansen, 2002; Vosse & Kempen, 2000) proposed that simply time delay influences the processing. The activation level of formerly presented elements decays and must be updated/activated when they are separated from their structural constituents. Fading or decay begins immediately at the presentation point of the element in the sentence. In Vosse and Kempen's (2000) model, the more recently entered elements in the unification space have more activation levels. In other words, the activation level of earlier entered elements is lower because it starts to decay immediately after the element is presented. Similarly, in MacDonald and Christiansen's (2002) model time has a crucial role in terms of the activation decay of nodes. According to these models, working memory, processing capacity are not independent variables as these emerge from network architecture and experience. This is one of the main dividing points between these connectionist concepts and the assumptions of Baddeley (1986) or Just and Carpenter (1992). However, the latter theories also emphasize the importance of time decay. The dissimilarity between these two neural network models (Vosse & Kempen, 2000; MacDonald & Christiansen, 2000) is that MacDonald and Christiansen (2002) built a standard feed-forward connectionist network with 31 input and output units as well as 60 hidden and context units. In contrast, Vosse and Kempen (2000) were operating with the activation and plausibility based competition of the rival links between the three ordered trees of words. The unification process used in this theory differs from head-driven phrase structure grammar and lexical-functional grammar in being non-recursive and involving only feature unification based on activation level that given by recency and semantic effects (Vosse & Kempen, 2000).

Since two partly overlapping ANs were elicited by memory and integration processes (Jolsvai et al, submitted) when distance was defined according to Gibson (2000), the second study was designed to test further the independency of the memory/holding online and integration/unification processes. The general aim of the second study is to test the neurophysiological characteristics of intra-sentential dependency divided by only time delay. The real exciting question is whether processing uncoupled elements without intervening sentence elements correlate with one integration negativity or rather one memory negativity. The former prediction would follow the DLT theory (Gibson, 2000) since it implies that the comprehension of a sentence relies on the integration process in the lack of additional uncompleted discourse referents; memory processes involved in predicting the subsequent syntactic category are not burden.

However, in case of integration and memory processes are representing only activation pattern differences of the same single network, the uncoupled dependency will correlate solely with activation decay related memory negativity.

Hypothesis

Subject-verb dependency separated by pure time, lacking any additional uncompleted discourse references, will obtain one memory LAN, since memory and integration are inseparable and cohesivity of elements emerge through statistical learning.

Methods

In this study a mixed design paradigm was implemented by manipulating subject-verb agreement integration without additional uncompleted referents, and manipulating the distance of intra-sentential dependency (time delay) as well.

The grammatical violation involved the mismatch of the verbal agreement of number disagreement between the sentence's initial noun and the sentence's final intransitive verb. Since the ERP responses were compared on the sentence's final verbs, they were kept in singular in all sentences. Half of the sentences had plural nouns (violated sentences) and half of them had singular nouns (correct sentences). Number-person violation was chosen because we wanted to keep the verb constants singular and manipulate the correctness of the sentences with the nouns only. Syntactic and semantic information of the head noun had to be held in memory until the verb was encountered and they could be unified.

Correct sentence: A katona *ül.* / *The soldier sits.*

Violated sentence: A katonák *ül. / The soldiers sits.*

The study also included a condition that manipulated the distance of subject-verb intra-sentential dependency. The subject verb dependency (The soldier sits) was decoupled by 1) short time delay ISI=800 ms or 2) long time delay ISI=2800ms.

Participants

Twenty right-handed, native Hungarian participants (1 male, average M= 21.6 years; SD= 3.99) took part in the study. Participants gave informed consent. Subjects were paid for their participation. Five subjects were eliminated from the data due to excessive movement artifact, thus 15 subjects remained in the final analyses. All subjects had normal or corrected-to-normal vision and were right handed, according to Marion Annett's Handedness Questionnaire. None of the subjects had difficulty reading, or a history of either neurological or psychiatric disorders.

Materials

The stimuli consisted of a list of 600 visually presented Hungarian sentences. Of these sentences 400 were critical for the study. The remaining sentences were fillers with similar sentences (verbs were plurals and nouns were singular or plurals) and were not included in the further analysis. All of the sentences contained a definite article ("the"), a noun, and a verb. Half of the sentences were short, i.e., there was a short time delay between the subject and the verb (ISI=400 ms). The other half were long sentences with a long time delay (ISI=2800 ms). In addition to time delay, we also manipulated grammaticality. Half of these sentences were grammatical (subject-verb agreed in number) and the other half ungrammatical. The critical sentences were short grammatical (80); short ungrammatical (80); long grammatical (80); long ungrammatical (80).

Sentences were presented word by word in a random order for a total of 12 blocks, with an additional first practice block with 20 sentences. To make sure that the subjects read the sentences properly, in 15% of the combined blocks test questions were presented randomly. The test questions were related to grammaticality ("Was this sentence correct?") and the participants pressed a Yes or No button. After each block we gave feedback to the subjects about their hit rates.

Procedure

Subjects were tested individually in a dimmed, sound-attenuating booth. They were seated in a comfortable reclining chair, and instructed to move as little as possible. Participants were told that they would be presented with a series of sentences. They were asked to process each sentence for comprehension. All sentences were self-paced, when the subjects were ready to start they pushed a button and the sentence presentation started. At the beginning of each trial a fixation cross was displayed for

400 ms. Sentences were presented on the center of the computer screen, word by word in white lowercase letters (Arial, font size 20) against a dark background. Viewing distance was approximately 100 cm and the stimuli subtended a visual angle of 3° horizontally and 0.5° vertically. Each word was presented for 400 ms, followed by a blank screen for 400 ms short condition and 2800 ms for long condition before the next word began. The testing session began with a short practice block. The experimental trials were presented in 12 blocks of approximately 8 min each. Subjects were given short breaks between the blocks.

Electrophysiological Recordings

The acquisition of bioelectrical signals was performed by means of a 32-channel electroencephalogram (EEG) recording system (BrainAmp amplifier and BrainVision Recorder Software, BrainProducts GmbH). Recording electrodes were mounted in an elastic electrode cap according to the 10-20 electrode placement system at the following positions Fp1, Fp2, F9, F7, F3, Fz, F4, F8, F10, FC5, FC1, FC2, FC6, T9, T7, C3, Cz, C4, T8, T10, CP5, CP1, CP2, CP6, P7, P3, P4, P8, O1, O2, P9, P10.The reference electrode was at Pz, whereas the ground was placed between Fz and Fpz, on the midline. Eye movements were monitored on sites Fp1 and Fp2. The contact impedance was kept below 5 kΩ at each electrode site. The recording bandwidth was 0.01-70 Hz and the sampling rate was 500 Hz.

Data Analysis

Since the behavioral data set was very limited by test questions representing 15 % of the whole sentence set, the behavioral data was simply used as selective criteria instead of being analyzed. Only subjects who performed a hit rate between 95-100 % were kept in the ERP analysis.

The EEG was re-referenced to the common average reference of all electrodes (Picton et al., 1998). Offline analysis of the data was done with Brain Vision Analyzer Software, BrainProducts GmbH comprising the following steps. After applying a band-pass filter (0.5-30 Hz, 24 dB/oct), the EEG data were segmented into epochs of 900 ms (-100 ms before onset and 800 ms after onset). Artifact correction criteria was set to ± 60 µV. The segmented epochs were then averaged by each condition and base-line corrected to the 100 ms pre-stimulus. The ERPs time-locked to the sentence's final main verbs were analyzed and compared. The main verbs were always in singular, as we controlled the number-person correctness with varying the number person suffixes on the sentence's first nouns. The grand means were used to define the latency of component peaks by visual inspection. To compare amplitude differences, three way repeated measures analysis of variance (ANOVA) with Statistica 6 software was conducted with factors of Time, (Short vs. Long), Grammaticality (Grammatical vs. Ungrammatical), and electrode (F9, F7, F3, Fz, F4, F8, F10, C3, Cz, C4 or P3, Pz, P4). The Greenhouse-Geisser adjusted univariate test was used to avoid possible violation of compound symmetry and sphericity assumption. Tukey HSD post-hoc analyses were performed to analyze further the significant effects. Time window was set for the statistical analysis between 280–320 ms for frontal LAN and 380–420 for central LAN.

Results

Frontal tendency of memory LAN

Figure 1. displays the grand mean ERP waveforms at Fz in each condition. LAN was the most negative at the left frontal electrode sites (main effect of electrode, [$F(2.4; 33.63) = 6.226$; $p = 0.003$].) (p always <0.035). LAN was larger for long time delay at F3 and Fz (Time x Electrode interaction [$F(2.4; 33.65) = 7.957$; $p = 0.008$]), although this effect was not significant according to the post hoc test. At the same time, however, at F4 and F8 electrode sites LAN became more positive for the time delay $p < 0.014$.

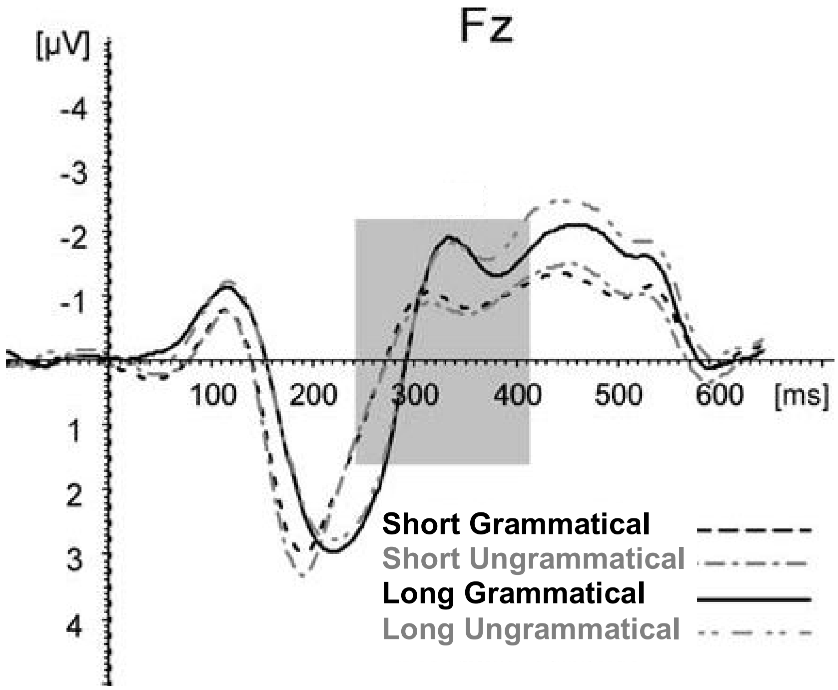


Figure 4. The left frontal negativity elicited in the four conditions. Negativity is upwards.

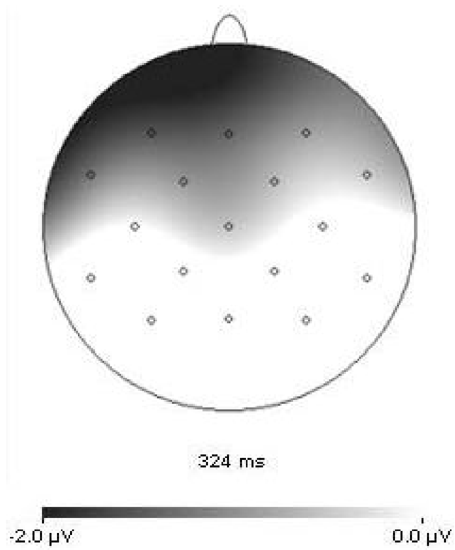


Figure 5. The Voltage map of frontal LAN at 324 ms.

Central Negativity

Long time delay between the two decoupled contingent sentence elements elicited a larger Central LAN (Time main effect, F(1.14)=9.072; p=0.009). LAN was largest at central electrode sites (Electrode main effect F(2, 28)=10.318; p<0.001), although this effect was not significant according to the post hoc test. A long time delay elicited a larger LAN at all electrodes (Time x Electrode interaction, $F(2, 28) = 3, 789$; $p = 0.034$) but was only significant at Cz ($p = 0.009$).

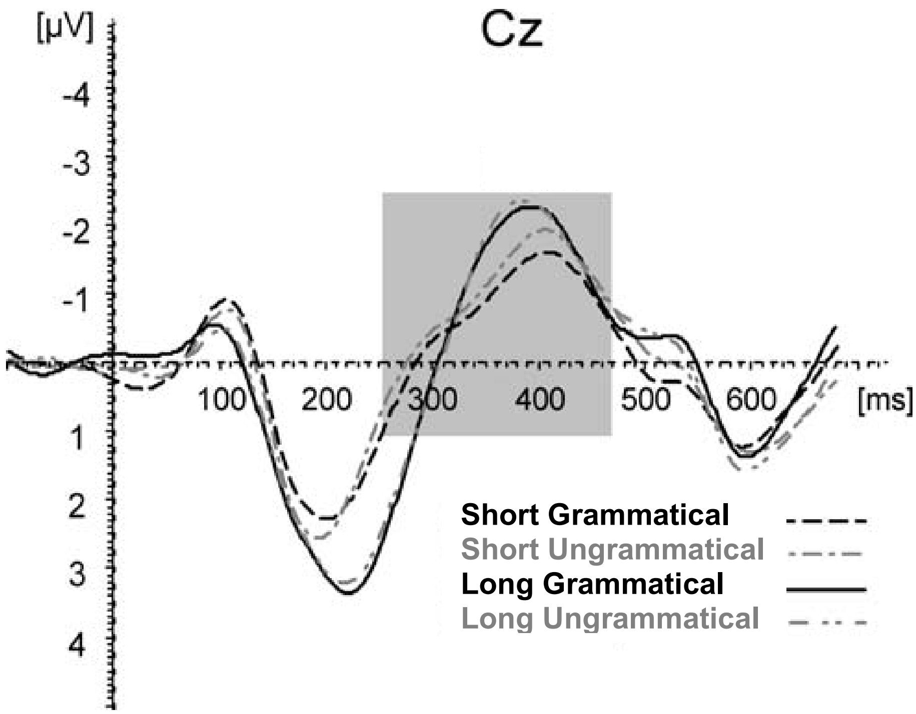


Figure 6. Central negativity elicited by Long time delay peaked at 404 ms and had a Cz maximum. Negativity is upwards.

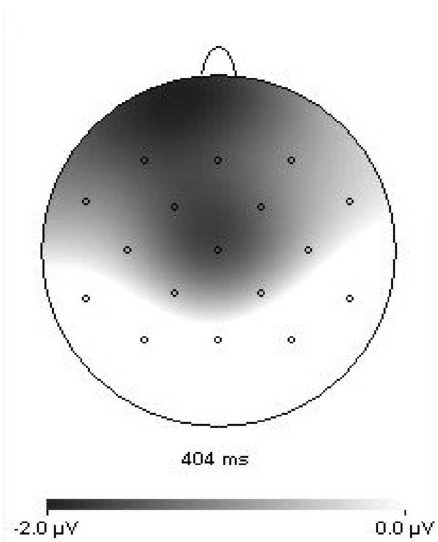


Figure 7 displays the Voltage map of central negativity at 404 ms.

Summary of Study 2 Results

Manipulating sentence comprehension by the time delay of constituent elements and agreement features engendered one central memory negativity. In spite of the fact that the frontal anterior negativity did not reach statistical significance, its presence was obvious. Moreover, it had also a tendency to be enlarged by the time delay. It is possible that these seemingly two different negativities index one single process that starts to be engendered by long time delay, but only around 400 ms when reaching the central sites are the involved neurons synchronized enough to reach statistical significance. Since subjects comprehended the sentences or in other word they integrated the constituent elements (performed a hit rate between 95-100%) these results indicate one neural network instead of separate processes.

Exploring Study 1 results further: Global Dissimilarity

To identify potential differences in the topographic distribution of the components elicited by memory and integration processes in Study 1, an index of Global Dissimilarity (GD) (Lehmann & Skrandies, 1984) was additionally obtained to test of the independence of the fronto-central memory and frontal-fronto-central integration negativities engendered in Study 1. While Global Field Power (GFP) (Lehmann & Skrandies, 1984) provides information on average across the electrode montage how strong a potential is being recorded, GD is an index of configuration differences between two vectors (each of which represents one electric field). The Global Dissimilarity index was obtained by calculating the square root of the mean of the squared differences between all corresponding electrodes, after normalizing the data by dividing the mean voltage by its own GFP (Lehmann & Skrandies, 1984; Murrey et al, 2008). The global dissimilarity provides an index of the degree of topographic similarity between two electrical fields. Numbers closer to 2 indicate topographic incongruence and numbers closer to 0 indicate topographic homogeneity.

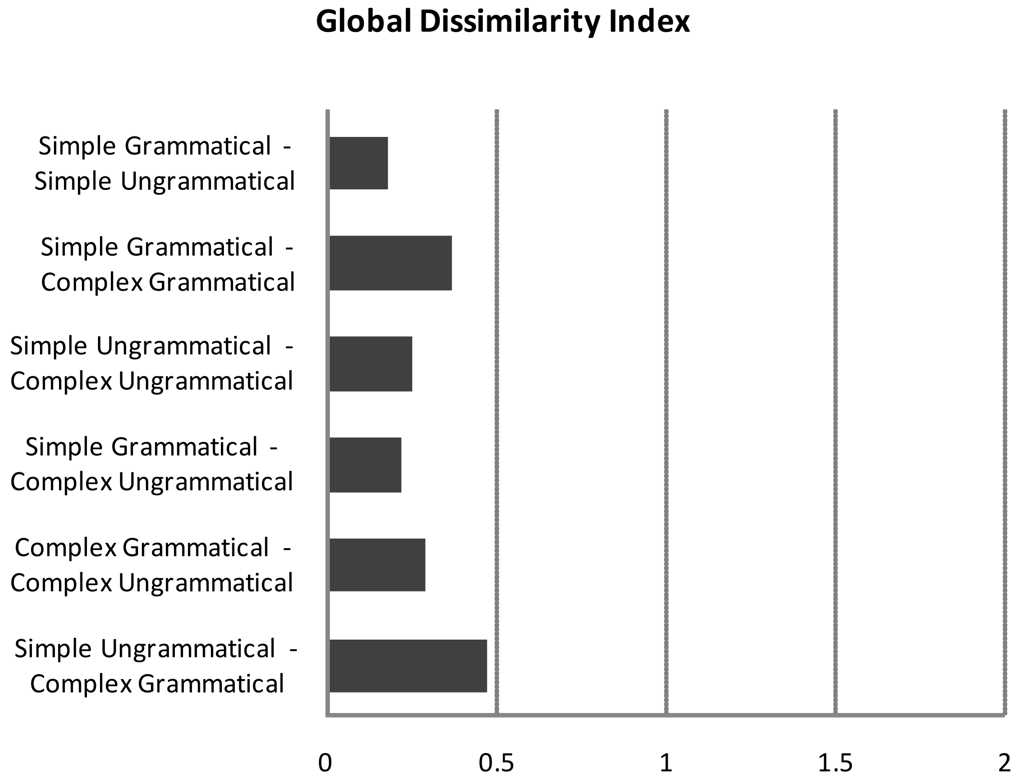


Figure 8. Global Dissimilarity index of the Anterior Negativities elicited in Study 1, conditions compared pairwise. Numbers closer to 2 indicate topographic incongruence and numbers closer to 0 indicate topographic homogeneity. All the Global Dissimilarity indexes were closer to 0, indicating topographic homogeneity between the anterior negativity effects in all compared conditions.

Figure 8 displays the Global Dissimilarity indexes for the Anterior Negativities (AN) evoked in each condition. Pairwise comparison of the negativities elicited by the different conditions resulted in GD values smaller than 0.5 in all cases. These results indicate topographic homogeneity between each compared electrical field. Analysis of the pattern of topographies between the six possible combinations of pairwise comparison of Grammaticality (Grammatical-Ungrammatical) and Complexity (Simple-Complex) conditions confirmed that negativities engendered by agreement feature violation and distance (defined by Gibson, 2000) were evoked in a common neural network. GD index allows assembling additional information beyond what is available from canonical waveform analysis thereby helping neurophysiological interpretation of a given data (Murrey, et al, 2008).

11. Discussion of Studies 1 and 2

Longer reading times were measured on the main verbs of complex sentences compared to simple ones (Gibson & Grodner 2005). However, this approach provides only a poor characterization of the integration and memory processes. Neurophysiological measurements enable a better understanding of the psycholinguistic processes, combining the two different experimental designs used for syntactic LAN (Friederici et al., 1996) and WM LAN (Kluender & Kutas, 1993). The aim of the first two studies was to investigate the neurophysiological characteristics of the co-existing integration and memory processes of uncoupled subject-verb dependencies. The studies tested the hypothesis that memory and integration are functionally inseparable and representing activation pattern differences of the same single language capacity network (MacDonald & Christiansen, 2002). A frontal-fronto-central negativity was elicited by the integration/unification (checking agreement features, Vosse & Kempen, 2000) and a partially overlapping fronto-central negativity was elicited by manipulating the memory variable in Study 1.

The main findings of the first two studies that negativities enlarged by memory and integration variables were overlapping at the fronto-central (Fc1 and Fc2) electrode sites and they did not indicate different neural mechanisms based on the CSD maps and GD index but rather activation pattern differences of the same single network (Jolsvai et al, submitted).

Moreover, the results of Study 2 indicate that in the case of simple subject-verb sentences, memory operations alone are suffice for comprehension without the need for an increased or separate integration processes. The memory process can unite the subject-verb pair alone, especially as the expectancy of a verb after a subject is the most frequent and expected in a natural language (Gibson, 1998). Instead of an auxiliary integration process in Study 2, the frontal tendency as well was increased by time delay and not integration, although it did not reach a significant level. In study 1 the two overlapping negativities imply a quantitative activation difference of the same network instead of indexing two separate processes. This hypothesis was further supported by the GD indexes. These studies emphasize the importance of using different calculations, since CSD maps and GD index can help to broaden views of the interpretation of surface-recorded event-related data. Calculating CSD maps and GD index provide not only information regarding when experimental conditions differ, but also how conditions differ in terms of likely underlying neurophysiologic mechanisms (Murrey, 2008). CSD is a reference free measurement, based on Laplacian derivation. Therefore, CSD derivations are intrinsically based on spatial gradients in the electric fields at the scalp, contrary to ERPs. Global Dissimilarity (Lehmann & Skrandies, 1984) also a reference-free method and includes all the electrodes, similar to CSD. By using multiple methods one can identify modulations in the underlying sources of responses (topographic modulations) and attempt to step over the limitations and pitfalls of canonical waveform analyses. A related issue with the canonical analyses of voltage waveforms concerns the interpretations of condition x electrode interactions (as observed in this study as well) observed in an analysis of variance (ANOVA). This issue has been extensively treated by McCarthy and Wood (1985) who emphasized how this analysis cannot differentiate modulations in topography from modulations in amplitude when data are not first scaled.

The Unification MUC model (Hagoort, 2003; 2005; 2007) suggests that during sentence comprehension all involved words' structural frames get activated and start to decay immediately. In this model linking up the lexical items is governed by the activation and plausibility effects (Vosse & Kempen, 2000). According to this approach when a syntactic structure does not match in, e.g., number feature, the integration of the lexical frames is more difficult (that is, indexed by LAN) and the completion between the possible candidates takes longer indexed by a P600 (Hagoort, 2003). Other linguistic theories also suggest that words can be difficult to integrate syntactically into sentences when they are distant from their dependents or when they are syntactically unexpected (Narayan & Jurafsky, 1998; 2002). Lau et al (2006) provided neurophysiological evidence to support the expectation model, suggesting that LANs elicited in syntactic or morphosyntactic violations index expectations previously built by the context. In the MUC model, the left superior temporal cortex is responsible for storage and retrieval of syntactic frames. The left prefrontal cortex has a significant role in holding online and binding frames together, unifying the whole utterance (Hagoort, 2007). Even though Hagoort‘s (2003) approach and the Unification model (Vosse and Kempen, 2000) do not explicitly discuss the relationship between holding online and checking agreement feature processes, they imply that these are functionally inseparable operations, in terms of fostering the unification of the sentence. One of the crucial functions of the prefrontal cortex is providing the computational resources to hold online and bind syntactic frames together that are stored in the left temporal cortex (Hagoort, 2003). This is a relevant point since LANs that were found on the main verbs for sentences with intra-sentential dependencies (e.g. Kluender and Kutas, 1993), when the first syntactic element (e.g. subject of the sentence) is needed to hold online or reactivate when encountering the second element (e.g. verb), were very similar to LANs elicited by morphosyntactic violations (e.g. Münte, 1993). Since the newly incoming words have a higher activation level than words appearing earlier in the utterance, it is obvious that the first element of the intra-sentential dependency needs to be kept active until the second element is encountered when they can bind together with the checking agreement features. In the Unification MUC model (Hagoort 2003; 2005; 2007), the decay of the activation level of a word is increasing over time. As seen in the CSD maps for the two different LANs in Study 1, there was no sharp difference between the two processes of holding online the first constituent frame of the syntactic structure until integration (the morpho-syntactic agreement checking). These results are compatible with the Unification MUC model, since the holding online and the checking agreement features are not functionally separated components and may be originated by different but possibly partially overlapping generators. However, not only has the functional significance of LAN generated an intense debate in the literature, but neither the frontal distribution nor the left lateralization are consistent findings (e.g. Coulson, King & Kutas, 1998; Günter & Friederici, 1999). It is also a testable and intriguing question whether the fronto-central memory negativities found in studies 1 and 2 are at least partially overlapping lexical retrieval processes indexed by N400 (Kutas et al, 2006). It is plausible to hypothesize that the activation (retrieval) of newly coming items are also contributing the fronto- central negativities measured in Studies 1 and 2. This is one of the future experimental questions I intend to test further.

According to Hagoort (2003), the P600 reflects the time and effort needed to link up the whole utterance. In other words, P600 is related to the time it takes to establish the unification links of sufficient strengths among the syntactic frames. In the present study, P600 showed sensitivity only for the morphological processes. While in English, grammatical relations are expressed mainly by word order, in Hungarian morphology (particularly case marking and agreement) plays a similar role. That is why it is possible in Hungarian, with its more significant morphology, that local morphosyntactic processes are more articulated (Pléh, 1998) than in the word order-operated holistic processing that the structure of English requires. This can be an explanation why P600 in Study 1 was only enlarged by the number integration process and not by the complexity manipulation. Nevertheless, P600 is also a disputed component, as it has been found for syntactic violations in pairs with LAN (Friederici et al, 1996), just as was elicited by complex sentences (Kaan and Swaab, 2006) or semantic anomalies (van Herten, Kolk, Chwilla, 2004). Since in Study 1 the late positivity was only sensitive for Grammaticality and participants were asked to judge Grammaticality it raises the question whether P600 indexes a P300 elicited in cognitive tasks when subjects are forced to respond (Picton, 1992; Polich, 1998; Pritchard, 1981) since it has been suggested that the P300 reflects some aspects of stimulus categorization (e.g., Johnson & Donchin, 1980; Mecklinger & Ullsperger, 1993).

The results of Studies 1 and 2 indicate that the hypothesis tying (L)ANs to modular syntactic processing mechanism can be questioned. Although, there have published many studies providing evidence to the syntactic view of LAN (e.g. Friederici, 2002) these studies always used exclusively the violation paradigm. It is worth making an important point to emphasize that the use of violation paradigms in electrophysiological studies has been questioned (Thierry et al, 2003). One might suggest that such experimental designs merely reveal the neural correlates of general processes, such as some kind of error detection, attentional shifts or repair processes, rather than those relating to a targeted psycholinguistic process (Thierry et al 2003). In fact, there have been some studies raising the important question as to whether these negativities are language specific in nature at all. Hoen & Dominey (2000) claim that the LAN reflects the operation of a general sequence-processing capability, in which special symbols encode structural information that, when combined with past elements in the sequence, allows the prediction of successor elements. In this view, negativities elicited by violations in language studies reflects a general neurocomputational capacity for treating structural complexity in rule-governed sequences (Hoen & Dominey, 2000), which might include but not be limited to sentences.

12. Functionally different associative memory and rule systems?

Study 3: Attend, Study 4: Non-attend and Study 5: Control

A number of different accounts have been put forward in order to explain the dissimilarity of automatic rule mechanisms combining abstract categories and controlled declarative memory processes of tokens (e.g. Pinker, 1991; 1994; Ullman, 2001; Waters & Caplan, 1999; Friederici, 2002). Evidence supporting the dual system memory view (Ullman, 2001) is emerging from neuropsychological data. Amnesic patients had no difficulties classifying distorted dotted patterns of a prototype of a category (implicit task), they were not able to recognize if they had seen a particular pattern before or not. These data are interpreted as a proof of two different memory systems (Knowlton & Squire, 1993). According to Knowlton and Squire (1993) implicit memory system is in charge of category-level information, and the declarative memory is responsible for the recognition of old/new exemplars.

However, Nosofsky and Zaki (1998) challenged this interpretation by showing that simple parameter settings of memory sensitivity can lead to similar results confirmed by a connectionist model introduced by the authors. They developed a neural net model that can perform on categorization and recognition tasks similarly to the patients by a parameter change in memory sensitivity of one single memory system. Similarly, Nosofsky & Zaki intended to “simulate amnesia” in an experiment by manipulating the time delay between the characterization (procedural) and recognition (declarative) of different group of normal participants, using the same test material used by Knowlton and Squire (1993). They found that “the parameter change of memory sensitivity”: namely, long time delay produced similar results than was found in Knowlton and Squire’s (1993) study: while categorization was not significantly different, recognition was significantly worse in the long time delay group than in the short time delay group. Their exemplar model performed consistent to their experimental data, a parameter change in the memory sensitivity could reproduce the dissociative performance by one network (Nosofsky & Zaki, 1998).

There is an agreement (Hagoort, 2003; Ullman, 2001) that the neural underpinnings of integration process of a sequence of abstract categories (sentence) can handle representations from other domains, (Hagoort, 2003; Ullman, 2001). There is dissimilarity in models in relation to whether they assign declarative memory and the integration processes to two functionally different mechanisms (Ullman, 2001) or allocate them to one network of interconnected connections (MacDonald & Christiansen, 2000).

Studies violating rules of abstract linguistic categories (e.g. Friederici et al, 1996 word category violation or agreement violation) or music strings: e.g. Koelsch, (2005) key violations result in similar brain responses, namely anterior negativities. In both of these types of studies, relations between abstract categories (musical or linguistic) were violated. These results are in line with the DP model (Ullman, 2004) that claims that the procedural rule based system is functionally separated from the associative declarative lexicon. However, from a connectionist point of view (e.g. Vosse & Kempen, 2000; in press) the traditional distinction between rules and declarative memory is blurred.

According to Wallis et al (2001), abstract rules are set of laws that are not bound to a specific context or stimulus but rather can be retrieved and applied to familiar and novel situations alike. Following this definition of abstract rule, my starting point was to create a sequence of distinctive tokens such as irregulars (Pinker, 1991) or second language elements (Ullman & Pierpont, 2005) with an associative nature and test whether they elicit similar neurophysiological responses as a string of words with combinatory rules (Friederici, 2002).

Strings of familiar melodies were used without the lyrics (but here I refer to them with the lyrics to simplify the details). For example, “One little Indian” consists of a specific length of particular notes in a particular order. Even if the right notes were not in their original order it would not be the same melody. However, the order of these notes or sounds is not given by any abstract rule based category membership, but only learned associations. From the sequence of the constituent tunes or sounds there is no way to extract any rule that could be used in a different melody or any situation. However, in this study, independent of how many times a person has heard a given children’s song, like “One little Indian” this string of notes is idiosyncratic and specific to a certain song and she or he cannot extract any rules out of this melody string that could be applied to new situations.

Furthermore, melodies were cut into pieces (melody phrases) and presented in pairs. Associations of the melody phrases are based on learned contingencies; therefore they are strings of contingent associations. When the first couple of notes are presented, subjects who are familiar with the melody tokens expect the rest of the notes to be heard. According to Ullman’s theory, word stems and the irregular verbs or all elements of a second language are stored in and rely on the associative declarative memory and they would not elicit similar brain responses elicited by strings with abstract rules (Ullman, 2004).

If violation of associative tokens can obtain similar brain responses to those that rule based strings of music or language elicit that would support the proposition rejecting the traditional distinction between declarative memory and rule system.

The main hypothesis of studies 3,4 and 5 is that processing uncoupled constituent elements occurs in one single distributed network and coherency of elements can emerge through encoding statistical associations that exist in the world. According to that, similar brain responses will be elicited by violating the expectancy of associative tokens to those were reported in violating abstract rules such as what word category can be presented after a category of “article” (Friederici, et al 1993. Furthermore, there is no relevance to predict functional differences based on controlled (declarative) or implicit (non- attend) mechanisms (Nosofsky & Zaki, 1998) therefore the same variables effect the ERPS in ignore (3) and attend (4) studies.

13. Study 3 – Non-attend

Hypothesis

Violation of contingent relationship of associative melody token pairs will elicit similar brain responses to abstract rule violation of any elements since unifying cohesive elements can based on statistical association of co-occurring elements.

Method

Familiar American children songs were cut into two parts, following their musical phrases as such: Melody part1 (M1) "One little two little three little Indians", and Melody 2 (M2) "five little six little seven little Indians". Only a learned contingent relationship linked these melody pairs together. The song phrases were presented to participants in pairs, separated by a time delay (short, ISI=800 ms, long ISI=2800 ms). Half of the pairs were contingent (the second melody part completed the first) and the other half were violated; the two phrases did not match since the second part was chosen from a different melody. Therefore, by the violation, the associational string was broken.

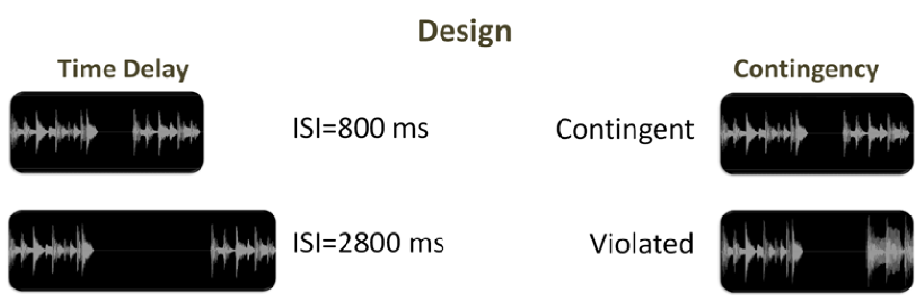


Figure 9. The experimental design consisted of: Time delay (short: ISI=800 ms and long: ISI=2800 ms) and Contingency (Contingent vs. Violated) of the melody pairs. The melody pairs were either contingent short or contingent with a long time delay or Violated with short or long time delay in between the melody parts.

There were four conditions: short delay with contingent phrases, short with non contingent/violated, long delay with contingent phrases, and long with non contingent/violated. The time delay variable was used to engender responses underpinning processes which were reported in the previous long distance dependency studies. There updating and unification processes were found inseparable

(Jolsvai et al, submitted). However, since the hypothesis is that unification of different representations can be executed by the learned co-occurrence of items without being members of an abstract category group the other variable is Contingency. Half of the melody pairs were violated the expectancy of the contingent second phrase (more details in the material section).

In Study 1 (Non active or Ignore), subjects did not pay direct attention to the auditory stimuli but watched a captioned movie. In Study 2 (Active or Attend), subjects paid direct attention to the stimuli and were asked to judge the contingency of the pairs. In both Ignore and Attend experiments, ERPs to the second phrase of the pairs were compared across the four conditions.

Participants

Thirteen American[11] adults (5 males) from 20-36 years old (M=28, 94 SD=4.35) were paid for their participation. The subjects were treated in accordance with the ethical research guidelines for human subjects at Albert Einstein College of Medicine, where the study was conducted. Participants gave informed consent. None of the subjects had difficulty hearing, or a history of either neurological or psychiatric disorders.

Materials

In both ignore and attend experiments the same material was used, therefore, I will only present it at the ignore study and not discuss the materials again at the attend study.

American Children songs were collected and then cut into various lengths (between 1800-2800 ms) of musical phrases retaining and following their melody phrases. Ten independent subjects were asked to rate the melody parts on a 1- 10 scale preceding the ERP experiments in reference to contingency. 10 indicates that the second element of the melody pairs completed the first and when the first element of a particular pair was presented subjects were expecting the second part to follow. Only the best pairs (only 10s from all the 10 subjects) were kept in the study. The length of

[11] Participants were all born and raised in the USA and none of them were first generation Americans.

the melody parts varied from about 1800 ms to 2800 as we were following their musical phrases (which included 5-msec of onset and offset linear ramps), equated for intensity at 80 dB SPL (calibrated with a Bruel & Kjaer 2209 sound level meter), which was the level at which they were presented. They were presented bilaterally through insert earphones (Eartone 3A) at a varying inter stimulus interval ISI short: 800 ms, ISI long: 2800 ms between the music phrases. There were 0, 3 or 6 piece(s) of 50 ms segments cut from the same melody and presented in between the M1 and M2 phrases. ITI was 4000 ms. The segments had the same spatio-temporal characteristics as the melody phrases they divided/separated, but they were too short to be recognizable. Segments were used to minimize refractoriness effect and also to make the design less predictable for the participants.

Each pair was presented only once, but there were different pairs presented from the same melody. All the different melody parts were normalized by Adobe Audition software.

Procedure

Subjects were tested individually in a dimly illuminated sound-attenuating booth. They were seated in a comfortable reclining chair, and instructed to move as little as possible. Participants were told that they would be presented with a series of melody pairs. They were watching a movie with captioning and were asked not to pay attention to the music. The stimuli were presented in four conditions (short contingent, short violated, long contingent, long violated).

Music phrase pairs were presented binaurally at 80 dB. The list of music pairs in the different conditions were presented pseudo-randomly and in varying order for each subject. The experimental trials were presented in 40 blocks of approximately 2 to, 4 min each. Subjects were given short breaks between the blocks if it was necessary and a long break in the middle.

Electrophysiological Recordings

EEG was recorded continuously with an electrode cap 32 scalp locations (modified 10-20 International system) Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp5, Cp6, Cp1, Cp2, P7, P8, P3, P4, O1, O2, plus electrodes placed on the right (RM) and left (LM) mastoids. Horizontal electro-oculogram (EOG) was recorded using F7 and F8 electrode sites, and vertical EOG was recorded using a bipolar montage between FP2 and an external electrode placed beneath the right eye. The reference electrode was attached to the tip of the nose. The EEG and EOG were digitized (Neuroscan Synamps amplifier, Compumedics. Corp., El Paso, Texas, USA) at a sampling rate of 500 Hz (bandpass 0.05-100 Hz),

Data analysis

Offline analysis of the data was done with Neuroscan 4.3 Software comprising the following steps. After applying a low-pass filter (15 Hz, 24 dB/oct), the EEG data were segmented into epochs of 1650 ms (-150 ms before onset and 1500 ms after onset); next the data were baseline-corrected on the entire epoch. Artifact correction criteria was set to ± 75 µV. The segmented epochs were then averaged by each condition and base-line corrected to the 150 ms pre-stimulus. The ERPs time-locked to the second melody parts were analyzed and compared. P1, N1 and P2 peaks were selected by Global Field Power. To compare amplitudes and latencies, repeated measures analysis of variance (ANOVA) with Statistica 9 was conducted with factors of Time Delay (Short vs. Long), Contingency (Contingent vs. Violated), and electrode (10 electrodes total: Fz, Cz, F3, F4, Fc1, Fc2, C3, C4, Cp1, Cp2 for the amplitude analysis and 3 electrodes (Fz: Cz, Pz) for latency analysis.

To analyze the slow potential Sustained Negativity (SN), it was cut into 14 x 50 ms parts between 800-1500 ms. Repeated measures of ANOVA was performed with the levels of Time delay (Short-Long), Contingency (Contingent-Violated), Electrode (Fz, Cz, F3, F4, Fc1, Fc2), Time window (14) was performed. A Greenhouse-Geisser adjusted univariate test was used to avoid possible violation of compound symmetry and sphericity assumption. Tukey HSD post-hoc analyses were performed to analyze further the significant effects.

ERP Results

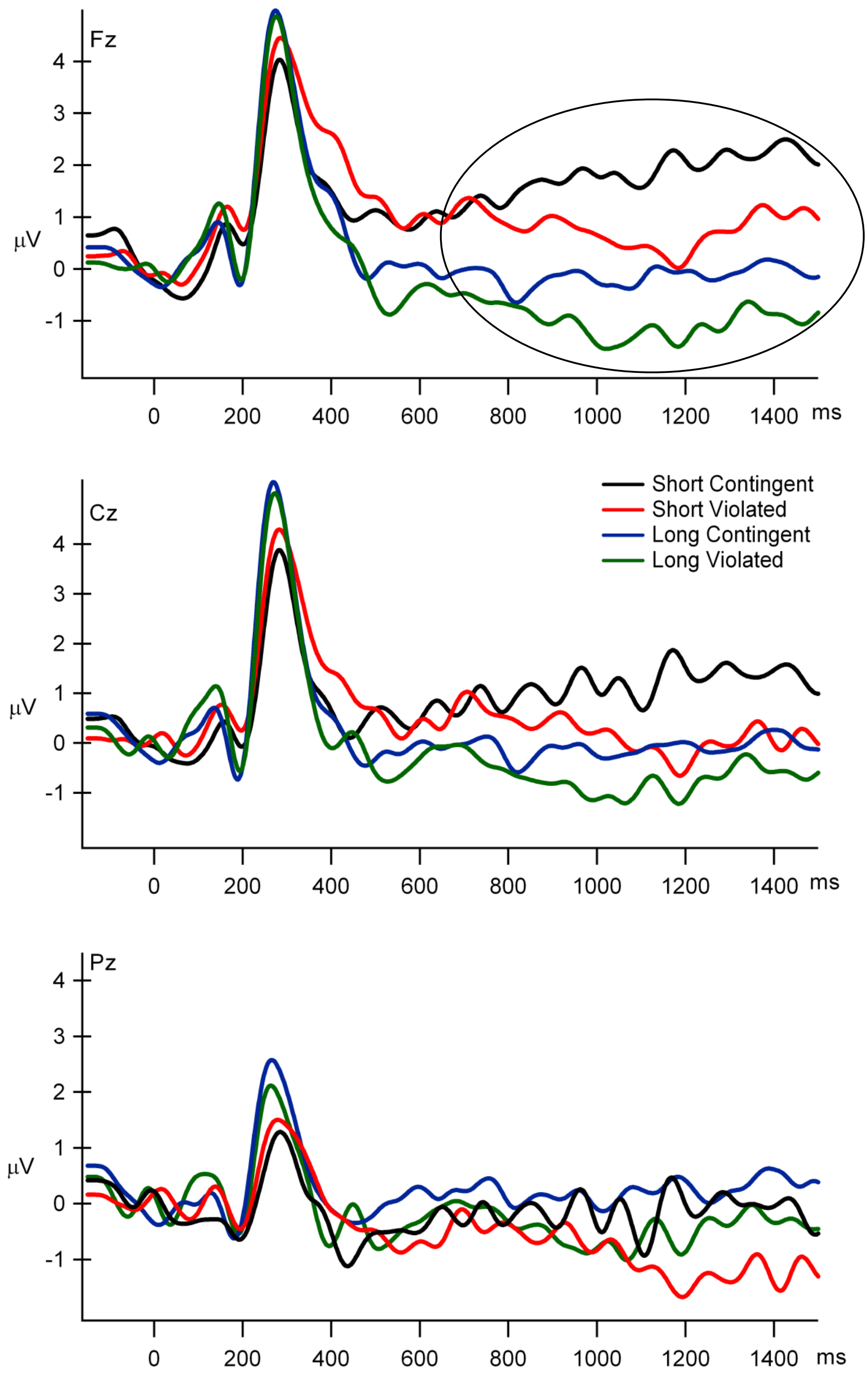


Figure 10. Frontal Sustained Negativity (SN) elicited by the Violation of Contingency and Long Time delay. Figure presents Fz, Cz and Pz electrodes. Negativities are downwards.

Although the obligatory components of the last three studies of the thesis are reported, due to the tradition of the language ERP field as well as since the focus of these studies is on the later effects similarly to Studies 1 and 2, the obligatory components (P1, N1 and P2) will not be discussed.

P1

P1 amplitude was smaller at Fc1 and Fc2 than the rest of the measured electrodes, Electrode main effect [$F(2.11, 25.43)=14.28$, $p<0.001$]. Long pairs elicited a larger P1 at Fc1 and Fc2 than the Short pairs, Contingency x Electrode interaction [F(2.60,31.24, 108)=3.295, $p<0.001$].

N1

Long time delay elicited a larger N1 amplitude than short time delay, main effect of Time delay,[$F(1, 12)=5.873$, $p=0.032$]. Short time delay elicited a longer N1 latency, main effect of Time delay,[$F(1, 12)=57.75$, $p<0.001$]. At Pz short time delay elicited a shorter N1 latency than at Fz or Cz, Contingency x Electrode interaction [$F(1.97,23.65)=4.447$, $p=0.02$].

P2

P2 amplitude was larger at Fz, F3, F4, Fc1, Fc2 and Cz than at C3, C4, Cp1, Cp2 (main effect of Electrode, [$F(2.66, 31.94)=27.957$, $p<0.001$]. Short time delay elicited a longer P2 latency at Fz and Cz than long time delay, as revealed by Time delay x Electrode interaction [F(1.33, 16.09)=5, 332, $p=0.012$].

Sustained Negativity

It is visible on the Grand mean ERPs that between 400- 600 ms a transient-like fronto-central negativity started to be engendered, however the GFP did not indicated a real peak in this time window.

Long time delay elicited a more negative SN than short time delay; main effect of Time delay, [$F(1, 12)=7.16$, $p=0.02$]. Violated pairs elicited a more negative SN than contingent pairs, revealed by the main effect of contingency [$F(1, 12)=8.55$, $p=0.012$].

Summary of Non-Attend Study (3)

Without direct attention to the melodies, violation of the expectancy of associative melody tokens elicited a Sustained Negativity (SN). Long time delay obtained a more negative SN than short delay pairs. It is important to emphasize that it was visible on the Grand mean ERPs that between 400- 600 ms a transient-like fronto-central negativity had started to be engendered by time delay. However it was not recognized as a real peak by GFP. At 800 ms, the SN was uncoupled by time delay and the short long delayed waves were uncoupled again by contingency; where the non-contingent, violated pairs obtained a more negative Sustained negativity in comparison to the contingent pairs.

14. Study 4 – Attend

Hypothesis

Similar but transient effects will be elicited by Contingency and Time delay in an active (attend) situation than was obtained in the non- attend study, since there is no relevance of qualitatively different unconscious (rule) and controlled (declarative) mechanisms in a single connectionist network.

Methods

Participants

Twelve adults (3 males) from 23 to 37 years of age (M: 26.5, SD=3.51) were paid for their participation. The subjects were treated in accordance with the ethical research guidelines for human subjects at Albert Einstein College of Medicine, where the study was conducted. None of the subjects had difficulty hearing, or a history of either neurological or psychiatric disorders.

Materials

The same material was used as in Study 3.

Procedure

Subjects were tested individually in a dimly illuminated sound-attenuating booth. They were seated in a comfortable reclining chair, and instructed to move as little as possible. Participants were told that they would be presented with a series of melody pairs. They were asked to pay attention to them and respond after each pair, indicating whether they were contingent (right button on the responding box) or violated (left button on the responding box). The stimuli were presented in four conditions (short contingent, short violated, long contingent, long violated).

Music phrase pairs were presented binaurally in 80 dB. The list of music pairs in each of the different conditions was presented pseudo-randomly and in varying order to each subject. The experimental trials were presented in 40 blocks of approximately 2-, 4 min each. Subjects were given short breaks between the blocks if it was necessary and a long break in the middle.

Electrophysiological Recordings

EEG was recorded continuously with an electrode cap 32 scalp locations (modified 10-20 International system) Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp5, Cp6, Cp1, Cp2, P7, P8, P3, P4, O1, O2, plus electrodes placed on the right (RM) and left (LM) mastoids. Horizontal electro-oculogram (EOG) was recorded using F7 and F8 electrode sites, and vertical EOG was recorded using a bipolar montage between FP2 and an external electrode placed beneath the right eye. The reference electrode was attached to the tip of the nose. The EEG and EOG were digitized (Neuroscan Synamps amplifier, Compumedics. Corp., El Paso, Texas, USA) at a sampling rate of 500 Hz (bandpass 0.05-100 Hz),

Data analysis

Offline analysis of the data was done with Neuroscan 4.3 Software in the following steps. After applying a low-pass filter (15 Hz, 24 dB/oct), the EEG data were segmented into epochs of 1650 ms (-150 ms before onset and 1500 ms after onset); next the data were baseline-corrected on the entire epoch. The artifact correction criteria was set to ± 75 µV. The segmented epochs were then averaged by each condition and base-line corrected to the 150 ms pre-stimulus. The ERPs time-locked to the second melody parts were analyzed and compared. To compare P1, N1, P2 and Anterior Negativity amplitudes, three way repeated measures analysis of variance (ANOVA) with Statistica 9 software was conducted with factors of Time Delay 2 levels (Short vs. Long), Contingency 2 levels (Contingent vs. Non-contingent), and electrode 6 levels (F3, Fz, F4, Fc1, Fc2, Cz, 6 electrodes total) for P1, N1 and P2 and 6 electrodes of total for the transient Negativity between 300-500 ms (Cz, C3, C4, Cp1, Cp2) and 7 electrodes for the late positivity (Pz, P7, P8, P3, P4, O1, O2) were selected. To compare P1, N1, P2 and Anterior Negativity latencies, a repeated measures analysis of variance (ANOVA) was conducted with factors of Time Delay 2 levels (Short vs. Long), Contingency 2 levels (Contingent vs. Violated), and electrode 3 levels (Fz, Cz, Pz).

Sustained Negativity (SN) measured between 500 and 1500 ms post stimulus followed by the transient anterior negativity (AN), was also cut into 50 ms segments and analyzed with factors of Time Delay 2 levels (Short vs. Long), 2 levels Contingency (Contingent vs. Non-contingent), and time windows(20), 6 levels electrode (F3, Fz, F4, Fc1, Fc2, Cz, 6 electrodes total). A Greenhouse-Geisser adjusted univariate test was used to avoid possible violation of compound symmetry and sphericity assumptions. Tukey HSD post-hoc analyses were performed to analyze further the significant effects.

Behavioral results

Reaction times were analyzed by a repeated measures of ANOVA (Statistica 9 software) with levels of Time delay (short-long) and Contingency (Contingent-Violated).

Reaction Time

	Short Contingent	Short Violated	Long Contingent	Long Violated
RT	442.68	474.57	445.81	506.25
SD	82.12	108.85	101.68	122.40

Table 3. The mean RTs and SD are presented in each condition.

Violated (Non-contingent) pairs elicited longer RTs, main effect of Contingency,[$F(1, 12)=5.155$, $p=0.042$]. Long violated elicited longer RTs than short Violated pairs revealed by a Contingency x Time delay interaction,[$F(1, 12)=4.829$, $p=0.048$].

ERP Results

P1

P1 amplitude had a Cz maxima, main effect of Electrode [$F(2.63, 29)=12.262$, $p<0.001$]. Long time delay elicited a larger P1 at Cz, than short time delay, revealed by Time Delay x Electrode interaction, [$F(2.92,32.12)=5.881$, $p<0.001$]. Long time delay elicited a shorter P1 latency, main effect of was [$F(1, 11)=58.686$, $p=0.001$].

N1

N1 amplitude was less negative at F3 than at Fz, Cz, F4, Fc1 and Fc2 (Electrode x Time delay interaction, [$F(2.10, 23)=6.63$, $p< 0.001$].Short time delay elicited a less negative

N1 F3 and F4 and Fz than the Violated condition, Time delay x Electrode interaction [$F(2.27, 25.03)=12.034$, $p<0.001$. Violated pairs elicited a longer N1 than the Contingent condition, Contingency main effect [($F(1, 11)=6, 508$, $p= 0.026$]. Short time delay elicited a longer N1 latency than long time delay, main effect of Time delay [($F(1, 11)=83, 864$, $p<0.001$]. N1 latency was the longer at Fz Electrode main effect [($F(1.29, 14.22)=6, 788$, $p<0.005$)].

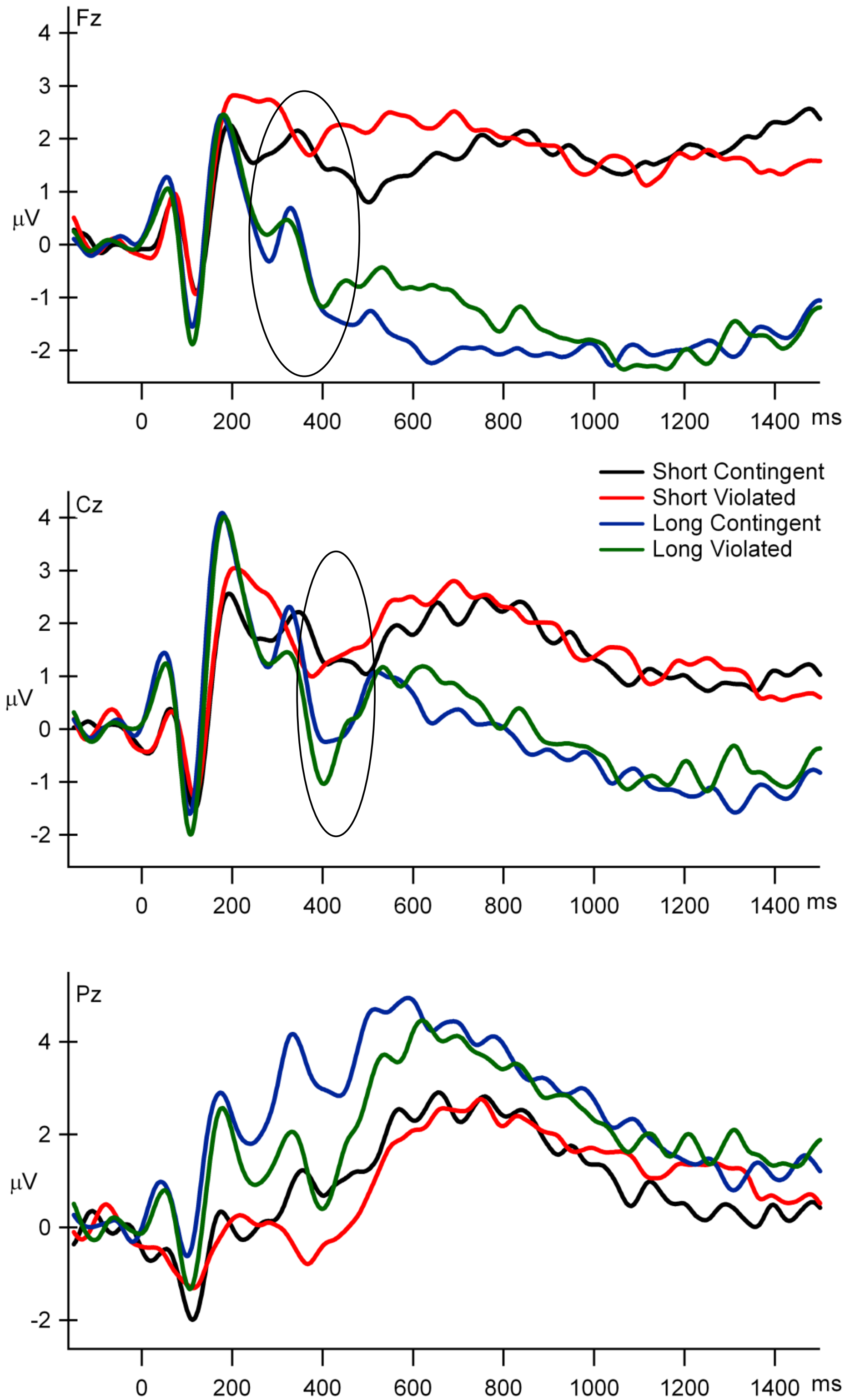


Figure 11 displays the grand average ERPs in all conditions at Fz Cz and Pz. Negativity is downwards.

P2

P2 was the largest at Cz, main effect of Electrode [F(1.78,19.58)=7.88,p<0.001. Long delayed pairs elicited a larger P2 amplitude at Cz, Time delay X Electrode interaction [F(2.75,30.35)=22.02, p<.001]. P2 latency was longer in the short time delay condition, main effect of Time delay, F(1, 11)=90.049, p<0.001.

Peaks	Short Contingent	Short Violated	Long Contingent	Long Violated
Negativity (AN)	408	374	414	404
P300	598	610	598	610

Table 4. The latencies of transient negativity referred as AN and the late positivity referred as P600 in each condition.

Negativity between 300-500 ms (AN)

Violated pairs elicited a larger AN Negativity (300-500 ms) all analyzed sites than contingent pairs, Contingency x Electrode interaction,[F(2.24, 24.67)=27.613, p<0.001]. Long delayed pairs elicited a larger AN at Cz, C3 and C4, Time delay x Electrode interaction, F(2.35, 25.88)=17.05, p<0.001.

Long Time delay correlated with a longer AN latency revealed by a Time delay main effect [F(1, 11)=26.956, p<0.001]. Non-Contingent (Violated) pairs correlated with a shorter AN latency, main effect of Time delay, [F(1, 11)=68.435, p<0.001]. Short Violated pairs correlated with a significantly shorter latency than the latencies in all the rest of the conditions revealed by a Time delay and Contingency interaction [F(1, 11)=25.274, p<0.001].

Sustained Negativity (500-1500 ms)

Long delayed pairs elicited a more negative SN, revealed by a Time delay main effect, [F(1, 11)=32.535, p<0.001].

P300

Long time delay elicited a larger P300 at parietal sites, main effect of Time delay,[F(1, 11)=23.249, p<0.001] and Time delay x Electrode interaction [F(2.29, 25.27)=10.595, p<0.001. Neither Time delay (F(1, 11)=.12090, p<0.734 nor Contingency F(1, 11)=2.2293, p=0.163 reveal any significant effect of P300 latency.

Summary of Attend Study (4) Results

A transient Negativity between 300-400 ms was enlarged by to both Contingency and Time delay. Long time delay correlated with a more negative Sustained negativity between 500-1500 ms than the short time delay. Parietal P300 was enlarged by the time delay.

Summary of Studies 3 and 4 Results

In Studies 3 and 4, melody tokens without any abstract category membership presented in two parts separated by time obtained very similar effects to those measured in neurophysiological language or music studies that violated an abstract rule of the string. In both ignore and attend experiments contingency and memory load correlated with similar effects.

In the non-attended study, a slow negative wave correlated with the time delay and contingency double uncoupling the slow wave from 800 ms. It is visible on the Grand mean ERPs that between 400- 600 ms a transient-like fronto-central negativity had started to engendered however it was not recognized as a real peak by the GFP calculation. It might be due to the fact that in acoustic studies the attention is fluctuates since participants can not completely ignore the stimuli but do watch a movie (Sussman, personal communication). In the attend study, memory and contingency affected a transient negativity. Both studies engendered similar effects; contingency and time delay both affected the SN in the non-attend and the transient negativity in the attend studies.

15. Study 5 – Control

Studies 3 and 4 showed that a series of two elements with simple learned contingent pairs of melody tokens with associative relationship elicit ERP effects that are similar to abstract rule violations of a string of linguistic stimuli.

To control that these effects reflect a similar process support the unifying of two constituent melody parts as were registered in Studies 1 and 2, a control study was conducted. Specifically, the main goal of the control study is to provide evidence that the unification effects registered in studies 3 and 4 do not reflect retrieval processes in long term memory, but rather indicate an indulgent unification process that can be based on statistical learning or contingency.

Neural processes indexed by the N400 component indicate the configuration of knowledge and retrieval processes in long term memory (Kutas et al, 2006). As Kutas and Federmeier, (2000) discussed extensively, the centro-parietal N400 indexes some sort of configuration or organization of memory. According to Kutas and Federmeier (2000), N400 reflects the semantic and probabilistic effects that indicate an organization of the arrangement of items in the knowledge base (Kutas et al, 2005). The more the test item matches to its previous context, the smaller the N400 amplitude is (Kutas & Federmeier, 2000).

Kutas and Hillyard's (1980) seminal study elicited a positive waveform between 200-600 ms for sentence final words that formed predictable completion, while the incongruent words elicited a large negative wave in this time range. The N400 effect is the difference wave of these two conditions (Kutas et al, 2006).

It is known that semantic and identical priming inversely correlate with N400; eliciting a smaller N400 amplitude (a positivity rather, Kutas & Hillyard, (1980)) in the case of semantic congruency or identical priming (Kutas et al, 2006).

The main goal of the control study is to provide evidence that the unification and contributing memory (holding online) processes measured in the previous experiments are separable form memory retrieval processes.

Further aims of the control study are to make sure that all the later effects are not the consequence of any refractoriness effect. Refractoriness affects the early obligatory components (Ritter et al, 1968) but will not affect the later processes.

Hypothesis

N400 effect reflects memory retrieval processes. Two identical musical phrases presented in pairs will elicit a positive wave between 200-600 ms. Contrary to the Music condition, there will be no effect of the Pure Tone and White Noise condition. Refractoriness will not affect the brain waves elicited following the obligatory components (N1 and P2) in all three conditions (music, noise, pure tone).

Methods

Participants

Ten adults (5 male) from 26 to 31 years of age (M:28. 3 years SD=3.98)) were paid for their participation. The subjects were treated in accordance with the ethical research guidelines for human subjects at Albert Einstein College of Medicine, where the study was conducted. None of the subjects had difficulty hearing, or a history of either neurological or psychiatric disorders.

Materials

One single music phrase element (with a length of 2000 ms) of the previously used musical phrases pair was presented with a short (800 ms) and a long (2800 ms) time delay in between the same melody presented twice with 4000 ms ITI.

Since the same melody part was presented twice (M1) separated by time delay (M1 again), in this control study the pairs lacked any learned contingency but were similar to an identical priming study. Similarly to the music condition, there were a noise pair condition and a pure tone condition, with the same lengths (2000 ms) and time delay (short=800 ms and long=2800 ms) between them with 4000 ms ITI.

The control study had three conditions: Melody pairs, White noise pairs, and Pure tone pairs. The three conditions were presented in three blocks (the order of the blocks rotated for all the subjects). One third of the stimuli were melody pairs (80 pairs), another third were noise pairs (80 pairs) and the last third were 1000 Hz pure tone pairs (80 pairs).

Procedure

Subjects were tested individually in a dimly illuminated sound-attenuating booth. They were seated in a comfortable reclining chair, and instructed to move as little as possible. Participants were told that they would be presented with a series of melody pairs or pure tone or white noise pairs. They were asked to watch a movie with captioning and not to pay attention to acoustic stimuli. The stimuli were presented in three blocks, music, noise, pure tone. Music, Noise and Pure tone pairs were presented binaurally in 80 dB. The experimental trials were presented in 3 blocks of approximately 30 min each. Subjects were given short breaks between the blocks or if it was necessary and a long break in the middle.

Electrophysiological Recordings

EEG was recorded continuously with an electrode cap 32 scalp locations (modified 10-20 International system) Fpz, Fz, Cz, Pz, Oz, Fp1, Fp2, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp5, Cp6, Cp1, Cp2, P7, P8, P3, P4, O1, O2, plus electrodes placed on the right (RM) and left (LM) mastoids. Horizontal electro-oculogram (EOG) was recorded using F7 and F8 electrode sites, and vertical EOG was recorded using a bipolar montage between FP2 and an external electrode placed beneath the right eye. The reference electrode was attached to the tip of the nose. The EEG and EOG were digitized (Neuroscan Synamps amplifier, Compumedics. Corp., El Paso, Texas, USA) at a sampling rate of 500 Hz (bandpass 0.05-100 Hz),

Data analysis

Offline analysis of the data was done with Neuroscan 4.3 Software comprising the following steps. After applying a low-pass filter (15 Hz, 24 dB/oct), the EEG data were segmented into epochs of 1650 ms (-150 ms before onset and 1500 ms after onset); next the data were baseline-corrected on the entire epoch. Artifact correction criteria was set to ± 75 µV. The segmented epochs were then averaged by each condition and base-line corrected to the 150 ms pre-stimulus. The ERPs time-locked to the second melody/noise/pure tone parts were analyzed and compared. To compare N1, P2, and P3 a three way of ANOVA with Statistica 9 software was conducted with levels of Condition (Music, Noise, Pure tone) Time Delay (short/long), Electrode (Fz, Cz, Pz). To compare the critical 500 and 1500 ms post stimulus time window, brain waves were cut into 50 ms segments and analyzed by a four way of ANOVA, with factors of Memory Delay 2 levels (Short vs. Long), 3 levels of Condition (Melody, Noise, Pure) 20 levels of Time (50ms segments x 20 in between 500- 1500 ms) and levels of electrodes (Fz, Cz, Pz)

A Greenhouse-Geisser adjusted univariate test was used to avoid possible violation of compound symmetry and sphericity assumption. Tukey HSD post-hoc analyses were performed to analyze further the significant effects.

	Music		**Noise**		**Pure**	
	Short	**Long**	**Short**	**Long**	**Short**	**Long**
N1	108 (93-123)	112 (97-127)	102 (87-117)	100 (85-115)	108 (93-123)	98 (83-113)
P2	186 (171-201)	202 (187-217)	186 (171-201)	196 (181-211)	164 (149-179)	184 (169-199)
P3	360 (345-375)	352 (337-367)	360 (345-375)	352 (337-367)	360 (345-375)	352 (337-367)

Table 5. Peak latencies of N1, P2 and Music P3. Comparing the amplitude differences elicited by Music condition, the same time windows were used in Noise and Pure Tone conditions. In brackets the latency ranges used for amplitude extraction are shown.

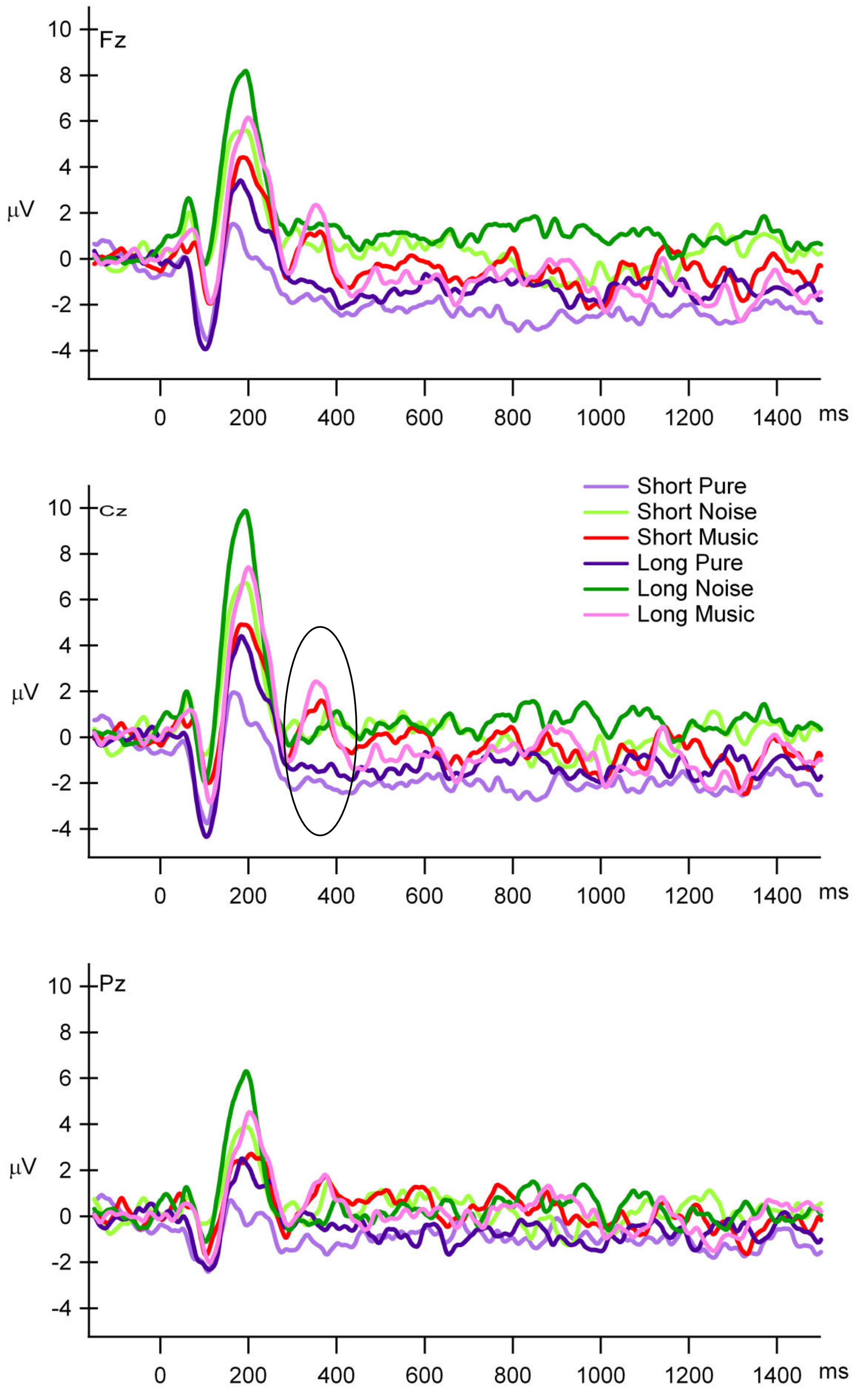


Figure 12. Grand mean averages in Short Music (SM) Short Noise (SN), Short Pure (SP), Long Music (LM), Long Noise (LM) and Long Pure tone (LP) conditions. Negativity is downwards.

Results

N1

N1 amplitude elicited by Noise was significantly less negative than N1 obtained for Music or Pure tones, Condition main effect [$F(1.37, 12.36)=11.049$, $p=0.001$]. N1 was the most negative at Cz as revealed by an Electrode main effect, $F(1.38, 12.04)=5.436$, $p=0.015$]. N1 elicited by Pure tone was larger at Fz and Cz than N1 elicited by Music or Noise, Conditions x Electrode interaction,[$F(2.76, 24.88)=7.749$, $p<0.001$]. N1 latency was the longest in the Music condition; Condition main effect [$F(2, 18)=13, 739$, $p<0.001$].

P2

Noise Condition elicited the largest P2, main effect, [$F(2, 18)=12.914$, $p<0.001$].

P2 was the most positive at Fz and Cz in Music and Noise conditions, Condition x Electrode interaction, [$F(4, 36)=3.681$, $p<0.012$].

Long time delay correlated with a longer P2 latency, main effect main effect of Time delay, $F(1, 9)=96.281$, $p<0.001$. Pure tone elicited a shorter P2 latency than Music and Noise conditions, main effect of Condition, $F(2, 18)=28.075$, $p<0.001$.

Music positivity around 300 ms

Music elicited a significant positivity at Fz and Cz revealed by Condition x Electrode interaction, $F(2.76, 24.85)=7.21$, $p<0.001$. Longer Time delay correlated with a longer Music positivity latency $F(1, 9)=6, 061$, $p=0.036$.

SN (500- 1500 ms)

There was no significant effect of Time delay, $F(1, 9)=.832$, $p=0.385$.

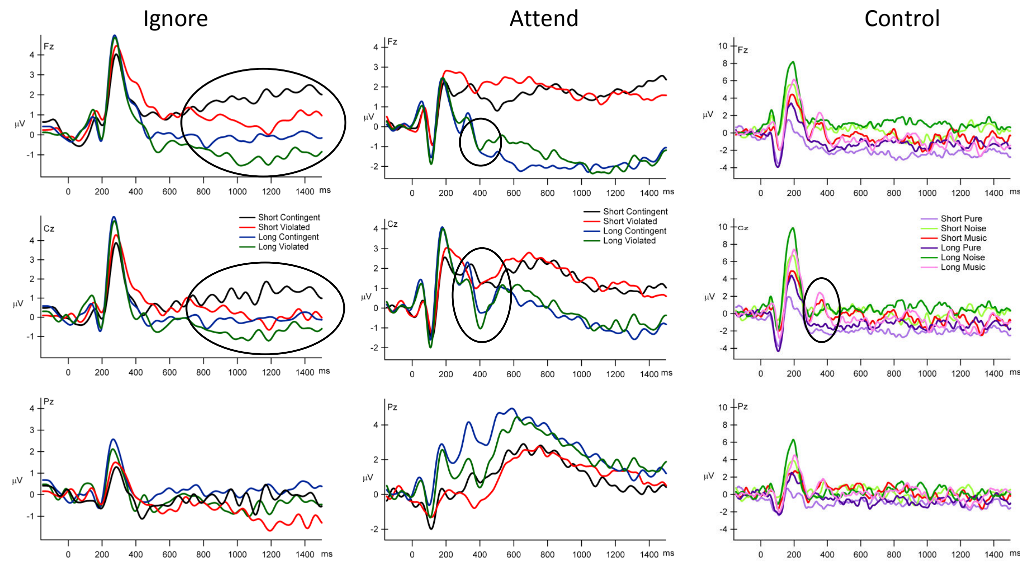


Figure 13. Summary of the ERPs in all the three: ignore, attend and control studies.

Summary of Study 5 Results

There were no sign of refractoriness effects (Ritter et al, 1968) after the obligatory components even in the Noise and Pure tone conditions. The music identical priming condition elicited a central positivity that is similar to what Kutas and Hillyard (1980) registered for the semantically congruent sentences. Since in the studies 3 and 4 the contingent pairs did not correlate with a similar positivity, it is convincing that the negativities registered in studies 3 and 4 are not responsible for retrieval processes in the knowledge base but rather index an indulgent unification process that can be based on contingent relationships.

16. Discussion of Studies 3, 4, and 5

The violation of a stream of musical phrases with associative relationships engendered very similar neurophysiological responses as those identified with abstract rules elicited by a stream of linguistic elements in previous studies (e.g. Friederici, 2002).

Without focused attention the violation of contingent musical phrases elicited Sustained negativity (SN) which was affected by the contingency violation and the time delay. It is visible on the grand mean ERPs that between 400- 600 ms transient-like fronto-central negativity had started to engender however this effect was not recognized as a real peak by the GFP. It is possible that due the fluctuation of attention, instead of an evolved transient negativity observed in the attend study, a slow negativity was obtained. In the attend study, similarly affected but transient waves were elicited: a negativity was engendered by both contingency and time delay, followed by a SN that was affected by time delay. A positive peak between 600-800 ms followed the negativity that was enlarged by time delay. SN following the transient negativity in the attend study might indicate that once the participants made their contingency judgment (around 400 ms according to the RT) they did not focused their attention to the melodies. However, the positivity measured around 600 ms was affected by the time delay and not contingency presumably because making contingency judgment on the pairs were more difficult if the pairs were separated by long time delay. That was supported by the RTs; long violated pairs elicited longer RTs than short violated pairs. These results support the theories of P300 indexes the neurophysiological activity required the ongoing model of the environment is revised or updated in working memory (Donchin, 1981; Donchin & Coles, 1988). In the control study there was no Sustained or transient Negativity obtained, but an increased positivity -the congruent positivity of the N400 effect- was elicited by the time delay in the music condition. The results of the control study provide evidence that the measured SN in the non active (ignore) contingency study (study 3) and transient negativity followed by the SN (study 4) reflect unification processes that can be based on learned contingency without any membership of abstract categories. This process is at least partially different from the memory retrieval process that is indexed by the positivity around 300-500 ms, elicited solely by the identical Music condition but neither by Noise or Pure Tone conditions nor by the Contingency pairs in studies 3 and 4.

Studies 3, 4 and 5 provided evidence of a compliant unification process that is responsible for merging consistent elements, where a cohesive relationship can be based on contingency (Jolsvai & Sussman, in preparation).

Models depict functional differences between mechanisms subserving rule and declarative systems argue that a rule system is an automatic operation while mechanisms associated with declarative a memory system are controlled (e.g. Friederici, 2002; Ullman, 2004; Waters & Caplan, 1999). On the contrary, the non active study (study 3) revealed similar neurophysiological responses to the active situation (study 4): in both studies ERPs were affected by memory and contingency. These results can challenge the sharp functional demarcation of the two rule and declarative memory systems. However, one should bear in mind that in acoustic non-attend studies unlike in visual experiments, the participants' attention to the stimuli fluctuate. This was obvious since after the ignore study participants were able to name number of melodies. Without focused attention, the participants still hear the melodies and unify them similarly to the attend study, however since they were not making decisions about the pairs, the negativity was a slow wave and there was no P300.

As studies 3, 4 and 5 made clear, violations of the expectations of associative melody parts can elicit similar brain responses to abstract rule violations of linguistic or music elements. These results indicate that temporal lobe memory and frontal unification system is one highly interconnected distributed network, similar to what connectionist network models propose (MacDonald & Christiansen, 2002). The identification of similar binding processes without any abstract rule operation, other than elicited by syntactic violations (e.g. Münte, 1993) suggests that declarative memory and procedural system are more permeable than traditionally suggested.

DP theory (Ullman, 2001) mostly relies on neuropsychological and functional imaging data. The problematic aspects of fMRI (Uttal, 2001; Dobbs, 2005) have been discussed above. Nevertheless, neuropsychological data of amnesic patients that are in line with the predictions of DP theory have also been challenged. Nosofsky and Zaki (1998) demonstrated that differences in performance on procedural and declarative tasks can be obtained by a single memory system.

Cowan (1999) also argues that the amnesic patients data of different systems can be interpreted as the damage of a hippocampal-cortex binding process. These results are in line with the neural network assumption (e.g. MacDonald and Christiansen, 2002) that long term memory is not separated from working memory or the processing itself, but that they are all implemented in the one single network.

17. General Discussion and Outstanding Issues

The main hypothesis of thesis was that processing uncoupled constituent elements occurs in one single distributed network. In this single distributed set of connections there are no separate mechanisms like WM and syntax or WM and knowledge but all of these features are implemented in one network. Furthermore coherency of elements can emerge through statistical learning (Rummelhart, 1989).

Processing of uncoupled subject-verb dependency elicited frontal-frontocentral interdependent fronto-central negativities indicating functionally inseparable cooperative processes of language comprehension (Jolsvai et al, submitted). The results of study 2 with a single memory process together with the CSD maps and GD indexes of Study 1 point out that unification and memory processes are not likely qualitatively dissimilar, they rather represent quantitatively different activation patterns of the same network (MacDonald & Christiansen, 2002). Furthermore, string of associative music tokens obtained similar neurophysiological responses to syntactic rule violations (e.g., Münte, 1993) or other abstract rule violation (Hoen & Domeney, 2000) advocating that the traditional separation of rule and memory system is dematerialized (Hagoort, 2009). The non active and active situations revealed similar unification effects denying the distinction of controlled declarative and automatic rule processing. Furthermore, these results suggest that coherency of elements can emerge through statistical learning (Rummelhart, 1989). Interestingly, similar statistical learning mechanism for speech and tone sequences segmentation has been suggested, in this manner raising the attractive opportunity of a common language and music learning device (Schön et al, 2008).This possibility is supported, from an evolutionary perspective (Besson & Schön, 2003). Moreover, this common learning device apparently can be still operated in adulthood (Jolsvai & Sussman, in preparation).

These results are in line with the neural network assumption (e.g. MacDonald & Christiansen, 2002) where long term memory, knowledge and processing are implemented in one network.

One should consider the neurobiological characteristics of the human brain when attempting to evaluate complex cognitive behavior by biological responses (Uttal, 2001; Dobbs, 2005). It is worth bearing in mind that neurophysiological measurements/components elicited by complex cognitive tasks may not represent one singular cognitive process because of the parallel and interconnected nature of the brain (Kutas et al, 2006; Uttal, 2001). The results of studies 3, 4 and 5 challenge the assumption of different functional mechanisms by emulating a syntactic dependency with a contingent dependency of associative music strings. This design mimicked second language processing in Ullman's view, where second language operation exclusively relies on the associative memory system (Ullman & Pierpont, 2005). In the absence of any abstract rule, the violations of the music tokens exclusively violated their learned contingent relationship. However, negativities elicited were similar to those obtained in studies 1 and 2, where morpho-syntactic features and the distance between them were manipulated. The similar results of active and non-active experiments can call into question the theories' separate declarative memory and rule based processing by attention requirements (Waters & Caplan, 1999). The control study showed that neither a refractoriness effect nor memory retrieval effects alone can explain the unification pattern found in studies 3 and 4. There results provided evidence for theories proposing that the separation of declarative memory and rule systems are blurred (Hagoort et al, 2009). Furthermore, it is plausible to think that unifying corresponding elements can occur based on emergent abstract categories stored in the memory (Vosse & Kempen, 2000) and by learned contingencies based on plausibility effects (Kutas et al, 2006; MacDonald & Christiansen, 2002).

One of the most intriguing aspects of the experiments presented above is the relationship between the memory mechanisms responsible for the retrieval/activation of lexical elements and those which are accountable for holding online elements for unification.

The human brain is able to store enormous amounts of different kinds of information; historical and self-involved facts, knowledge about other people and their attitudes, ways of making different meals, directions for the shortest route to downtown, or our mother's favorite folk songs. Much has been revealed about this organizational structure of the brain in categorization studies (Kutas & Federmeyer, 2000). These studies suggest that the human brain consistently groups together items that share perceptual and/or functional features (Kutas & Federmeier, 2000). The lexical-grammatical hypothesis claims that the lexical system in the brain is organized into sub-categories corresponding to verbs and nouns (Caramazza & Hillis, 1991). It is possible that all of these categories emerge through experience (MacDonald & Christiansen, 2002) and unification of emergent categories is processed similarly to the unification of idiosyncratic tokens with solely contingent association relationship (Jolsvai & Sussman, in preparation).

Retrieval can be seen as a function of long term memory (Kutas & Federmeier, 2000) while other theorists link retrieval processes to working memory (Baddeley & Lodge, 1999, 2001; Ullman, 2001). Moreover, it is not clear whether processes that keep active the representations needed for an ongoing integration process are separate from those activate or retrieve elements from the knowledge base previously. In the view of models arguing that working memory is the activated portion of long term memory (Cowan, 1998; Ruchkin, 1999; MacDonald & Christiansen, 2002), mechanisms responsible for holding online the uncoupled dependencies represent only different activation pattern of the same network (MacDonald & Christiansen, 2002).

Some studies suggest that N400 reflects the degree of match of the eliciting word and context-based information currently held in working memory (Brown & Hagoort, 1993; Rugg et al, 1994). Comprehension is easier, as indicated by the decreased N400 amplitude when the features of an item (e.g. a word) match with the previous context, suggesting a close link between N400 amplitude and plausibility (Kutas & Federmeier, 2000). In addition, it has been also suggested that N400 amplitude is also sensitive to the ease of accessing information from long-term memory (Kutas & Federmeier, 2000). In other words, N400 correlating with words presented without context depends on frequency and repetition (Neville, 1986). Kutas and Federmeier, (2000) claim that "both the immediate language context held in working memory and the context-independent relationships between items in long-term semantic memory affect the neural processes reflected in the N400" (p.466). If so, working memory processes measured by LAN during online sentence comprehension must at least partially overlap with effects indexing N400. Studies 1 and 2 reinforced the conclusion that a fronto-central negativity reflects some kind of working memory (King and Kutas, 1995) or holding online and unification processes. However, one might wonder, on the basis of other suggestions of WM representing only the activated part of long term memory (e.g. Cowan, 1998, MacDonald & Christiansen, 2000), how closely these two components are intertwined during real life comprehension as well as in experiments that attempt to measure the processes indexed by LAN or N400 components. In an active (attend) study examining uncoupled coherent elements, one cannot be sure of measuring exclusive holding online working memory (LAN) or retrieval (N400) processes, since during comprehension one must retrieve elements and hold them online while retrieving subsequent elements in the lexicon. Similar methodological concerns were articulated by Uttal (2001) and Kutas et al, (2006) who argue that any particular time interval may reflect more than one brain regions or functional processes.

18. References

Aboitiz, F., Bosman, C., Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain and Language*, 98 (387), 40-56.

Anderson, J. E., Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and Language*, 94,(2), 200–216.

Atkinson, R. C., Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In Spence K. W. (Ed.) *The psychology of learning and motivation: Advances in research and theory* (pp. 89–195) New York: Academic Press.

Baddeley A.D, Logie R.H. (1999). Working memory: The multiple component model. In: Miyake A., Shah P., (Eds). *Models of Working Memory*, (pp. 28–61).New York: Cambridge University Press.

Baddeley, A. D. (*1986*). *Working Memory*. *Oxford Univ*. *Press*, *Oxford*.

Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, 56, 851–864.

Baddeley, A.D. (2003). Working memory and language: an overview. *Journal of Communication Disorders* 36, 189–208.

Baddeley, A.D., Hitch, G. (1974). Working memory. In G.H. Bower (Ed).*The psychology of learning and motivation: Advances in research and theory*. (pp. 47--89).New York: Academic Press

Bedini, A. Tonazzini, A. (1992). Image restoration preserving discontinuities: the Bayesian approach and neural networks, *Image and Vision Computing* 10(2), 108-118.

Berwick, R., A. Weinberg (1984). Deterministic Parsing: A Modern View. *Proceedings of the Northeast Linguistics Society,*15-29.

Besson, M., & Schön, D. (2003). Comparison between language and music. In Peretz I. & Zatorre R.(Eds), *Neurosciences and music.* (pp. 269–293).Oxford University Press.

Bresnan, J., Kaplan, R.M. (1982). Introduction: Grammars as Mental Representations of Language. In J. Bresnan (Ed). *The mental representation of grammatical relations.* (pp. xvii-lii),Cambridge, Mass.: The MIT Press.

Brown, C. M., Hagoort, P. (1993) The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience, 5*, 34–44.

Brown, C. M., Hagoort, P. (1999). The cognitive neuroscience of language: Challenges and future directions. In Brown, C. M. Hagoort, P. (Eds). *The neurocognition of language* (pp. 3-14). Oxford: Oxford University Press.

Cabeza, R., Nyberg, L. (2000). Neural basis of learning and memory: Functional neuroimaging evidence. *Current Opinion in Neurology*, 13, 415-421.

Caplan D., Waters G.S (1999). Verbal working memory and sentence comprehension. *Behavior and Brain Sciences,* 22, 77–126.

Caplan, D., Alpert, N., Waters, G., Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Human brain mapping, 9(2), 65-71.

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. San Diego, CA. Academic Press.

Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49(4), 1024-1043.

Chomsky, N., (1965). *Aspects of the Theory of Syntax.* Cambridge: The MIT Press.

Chomsky, N., (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N., Miller. G. (1965) Finite State Languages. In.Luce, R., Bush, R., Galanter, E (Eds) Readings in Mathematical Psychology 2 (pp. 156-171). New York: Wiley and Sons.

Christiansen, M.H. Chater N. (Eds.) (2001). Connectionist psycholinguistics. Westport, CT: Ablex.

Churchland, P. S., Sejnowski, T. (1994). *The Computational Brain.*MIT Press (Cambridge, Mass.).

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence,* 30, 163 – 183.

Coulson, S. (2007). ERP data and cognitive linguistics. In: Gonzales-Marquez, M., I. Mittelberg, S. Coulson and M. Spivey (Eds.), *Methods in Cognitive Linguistics (pp.* 400-423) Benjamins.

Coulson, S., King, J. W., Kutas, M. (1998). Expect the unexpected: event related brain response to morphosyntactic violation. *Language and Cognitive Processes,* 13, 21–58.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin* 96. 341–70.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series, 26. Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In Miyake, A., Shah, P. (Eds.) Models of Working Memory: Mechanisms of active maintenance and executive control (pp. 62-101). Cambridge, U.K.: Cambridge University Press.

Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.

Craik, F.I.M., Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671-684.

Damasio, A. R. and Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences* 90, 11, 4957-4960.

Danneman, M., Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.

DeLong, K. A., Urbach T. P., Kutas M. (2005). Probabilistic word preactivation during language comprehension inferred from electrical brain activity. *Nature Neuroscience,* 8, 1117-1121.

Demberg V., Keller, F. (2008). Data from Eye-tracking Corpora as Evidence for Theories of Syntactic Processing Complexity. *Cognition*, 109(2), 193-210.

Dennett D. (1991). *Consciousness Explained*. Little, Brown and Co. (Boston).

Desmond J.E., Fiez J.A. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in Cognitive Science* 2, (9), 355-362.

Dikker, S., Rabagliati, H., Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, 110, 293–321.

Dobbs, D. (2005). Fact or phrenology? Scientific American Mind 16, 24–31.

Donchin, E. (1981). Event-related potentials in psychological research. *Japanese. Journal of EEG EMG,.* 143-148.

Donchin, E., Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *The Behavioral and Brain Sciences*, 11, .355-425.

Dror I. E., Gallogly D. P. (1999). Computational analyses in cognitive neuroscience: in defense of biological implausibility. *Psychonomic bulletin & review* 6, (2), 173-82.

Elman, J. L. Bates, E.A. Johnson, M.H., Karmiloff-Smith, A. Parisi D., Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*, MIT Press.

Fodor, J. A. (1983). *The Modularity of Mind*, The MIT Press.

Fodor, J. A., Pylyshyn, Z. W. (1988). *Connections and Symbols, chapter Connectionism and cognitive architecture: A critical analysis.* MIT Press, Cambridge, MA

Frazier, L. (1987). Sentence processing: a tutorial review, in: M. Coltheart (Ed.), *Attention and Performance 12: the psychology of reading.* (pp. 559-586). Lawrence Erlbaum Associates, Hove, UK.

Friederici, A. D. (2002): Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 678–684.

Friederici, A. D., Hahne, A., Mecklinger, A. (1996).The temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 22, 1219-1248.

Friederici, A. D., Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, 25, 157-176.

Friederici, A. D., Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, 25, 157-176.

Friederici, A. D., Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain Research*, 1000, 72-77.

Friederici, A. D., Pfeifer, A., Hahne, A., (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1, (3), 183–192.

Friederici, A. D., Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of psycholinguistic Research*, 28, 467-484.

Fuster, J. M. (1995). *Memory in the cerebral cortex*. Cambridge, MA: MIT Press.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. Cognition, 68, 1-76.

Gibson, E. (2000). The dependency locality theory: A distance- based theory of linguistic complexity. In: Marantaz, A. Miyashita, Y., O'Neil W. (Eds.), *Image, language, Brain* (pp. 95-126). Cambridge, MA: MIT Press.

Grodner, D., Gibson, E. (2005). Consequences of the Serial Nature of Linguistic Input for Sentenial Complexity. *Cognitive science*, 29(2), 30, 261-291.

Günter, T. C., Stowe, L. A., Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, 34, 660-676.

Günter, T.C., Friederici, A.D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology,* 36, 126-137.

Günter, T. C., Friederici, A. D., Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience,* 12, 556-568.

Haarmann, H. J., Just, M. A., Carpenter, P. A. (1997). Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain and Language,* 59, 76–120.s

Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *NeuroImage*, *20,*. 18-29.

Hagoort, P. (2005). Broca's complex as the unification space for language. In Cutler, A. (Ed.) *Twenty-first century psycholinguistics: Four cornerstones.* (pp. 157-173). Mahwah, NJ: Erlbaum.

Hagoort, P. (2007). The memory, unification, and control (MUC) model of language. In Meyer, A. S., Wheeldon, L., Krott, A. (Eds.) Automaticity and control in language processing(pp. 243-270). Hove: Psychology Press

Hagoort, P. (2009). Reflections on the neurobiology of syntax. In Bickerton, D., Szathmáry, E. (Eds.) *Biological foundations and origin of syntax* (pp. 279-296). Cambridge, MA: MIT Press.

Hagoort, P., Brown, C. (1994). Brain responses to lexical ambiguity resolution and parsing. In Clifton Jr C., Frazier, L., Rayner, K. (Eds.), Perspectives on sentence processing (pp. 45-81). Hillsdale, NJ: Erlbaum.

Hagoort, P., Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. Neuropsychologia, 38, 1531-1549.

Hagoort, P., Brown, C., Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing*. Language and Cognitive Processes, 8*, 439-483.

Hagoort, P., Wassenaar, M., Brown, C. M. (2003). Syntax-related ERP-effects in Dutch. *Cognitive Brain Research, 16(1),* 38-50.

Hahne, A. Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. Cognitive Brain Research, 13, 339-356.

Hahne, A., Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled. processes, Journal of Cognitive. Neuroscience., 11, 194-205.

Hazy, T.E., Frank, M. J. O'Reilly, R.C. (2007). Toward an executive without a homunculus: Computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society,* 362, 1601-1613.

Hebb, D. O. (1949). *The Organization of Behavior: A neuropsychological theory.* New York: Wiley.

Hickok, G., Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4, 131-138.

Hillis, A. E., Caramazza, A. (1991). Category-specific naming and comprehension impairment: A double dissociation. *Brain,* 114, 2081-2094.

Hoen, M., Domeney, P.F. (2000). ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *NeuroReport*, 11, 3187-3191.

Ivry R.B, Fiez J.A. (2000). Cerebellar contributions to cognition and imagery. In: Gazzaniga MS, (Ed). *The new cognitive neurosciences.* 2nd ed. (pp. 999–1011) Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.

James, W. (1890). *The Principles of Psychology*. Dover Publications.

Johnson, R. Jr., Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, 17, 167-178.

Jolsvai, H., Csépe, V., Mészáros, É., Csuhaj, R. (2006). Number-person agreement in Hungarian and the role of time delay in lexical access. An ERP study. *Pszichologia*, 26, (3), 227-246.

Just, M. A., Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 98, 122–149.

Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, R. K. (1996). Brain activation modulated by sentence comprehension. *Science*, 274, 114–116.

Kaan E., Swaab T.Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*. 6(8), 350-356.

Kempen, G., Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science*, 1, 273-290.

Kinberg, D.Y., Farrah, M.J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex organised behaviour. *Journal of. Experimental Psychology.* 122, 411–428.

King, J. W., Kutas, M. (1995). Who did what and when? Using word and clause level ERPs to monitor working memory usage in reading. Journal of Cognitive Neuroscience, 7, 376–395.

King, J., Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory*. Journal of Memory and Language*, 30, 580–602.

Kluender, R., Kutas, M. (1993). Bridging the gap: evidence from ERPs on the processing of unbounded dependences. *Journal of Cognitive Neuroscience,* 5, 2, 196–214.

Knowlton, B. J., Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science.* 262., 1747-1749.

Kukla, A., Walmsley, J. (2006). *A historical and philosophical introduction to the major theories*. Indianapolis. Hackett Pub. Co.

Kutas, M., Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Science, 4, 463-470.

Kutas, M., Hillyard, S. A.(1980).Reading senseless sentences: brain potentials reflect semantic incongruity *Science,* 207, 203-205.

Kutas, M., Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*. 11, 539–550.

Lau, E., Stroud, C., Plesch, S., Phillips, C. (2006).The Role of Structural Prediction in Rapid Syntactic Analysis. *Brain & Language*, 98, 74-88.

Lehmann, D., Skrandies, W. (1984). Spatial analysis of evoked potentials in man–a review. Prog. Neurobiology. 23 (3), 227–250).

Levelt, W.J.M. (1999). Producing spoken language: A blueprint of the speaker. In Brown, C.M., Hagoort, P. (Eds.), *The neurocognition of Language* (pp. 83-121). Oxford: Oxford University Press.

Luria, A. R. (1966). *Higher Cortical Functions in Man*. Basic Books, New York.

MacDonald, M. C., Christiansen, M., H., (2002). Reassessing Working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review,* 109, 1, 35–54.

Marslen-Wilson, W.D., Tyler, L.K.(1980). The temporal structure of spoken language understanding, *Cognition*, 8, 1-71.

Martin-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., Sommer, W. (2006). Semantics prelevance over syntax during sentence processing. A brain potential study of noun-adjective agreement in Spanish. *Brain Research*, 1093, 178-198.

Mecklinger, A., Ullsperger, P. (1993). P3 varies with stimulus categorization rather than probability. *Electroencephalography and Clinical Neurophysiology,* 86, 395-407.

Menon, V., Anagnoson, R. T., Glover, G. H. Pfefferbaum, A. (2000). A. Basal ganglia involvement in memory-guided movement sequencing. *Neuroreport.* 11(16), 3641-3645.

Mesulam, M. M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. In Principles of Frontal Lobe Function (Stuss, D.T. and Knight, R.T., eds), (pp. 8–31), Oxford University Press.

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., EEG source imaging. Clinical Neurophysiology. 115 (10), 2195–2222.

Mildner V. (2008). The Cognitive Neuroscience of Human Communication. New York. Laurence Erlbaum Associates.

Miller, G. A. (1954). Communication. *Annual Review of Psychology* 5, 12, 401—420.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Miller, G. A. and Chomsky N. (1963). Finitary models of language users. In Luce, L.D. Bush, R.R., Galanter, E. (Eds). *Handbook of mathematical psychology Volume 2*. (pp. 419–491)

Miyake, A., Shah, P. (Eds.) (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Murrey, M., M., Brunet, D., Michel, C. M. (2008): Topographic ERP Analyses: A Step-by-Step Tutorial Review. Brain Topography, 20, 249–264.

Müller, H. M., King, J. W., Kutas, M. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research*, 5, 193–203.

Münte, T. F., Anvari, E., Matzke, M., & Johannes, S. (1995). Brain potential correlates of number errors in the Turkish language. *Neurosci.Lett*., 199, (pp. 57-60).

Münte, T. F., Heinze, H. J. (1994). ERP negativities during syntactic processing of written words. In Heinze, H. J. Münte, . T. F., Mangun G. R. (Eds), *Cognitive electrophysiology (*pp. 211–238) Boston: Birkhauser.

Münte, T. F., Heinze, H. J., Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, 5, 335–344.

Münte, T. F., Szentkuti, A., Wieringa, B. M., Matzke, M., & Johannes, S. (1997). Human brain potentials to reading syntactic errors in sentences of different complexity. *Neuroscience Letters.*, 235, (pp. 105-108).

Näätänen, R., Gaillard, A.W.K., Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. Acta Psychologica, 42, 313-329.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, 53, (pp. 53-81).

Narayan, S., Jurafsky, D. (2002). A Bayesian model predicts human parse preference and reading time in sentence processing. *Advances in Neural Information Processing Systems, 14, 59-65.*

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event- related brain potentials. Journal of Cognitive Neuroscience, 3, 151–165.

Nosofsky, R. M., Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, 9(4), 247-255.

Opitz, B., Friederici, A.D. (2007). Neural basis of processing sequential and hierarchical syntactic structures. Human Brain Mapping, 28, 585-592.

Palohati, M., Leino, S., Jokela, M., Kopra, K., Paavilainen, P. (2005). Event related potentials suggest early interaction between syntax and semantics during on-line sentence comprehension. Neuroscience Letters, 384, 222-227.

Patel, A.D. (2008). *Music, Language, and the Brain*. New York : Oxford Univ. Press.

Pavlov, I. (1927). Conditioned reflexes. London: Routledge and Kegan Paul.

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72 (2), 184–187.

Phillips, S.,Takeda, Y. (2009). Greater frontal-parietal synchrony at low gamma-band frequencies for inefficient than efficient visual search in human *EEG. International Journal of Psychophysiology, 73, 350-354.*

Phillips,C.,Kazaninaa, N., Abada, S.H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22 (2005) 407– 428

Picton, T.W. (1992).The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9,456–479.

Picton, T. W., Lins, O. G., Scherg, M. (1995). The recording and analysis of event-related potentials. In BollerF., Grafman,J. (Eds.) *Handbook of neuropsychology* (vol. 10, pp. 3-73). Amsterdam: Elsevier.

Picton, T. W. (1998). The endogenous evoked potentials. In: Basar, E. (Ed.). *Springer series in brain dynamics*. Berlin, Springer-Verlag.

Pinker, S. (1991). *Rules of Language*. *Science*, 253, 530-535.

Pinker, S. (1994). *The Language Instinct.* New York: William Morrow and Co.

Pleh, Cs. (1998). A *Mondatmegértés a Magyar Nyelvbe*n: Pszicholingvisztikai Kísérletek és Modellek. Osiris Kiadó, Budapest.

Polich, J. (1998). P300 clinical utility and control of variability. *Journal Clinical Neurophysiology* 15.1, 14–33.

Postman, L. (1975). Tests of the generality of the principle of encoding specificity. *Memory* & *Cognition,* 3, 663-672.

Pritchard, W. S. (1981). *Psychophysiology* of the *P300.Psychol Bull 89,* 504–540.

Pulvermüller, F. (1992). Constituents of a neurological theory of language. Concepts in Neuroscience 3, 157-200.

Ramsey, W., Stich, S. and Rumelhart, D. (1991). *Philosophy and Connectionist Theory*. Lawrence Erlbaum Associates (Hillsdale, NJ)

Rickard, T.C., Grafman J. (1989). Losing their configural mind. Amnesic patients fail on transverse patterning. *Journal of Cognitive Neuroscience,* 10, 509–524.

Ritter, W., Vaughan, H.G., Jr, Costa, L.D. (1968) Orienting and habituation to auditory stimuli: a study of short term changes in average evoked responses. *Electroencephalogr Clin Neurophysiol* ,25, 550–556

Ruchkin, D. S., Grafman, J., Cameron, K., Berndt, R. S., (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences,* 26, 709–728.

Rugg, M. D. and Doyle, M.C. (1994). Event-related potentials and stimulus repetition in direct and indirect tests of memory. In Heinze, H J. et al (Eds) Cognitive Electrophysiology. (pp. 124–148) Birkhäuser.

Rumelhart, D. E. (1989) The architecture of Mind: A connectionist Approach. In M. Posner (ed.) Foundations of Cognitive Science, Cambridge, MA: MIT Press

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning internal representations by error propagation. In Rumelhart, D.E., McClelland, J.L., and the PDP Research Group, (Eds). Paralled Distributed Processing. Explorations in the

Microstructure of Cognition. Volume 1: Foundations, (pp. 318-362.). The MIT Press, Cambridge, MA.

Seidenberg M S. (1997). Language acquisition and use: learning and applying probabilistic constraints. Brain and language, 7;57(2), 254-79.

Semendeferi, K., Damasio, H., Frank, R., Van Hoesen, G.W. (1997). The evolution of the frontal lobes: a volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution*, 32, 375-388.

Shapiro, K., Caramazza, A. (2003). The representation of grammatical categories in the brain. *Trends in Cognitive Sciences.* 7, 201-206.

Shön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition* 106, 975–983.

Singer, W. (2001). Consciousness and the binding problem. *Ann N Y Acad Sci*.,929, 123-46.

Smith, E.E., Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33, 5-42.

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, 19, 1493-1503.

Streb, J., Rösler, F., Henninghausen, E. (1999): Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. Brain and Language, 70, 273–286.

Stromswold, K., Caplan, D., Alpert, N., Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomogra- phy. *Brain and Language*. 52., 452–473.

Sussman E. (2007). A new view on the MMN and attention debate: Auditory context effects. J Psychophysiology, 21., 164–175.

Swinney, D., Zurif, E. (1995). Syntactic processing in aphasia. *Brain and Language*, 50, 225-239.

Swinney, D., Zurif, E., Prather, P., Love, T. (1996). Neurological distribution of processing operations underlying language comprehension. *Journal of Cognitive Neuroscience*, 8:2, 174-184.

Thagard, P.(2002). How Molecules Matter to Mental Computation. *Philosophy of Science*, 69, 429–446.

Thibadeau, R. H., Just, M. A., & Carpenter, P. A. (1984). A model of the time course and content of reading. *Cognitive Science*, 6, 157–203.

Thibodeau, R., Just, M. A., Carpenter, P. A., (1984). A model of the time course and content of reading, Cognitive Science 6, 157-203.

Thierry, G., Cardebat, D. (2003). E lectrophysiological comparison of grammatical processing and semantic processing of single spoken nouns. *Cognitive Brain Research,* 17, 535-547.

Tomasello, M. (2003). Constructing a Language: A Usage-Based Approach to Child Language Acquisition. Cambridge, MA: Harvard University Press.

Townsend, D. J., Bever, T.G. (2001) *Sentence Comprehension: The Integration of Habits and Rules*. MIT Press

Turing A. M. (1950). Computing Machinery and Intelligence. *Mind 49*, 433-460.

Ueno, M., Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, 86, 243-271.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1), 37-69.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/ procedural model. *Bilingualism: Language and. Cognition,* 4(1), 105-122.

Ullman, M. T. (2004). Contributions of neural memory circuits to language: The declarative/procedural model. Cognition, 92(1-2), 231-270.

Ullman, M. T., Pierpont, E. I. (2005) Specific Language Impairment is not specific to language: the Procedural Deficit Hypothesis. *Cortex,* 41 (3), 399-43.

Uttal W. R. (2002). Précis of the New Phrenology: The Limits of Localizing Cognitive Processes in the Brain. *Brain and Mind* 3 (2), 221-228.

Uttal, W.R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain.* MIT Press.

Van Berkum, J. J. A., Brown, C. M., Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language,* 41(2), 147-182.

Van Berkum, J. J. A., Brown, C. M., Hagoort, P., Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, 40(2), 235-248.

Van Herten, M., Kolk, H.H.J., Chwilla, D.J., Vissers, C.T.W.M. (2004). P600 as a monitoring component: The effect of partial sentence plausibility. *Journal of Cognitive Neuroscience, 16*, D- 103.

von Neumann, J. (1951). Various techniques used in connection with random digits. *National Bureau of Standards Applied Math Series* 12(36).

Vos, S. H., Günter, T. C., Kolk, H. H. J., Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*, 38, 41-63.

Vosse, T., Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 75, 105-143.

Vosse, T., Kempen, G. (in press). The Unification Space implemented as a localist neural net: Predictions and error-tolerance in a constraint-based parser. Cognitive Neurodynamics, 3.

Waters, G. S., Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter1992. Psychological Review, 4, 761–772.

Watson, J. B. (1924). *Behaviorism*. New York: Norton.

Willems, R. M., Hagoort, P. (2009). Broca's region: Battles are not won by ignoring half of the facts. *Trends in Cognitive sciences,* 13(3), 101.

Woldorff, M.G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, 30, 98-119.

19. Magyar összefoglaló

Összetartozó de a szekvenciában szétválasztottan szereplő elemek egyesítése

Valamennyi kísérlet alap hipotézise, hogy összetartozó elemek feldolgozása egy elosztott halózatban történik, ahol a munkamemória és tudás, valamint maga a feldolgozás a hálózatban implementált.

Az első két kísérlet a munkamemória és integráció funkcionális kapcsolatát vizsgálja a nyelvfeldolgozásban. Kombinált pardigmánkban morfoszintaktikai sértést kombináltuk az elemek közötti távolság növelésével (Kísérlet 1: komplexitás és Kísérlet 2: Idő). Részben átfedő fronto-centrális negativitásokokat regisztráltunk. Mindkét változó (munkamemória és integráció) konzisztensen korrelált Fc1 és Fc2 elektródahelyek feletti negativitással amíg a integració frontal-frontocentrális és munkamemória fronto-centrális negativtással korrelált. A különböző kondiciókban kiváltott komponensek topográfiai eloszlásának lehetséges különbségeinek vizsgálata érdekeben Globális disszimilaritás indexet (Lehmann & Skrandies, 1984) számoltunk. Az egyes feltételekben kiváltott komponensek páros összehasonlítása topográfiai homogenitést indikált. Ezek az eredmenyek a CSD térképekkel együtt arra utalnak, hogy munkamemória funkcionálisan elválaszthatatlan az integrációs folyamattól, vagyis ezek egy szétosztott hálózatban valósulnak meg, elkülönült funkcionális feldolgozási egységek helyett.

A másik három kísérlet a deklaratív memória és a procedurális szabályrendszer (Ullman, 2001) funkcionális elkülönülésének relevanciáját vizsgálja. Néhány elmélet egy automatikus szabályrendszer és egy kontrollált deklaratív memória elkülönítését hírdeti (pl. Ullman, 2004, Pinker,1991), míg más elépzelések tagadják ezt (Vosse & Kempen, 2000). Neurális hálózatmodellek szerint kategóriák nem egy elkülönült szabályrendszer által kezelődnek, hanem kiemelkednek a környezetben előforduló statisztikai assziációk alapján (pl. Rummelhart, 1999).

Tanult asszociatív ismerős dal szekvenciák minden absztrakt szabályt nélkülözve hasonló neurofiziológiai válaszokat eredményeztek figyelt és nem figyelt helyzetekben, mint amilyeneket az első két kísérlet absztrakt nyelvi kategóriái kiváltottak. Ezek az eredmények a konnekcionista elképzeléseket támogatják.