Attentional and Semantic Anticipations in Recurrent Neural Networks

Frédéric Lavigne[1] and Sylvain Denis
Laboratoire de Psychologie Expérimentale et Quantitative
Université de Nice – Sophia Antipolis
24 Ave des Diables bleus 06357 Nice Cedex 4, France

Abstract

Why are attentional processes important in the driving of anticipations? Anticipatory processes are fundamental cognitive abilities of living systems, in order to rapidly and accurately perceive new events in the environment, and to trigger adapted behaviors to the newly perceived events. To process anticipations adapted to sequences of various events in complex environments, the cognitive system must be able to run specific anticipations on the basis of selected relevant events. Then more attention must be given to events potentially relevant for the living system, compared to less important events.

What are useful attentional factors in anticipatory processes? The relevance of events in the environment depend on the effects they can have on the survival of the living system. The cognitive system must then be able to detect relevant events to drive anticipations and to trigger adapted behaviors. The attention given to an event depends on i) its external physical relevance in the environment, such as time duration and visual quality, and ii) on its internal semantic relevance in memory, such as knowledge about the event (semantic field in memory) and anticipatory power (associative strength to anticipated associates).

How can we model interactions between attentional and semantic anticipations? Specific types of distributed recurrent neural networks are able to code temporal sequences of events as associated attractors in memory. Particular learning protocol and spike rate transmission through synaptic associations allow the model presented to vary attentionally the amount of activation of anticipations (by activation or inhibition processes) as a function of the external and internal relevance of the perceived events. This type of model offers a unique opportunity to account for both anticipations and attention in unified terms of neural dynamics in a recurrent network.

Keywords - anticipations – attention - context - dynamic memory - recurrent neural networks - semantic priming -

[1] To whom requests for reprints should be addressed: E-mail: lavigne@unice.fr

1 Semantic and attentional anticipations

Why are attentionnal processes important in the driving of anticipations?

1.1. Associative semantic anticipations

Anticipatory processes allow living systems to rapidly adapt their behaviors to events encountered in their environment (e.g., objects, scenes, and behaviors). Behavioral responses adapted to a given event perceived in the environment can be more rapid and accurate when the perceived event was anticipated by the cognitive system. According to the general concepts of anticipation (Dubois, 1998a; Rosen, 1985), anticipations are driven in memory on the basis of semantic knowledge (Lavigne & Lavigne, 2000). Then internal representations about the relations between events occurring in the environment and possible future events are important for the living system to anticipate adapted behavioral responses (see Lavigne & Lavigne, 2000 for a presentation). For this the cognitive system stores associations between events perceived in sequences. Events frequently occurring closer in sequences are learned as associated in memory. Within the framework of experimental studies of reading behavior in cognitive psychology, semantic knowledge is based on associations in memory between word meanings (concepts), which correspond to perceived events during the activity of reading. The associative strength between (words) events is learned from the (textual) environment and depends on their frequency of co-occurrence (Conrad, 1972; Freedman & Loftus, 1971; Landauer, Foltz & Laham, 1998; Foltz, Landauer & Dumais, 1997; Perlmutter, Sorce, & Myers, 1976; Spence & Kimberly, 1990). When an event occurring in the environment is perceived (e. g., 'an approaching snake'), dynamic activation processes propagate through associations in memory. Then the cognitive system can activate (i. e. anticipate) events which has not yet occurred in the environment but which are likely to occur in the very near future on a probabilistic basis (e. g., 'a dangerous attack').

Anticipated (words) events being already activated in memory when they are actually perceived, their perceptive processing (lexical access in memory) can be accomplished more rapidly. Then reading behavior can be enhanced by shortening fixation durations or lengthening saccades sizes. Then a reader's oculomotor behavior can be finely adapted to (words) events perceived in a sentence as a function of anticipations triggered in memory by previously perceived (words) events (Balota & Rayner, 1991; Keefe & Neely, 1990; Neely, 1991; Neely & Keefe, 1989; Neely, Keefe & Ross, 1989; Rayner & Balota, 1989). For example, a perceived target word ('attack') is more rapidly processed (about 550 ms) if already activated in memory according to an associated preceding context ('snake'), and is more slowly processed (about 600 ms) when preceded by a non-associated context ('cloud').

The natural environment of a living system consists of simple and complex events occurring in sequences or perceived in sequences. Events in the environment correspond to simple objects (natural, artifactual, living, etc.), complex situations (correlations of objects, interactions between the living system and other objects, etc.), or are abstract concepts elicited by perceived events (social representations, etc..). They are all represented as events at different levels of abstraction and are memorized as sequences organized in time. The memorization of sequences in memory allow the cognitive system to anticipate possible future events from actually perceived events. The simplest type of sequence of events is of a perceived first event triggering anticipation in memory of a second event not yet perceived. This type of anticipation needs to take into account only one perceived first event to activate associated ones in memory, the more activated (associated) corresponding to the one more likely to occur. However natural environments are complex, within which living systems are surrounded by sequences of numerous events. When several events are perceived, the cognitive system can trigger several anticipations in parallel, which can be coherent with each other or not (i. e., leading to compatible behavioral responses or not) depending on the associations between the sequentially perceived events (Lavigne & Vitu 1997; Lavigne & Lavigne, 2000; Masson, 1991, 1995).

In addition to simple activation of an event in memory from a perceived one, anticipations in complex environments imposes the cognitive system to select the most adapted anticipations among a set of several events perceived in sequences in the environment. Attention given to perceived events is therefore important to evaluate the relevance of a perceived event, and to trigger anticipations leading to behavior adapted to the more important events encountered (see Laberge, 1995; Lecas, 1992; Jones & Yee, 1993). Perceived event's evaluated relevance can then help selection processes of the more adequate anticipation leading to the more adapted behavior.

The purpose of this article is to present experimental results, theoretical views and a neural network model of anticipatory semantic and attentional processing. In Section 1 we present anticipatory semantic and attentional processes allowing several different anticipations from sequences of several perceived events, attentional evaluation of the relative relevance of the perceived events, and selection of the more adequate anticipations for adapted behavior. In Section 2 we define attentional factors allowing to evaluate the relevance of perceived events to trigger anticipations, such as processing time of the event, time elapsed after processing the event, processing load elicited by the event in memory. This allows defining common associative and temporal properties of semantic and attentional anticipatory processes, as well as their adaptive properties. In Section 3 we present an attractor neural network model giving simulations of the functioning of both semantic and attentional anticipatory processes based on a common and unique neural architecture.

1.2. Attentional drive of semantic anticipations

When at least two events triggering different anticipations are perceived at the same time or close in a sequence, the cognitive system must select the best possible anticipation activating the more probable event to occur. When perceiving sequences of events the system must not only (i) anticipate events from every perceived event, but must also (ii) anticipate events from the whole sequence of events, or at least (iii) anticipate from the most relevant event in the sequence.

(i) Every perceived event ('snake') triggers automatic propagation of activation in memory through links to associated events ('attack') (Anderson, 1983; Collins & Loftus, 1975; Collins & Quillian, 1969; Neely, 1991; Thompson-Schill, Kurtz, & Gabrieli, 1998; VanVoorhis, & Dark, 1995). These rapid and unconscious activations of associated events in memory (see Posner & Snyder, 1975a,b) are anticipations which do not last very long in memory (Keefe & Neely, 1990; Neely, 1991; Neely & Keefe, 1989; Neely, Keefe & Ross, 1989; Neely, 1976, 1977). Indeed, when subliminal (i. e., processed under the threshold of conscious perception), events allow only unconscious semantic anticipations (i. e., anticipations activated in memory under the threshold of consciousness). They can predict events only very closely related in time (a few milliseconds, Greenwald, Draine & Abrams, 1996). However, when supraliminal events allow conscious attentional control of the anticipations in memory, anticipations can be sustained longer and predict events far off in time (Balota, Black, & Cheney, 1992; Fuentes, Carmona, Agis, & Catena, 1994; Fuentes, & Ortells, 1993; Fuentes, & Tudela, 1992; Neely, 1991). So a role of attention is to maintain anticipations activated in memory longer to give them more predictive efficacy with time.

(ii) Two or more events perceived in the environment can already be associated together ('snake' and 'fang') and have common associates in memory ('attack'). Then, when perceived at a same time or close in a sequence, they are coherent and can trigger compatible anticipations leading to behaviors adapted to both of them. Furthermore, anticipation triggered by one event is amplified by other events triggering similar anticipations. This corresponds to additive activation processes triggered by two or more words on common associates in memory (Balota & Paul, 1996; Brodeur & Lupker, 1994; Lavigne & Vitu, 1997). The anticipation additively activated by several (words) events is then more available in memory for further attentional processing.

(iii) Two or more events can be incoherent if not associated in memory or not sharing common associates ('snake' and 'wasp'). Then they can trigger different anticipations leading to different motor responses corresponding to incompatible behaviors, each adapted to only one of the perceived events ('to walk back away the snake' and 'to wave off the wasp', respectively). Under the assumption that one behavior can be accomplished at a time, the actual adopted behavior must be adapted to the most relevant event with regard to its effects on the survival of the living system. This implies a selection among several anticipations of possible behaviors (Glenberg, 1997; see Berthoz, 1996), which correspond to incursion and hyperincursion in memory, for which one future state is selected in the system among several potential ones (Dubois, 1996, 1998b). The attentional selection in memory must involve inhibitory processes operating on activated anticipations, to eventually maintain activated only the selected anticipation corresponding to the most relevant event perceived (Laberge, 1995; Lecas, 1992; Jones, 1976; Jones & Boltz, 1989; Jones & Yee, 1993; Neely, 1991; Posner & Snyder, 1975). In this case attention plays a role in selecting anticipations in memory by activating the appropriate ones and inhibiting the inappropriate ones.

2. Attentional and semantic factors driving anticipations

What are useful attentional factors in anticipatory processes?

2.1. Attentional relevance of perceived events to drive anticipations

The relevance of events in the environment depends on the effects the events can have on the survival of a living system. With regard to their associated potential danger, some events have no or weak effects (e. g., 'a snake' or 'a wasp') on the internal state of a given system. Some events can kill the system and must be given priority in anticipating adapted behavioral responses. The cognitive system must then be able to detect the most relevant event in a sequence, in order to drive adequate anticipations and to adopt a behavior adapted to this event (see Laberge, 1995; Lecas, 1992; Jones, 1976; Jones & Boltz, 1989; Jones & Yee, 1993).

Attention must then be allocated to relevant events to orient anticipations by the system to behaviors presenting the greater adaptive value. To achieve this aim a fundamental role of attention is the evaluation of the relevance of the perceived events, in order to select behaviors adapted to relevant events and avoid behaviors not in relation to relevant events (Broadbent, 1982). More attention is then given to relevant events in order to drive anticipations. The attention given to an event depends on i) its external physical relevance in the environment, such as persistence of the event in the environment and visual perceptibility, and ii) on its internal semantic relevance in memory, such as anticipatory power (associative strength to anticipated associates) and knowledge about the event (familiarity and semantic field in memory) (see Broadbent, 1971; Shiffrin, 1988).

2.1.1. External physical relevance in the environment

Physical properties of the events themselves define their perceptive salience, and can be cues for attentional processes in their perceptive selection, independently of the semantic knowledge the cognitive system has about them (cf., the signal detection theory: Tanner & Swets, 1954; Green & Swets, 1966). Two physical properties are of particular importance for attentional processing: (iv) processing time and (v) perceptibility.

(iv) During reading, behavioral responses (e. g. eye movements or identification times, see Lavigne & Lavigne, 2000) adopted on anticipated target-words are influenced by the duration of the preceding prime words which led to the anticipations (Greenwald et al., 1996; Lorch, 1982; McNamara, 1994; Ratcliff & McKoon, 1981). The longer the prime-word is perceived the more it activates associated target-words in memory and facilitates behavioral responses to target-words. More generally, the longer attention is given to a (word) event, the more it can lead to anticipations.

When two prime-words are perceived in a sequence during reading, they trigger different and incompatible anticipations if they do not share common associates. Then the prime which is processed longer benefits more attention and can cancel anticipations made from the other (Lavigne & Vitu, 1997; see Lavigne & Lavigne, 2000; Neely, 1991 for reviews). This corresponds to general properties of attention in which events perceived longer activate their associates in memory at the expense of other perceived events (Posner & Snyder, 1975a,b; Posner, 1980; Posner & Cohen, 1984).

Therefore the amount of time during which an event is perceived determines the amount of attention given to this event and its ability to trigger anticipations. When perceiving different events triggering competing anticipations in parallel in memory (e. g., 'a snake' and 'a wasp'), processing time would be a cue of relevance of an event (e. g., 'a still snake' *vs*. 'a rapidly fleeing wasp'). Increasing attention with processing time would lead to selection processes maintaining the most relevant anticipation activated (e. g., 'walking back from the still snake') and inhibiting the other ones (e. g., 'waving off the fleeing wasp').

(v) Perceptibility of an event can also influence attentional processing of the event. During reading, anticipatory activation of associated target-words in memory depends on quantitative and qualitative roles of attention as a function of the perceptibility of a preceding prime-word. In case of very shortly perceived words (10 to 30 ms), perceptibility is diminished when the word is visually masked by non-verbal visual stimuli (e. g., a row of X's, random dots or a random letter string like 'skefgklj'; see Holender, 1986 for a review). Shortly presented and masked prime words lead to unconscious processing where no attentional control is possible. Only automatic processes occur to generate semantic anticipations on associated target-words in memory, these priming effects being weaker than when the prime-words are fully and attentionally processed (Greenwald et al., 1996; Holender, 1986). Furthermore, during reading as well as in many situations of perceiving events in the environment, (word) events can be foveally or parafoveally perceived. Parafoveally perceived words are unconsciously processed and benefit less attention, leading to weaker anticipatory priming effects than foveally perceived words which benefit from greater attentional processing (Fuentes, Carmona, Agis, & Catena, 1994; Fuentes, & Ortells, 1993; Fuentes, & Tudela, 1992; Lavigne & Dubois, 2000; Lavigne, Vitu, & d'Ydewalle, 2000). This is coherent with the general effects of perceptive salience influencing attention to a perceived event (Tanner & Swets, 1954; Green & Swets, 1966). In this case attentional processing can play a quantitative role in enhancing semantic anticipations.

A more qualitative role of attention arises from experimental studies on foveal and parafoveal word processing. Both consciously and unconsciously perceived words lead to automatic and unconscious anticipations (Neely, 1991), their strength varying with attention (quantitative effect of attention). However, when two prime words are perceived at the same time and trigger incompatible anticipations (e. g., 'snake and wasp'), only foveally perceived words (e. g., snake'), which benefit from more attention, can lead to anticipations (e. g., 'walking away') that inhibit other anticipations (e. g., 'waving off') (Fuentes, Carmona, Agis, & Catena, 1994; Fuentes, & Ortells, 1993; Fuentes, & Tudela, 1992). Then the ability of attention to not only increase the strength of anticipations but also to inhibit other anticipations corresponds to a qualitative role of attentional drive of anticipations.

To resume, attentional drive of semantic anticipations can vary as a function of external physical properties of perceived events, such as perception time and perceptibility. Exogenous selective attention allows a selection of relevant events (see Laberge, 1995; Sperling & Reeves, 1980; Weichselgartner & Sperling, 1987), to drive semantic anticipations in memory differentially as a function of their physical salience.

2.1.2. Internal semantic relevance in memory

In addition to external properties, internal properties can modulate the relevance of the perceived events. These internal properties depend on the knowledge the cognitive system has about the events in its environment. They are learned from previous encounters with the events as a function of (vi) their frequency of occurrence, (vii) the frequency of co-occurrence of two events, and (viii) the frequency of co-occurrence of several events.

(vi) Events can be encountered and learned more (e. g., 'a peach') or less (e. g., 'a cherimoya') frequently in the environment. The more frequently a word is read, the more knowledge we have about it as a visual form. It can then be more rapidly accessed in memory and identified for further reading, because of its higher level of activation in memory than other less frequent words (see Monsell, 1991 for a review). Although high-frequency words are more activated, activation thresholds put a limit to this level (Rumelhart & McClelland, 1981, 1982). Therefore, an important parameter is that low-frequency (words) events need more time to be accessed in memory and are processed longer (Rayner & Balota, 1989; Vitu, 1991; Lavigne, Vitu & d'Ydewalle, 2000). A consequence is that a less frequently encountered event can activate anticipations of associated events in memory for longer time. This is consistent with the effects of habituation to frequently encountered events, and of attention given to less frequently encountered events (Tipper, Bourque, Anderson & Brehaut, 1989), in the sense that anticipations which benefit from longer activation are given more attention. To generalize, infrequent events (e. g., 'a cherimoya') are privileged as more relevant by the attentional system to drive anticipations at the expense of frequent events (e. g., 'a peach').

(vii) Event frequency alone can not account for every attentional drive based on internal knowledge about the event. Indeed, one can drive strong anticipations from both infrequent and frequent events (e. g., 'a cherimoya' and 'a peach' respectively) if one has strong knowledge about them (e. g., 'it tastes very good'). Knowledge about the taste of a fruit depends on the co-occurrence of the eating of the fruit and of its flavor. The strength of the association between two events (or words or concepts) in memory is largely determined by their frequency of co-occurrence (Conrad, 1972; Freedman & Loftus, 1971; Foltz, Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998; Perlmutter et al., 1976; Spence & Kimberly 1990). The more two events are encountered together at the same time or close in time, the more they are learned together (associated) and the more one of the two events can trigger strong anticipations of the other event (Becker, 1979; Lorch, 1982). More generally, events strongly associated in memory to a perceived event benefit from more activation and are given more attentional relevance during anticipatory processes (see Jones & Yee, 1993).

(viii) Basic knowledge based on binary associations is not the only semantic factor which can influence attention given to anticipations. Associative norms are constructed experimentally in collecting words given by persons as associated to prime-words (see Lavigne & Lavigne 2000; Ferrand & Alario, 1999). The more a word is associated in memory to a prime- word, the more it is given as the first associate to come to mind when reading the prime-word. These associative norms show the variability in associative strength between words, and also that the number of different words given as associates can greatly vary among different prime-words. A given word (e. g., 'snake') is represented in memory through associations to several others (e. g., 'fangs, 'tail', 'reptile', 'rapid', 'dangerous', 'attack', 'poison', 'death', etc.). Not only binary associations between a prime-word and one of its associates (e. g., 'snake' and 'attack'), but many associations between a prime-word and all its associates (e. g., 'snake, 'fang', 'poison', …, 'attack') define the semantic field of the prime-word. Depending on the learned co-occurrences between a prime-word and a variable number of co-occurrent words, the semantic field size may vary from large ('snake' has many associates) to small ('auburn' has few associates). The larger the semantic field, the more activation propagates within the field from the prime-word to many other associates. Because many associates transmit their activation to a given associate in the field, the level of activation of each associate is higher when the field size is large than when it is small (Lavigne et al., in preparation). Then perceived events for which one has the more knowledge (i. e., which have large semantic fields) are more relevant for attentional processes and lead to stronger anticipations.

To resume, internal cognitive factors determine attention which is sustained through time to anticipate possible upcoming events in a sequence (Jones, 1976; Jones & Boltz, 1989; Jones & Yee, 1993).

2.2. Common structures and processes for semantic and attentional anticipations

Attention is a well-defined concept in cognitive models leading to various fundamental processes in semantic anticipations. However it is important to define theoretical properties of attention in terms of actual structures and processes in order to propose common properties of a neural network model of both attention and semantic anticipations. The discussion of experimental results and theoretical views strongly suggests that attentional drive of semantic anticipations involves levels of activation of event representations in memory. A common associative structure for semantic and attentional anticipations can be proposed: event representations are associated in memory, and activation propagates through the associative network from activated events to associated ones. The variable level of activation of the event representation can be determined by semantic anticipations themselves as well as by attentional control. Then semantic anticipations, running on associations between events, and attentional drive, based on physical and semantic properties of the events, interact to modulate the degree of activation of the representations of anticipated events. Through these interactions of dynamic processes based on a common associative structure, attention appears to be influenced by semantic structures in memory and semantic anticipations are influenced by attentional modifications of event representations. A common neural structure can then be presented that precisely models both attention and semantic anticipations in terms of common neural networks dynamics.

3. Recurrent attractor neural network model with delayed neuronal activities

How can we modelize interactions between attentional and semantic anticipations?

From previous models able to code temporal sequences of perceived events as associated attractors in memory (Amit, 1989; Amit et al., 1994, Brunel, 1994, 1996), a modified and extended version of a recurrent neural network was presented to modelize semantic anticipatory processes (Lavigne & Lavigne, 2000). Mathematical properties of a new model are presented as well as simulations of interactions between attentional and semantic anticipations.

3.1. Network architecture

The network is a local module similar to a cortical column connected to other areas of the cerebral cortex (see Brunel, 1996). It is made of 1000 neurons, 750 excitatory $_{(E)}$ and 250 inhibitory $_{(I)}$ neurons, with equal probability of having a synapse on any other neuron. (connectivity parameter c = 0.1). The network has then S_{EE} = 56250 excitatory to excitatory synapses, S_{EI} = S_{IE} = 18750 excitatory to inhibitory and inhibitory to excitatory synapses, and S_{II} = 6250 inhibitory to inhibitory synapses.

Excitatory neurons code for events perceived by the network and inhibitory neurons prevent runaway propagation of activation throughout all the excitatory neurons and maintain stable states in the network.

3.2. Neuron properties

Neurons are connected through four types of pre-synaptic (j) to post-synaptic (i) synapses. Synaptic efficacies correspond to post-synaptic potentials (mV) provoked by a spike. They are initially randomly defined as follows with respective means Jij_{EE}=0.04 mV (excitatory to excitatory), Jij_{EI}= 0.05 mV (excitatory to inhibitory), and Jij_{IE}=Jij_{II}= O.14 mV (respectively inhibitory to excitatory and inhibitory to inhibitory), with a synaptic variability Δ=J.

3.3. Neuron dynamics

All neurons in the network are leaky integrate-and-fire neurons converting input currents I (mV) in firing rates νi (spikes.s-1), according to the transduction function

$$\nu_i = \Phi(I) = O\,/\!\!/_{[0,I1[} + [(I-\alpha)(\beta I + \chi)]\,/\!\!/_{[I1,I2[} + [\delta I - \varepsilon]\,/\!\!/_{[I2,+\infty[} \tag{1}$$

approximating Brunel's (1996) values for Ricciardi's (1977) transduction function, with $/\!\!/_{[Ix,Iy[} = 1$ for the corresponding intervals of I, O if not, I1 = 15, I2 = 25, $\alpha = 13$, $\beta = 0.2$, $\chi = 11$, $\delta = 4$, $\varepsilon = 40$.

A neuron receives a total input intensity

$$Ii_{(tot)} = Ii_{(ext)} + \tau_E \Sigma\, \nu_{j(E)} J_{ij(E)} - \tau_I \Sigma\, \nu_{j(I)} J_{ij(I)} + \tau_{(t)} Ii_{(\mu)} \tag{2}$$

$Ii_{(ext)}$ is the external input current received by 50% of the neurons from the other cortical areas outside the network. The distribution of $Ii_{(ext)}$ has mean $I_{(ext)}$ = 11 mV and σ = 0.9 mV.

$\tau_E \Sigma\, \nu_{j(E)} J_{ij(E)}$ is the internal input current received by the neurons from excitatory neurons; and $\tau_I \Sigma\, \nu_{j(I)} J_{ij(I)}$ is the internal input current received by the neurons from inhibitory neurons; with τ_E = 0.01 and τ_I = 0.002 the time constants for excitatory and inhibitory neurons respectively, ν_j the spike rates of neuron i and s and J_{ij} the synaptic efficacies from neuron j to neuron i.

$\tau_{(t)} Ii_{(\mu)}$ is the external input current when an event μ is perceived, applied to excitatory neurons coding for the corresponding event μ. $\tau_{(t)}$ is the time variable slowly increasing with perception duration (t) of the event, which guarantees slow spike rate dynamics during event perception.

3.4. Learning dynamics

Synapses connecting excitatory neurons (JEE) coding for perceived events are plastic and sensitive to hebbian learning. Synaptic dynamics incorporates both associative long term potentiation (LTP) and depression (LTD) defining modifications of the synaptic efficacies J_{ij} between neurons j and i (Amit & Brunel, 1995):

$$\tau_c dJ_{ij}/dt = - J_{ij} + C_{ij} + J_{0/1} \tag{3.1}$$

calculated in the network as

$$J_{ij(t+1)} = (\tau_c - 1)J_{ij(t)} / \tau_c + C_{ij(t)}/\tau_c + J_{0/1}/\tau_c \tag{3.2}$$

J_{ij} vary according to the time constant $\tau_c = 20$.

$J_{0/1}$ takes the minimum (J_0 = 0.04) or maximum (J_1 = 0.15) values when J_{ij} crosses (getting respectively lower or upper) a threshold w_{ij}, which stochastically vary between $J_0 + \theta$ and $J_1 - \theta$, with $\theta = 0.04$ with steps of $\xi = 0.01$ mV.

Potentiation or depression of the synapse is given by the values of $C_{ij(t}$ defined by the Hebb learning rule according to Brunel (1996):

$$C_{ij(t)} = \lambda_{+} v_{i(t)}\ v_{j(t)} - \lambda_{-}[v_{i(t)} + v_{j(t)}] \tag{4}$$

$v_{i(t)}$ and $v_{j(t)}$ are the spike rates of neurons i and j respectively, and $\lambda_{+} = 0.0005$ and $\lambda_{-} = 0.004$ are the potentiation and depression parameters respectively.

3.5. Network dynamics

Each cycle in the network consists in a random updating of the spike rates of the neurons as a function of the intensities they receive. When currents are received only from outside the network and from other excitatory and inhibitory neurons (equation 2), neurons emit about 3.9 spikes per second (equation 1) and the network has a stable state of spontaneous activity.

In order to simulate slow variations of attentional activation of the attractors in the network, slow network dynamics are guaranteed by a variable increase of input intensity $Ii_{(\mu)}$. A perceived event slowly increases the total input intensity $Ii_{(tot)}$ to simulate attentional activation as a function of perception duration.

Before learning, synaptic efficacies are randomly distributed, and no or few changes occur when spike rates are low. Before learning, the network has no structured attractor corresponding to events stored in memory. After learning of sequences of events, learned attractors coding for each event correspond to neurons activated by the event, which are strongly associated. When perceiving the corresponding event, neurons in a same attractor transmit activation within the attractor, the activation being sustained and progressively decreasing through time after removal of the perceived event.

4. Network simulations of attentional and semantic anticipations

The neural network model presented allows to define long term and short term memories as different internal states (association/activation) of attractors coded in a same neural structure. This type of model presents several interests including its neurobiological plausibility, its ability to fit the external behavior of the system such as associative learning and activation processes, and most importantly its accounting for internal cognitive properties of the system such as the time course of activatory and inhibitory processes as well as attentional processes. This last feature gives the model a strong cognitive plausibility, making it an explicative model of internal processes not limited to predictive abilities of the end product of the processes (see Perfetti, 1998). Indeed, this model internally functions in accordance to basic properties of the cognitive system, a crucial point when attempting to model attentional drive of anticipatory processes.

4.1. Semantic and attentional learning of co-occurrences and similarities

Each event perceived by the network is coded as patterns of activation across a subset (10 neurons) of the entire network (1000 neurons). Events are coded in a distributed way by several neurons so that each event can be a complex event, corresponding to conjunctions of sub-events coded by individual neurons or small groups of neurons. Patterns memorized by the network are also non-orthogonal in the sense that they do not share neurons. This means that learned events are not associated in a pre-defined way by common neurons, but are associated through learning depending only on co-occurrence of events in the network environment.

Given that the attractor of a perceived event decreases slowly through time after removal of the event, neurons coding for a first event are still activated when a following event is perceived. This property of the network allows it to associate attractors corresponding to events occurring frequently in temporal sequences, that is to co-occurrent events. The model is then able to perform (ix) semantic learning based on the events encountered, and (x) attentional learning based on its internal cognitive states.

(i) Semantic learning is achieved by an unsupervised learning mechanism involving the Hebb-like rule (equation 4) varying synaptic efficacies and associating neurons coding for successive events (equations 3.1. and 3.2.; see Brunel, 1996; Lavigne & Lavigne, 2000). After learning, the network has many attractors corresponding to learned events. The attractors are associated as a function of the temporal co-occurrence between the perceived events. Semantic learning in the network then corresponds to binary associations between representations of events perceived in temporal contiguity (e. g., ‘peach’ and ‘good taste’). These binary associations based only on co-occurrences between events are not sufficient to account for semantic learning and processes (see Perfetti, 1998). Similarity relations are also explained by the model, based on relationships between events which do not directly co-occur (e. g., ‘peach’ and ‘cherimoya’) but which co-occur with a common other contextual event (e. g., ‘good taste’). These indirect co-occurrences lead to associations between events on a similarity basis due to a common contextual event surrounding the perception of the associated events. Furthermore, a contextual event leading to similarity relations between two non co-occurrent events (e. g., ‘peach’ and ‘cherimoya’) can be activated through the perception (‘good taste’) or through the internal activation of the non co-occurring event as an associate (e. g., ‘grow in trees’) to the perceived events (‘peach’ and ‘cherimoya’). The general property to store events in association with a surrounding contextual event allows the network to represent not only binary association but also semantic similarity relations not directly dependent on the encountered co-occurrences in the environment.

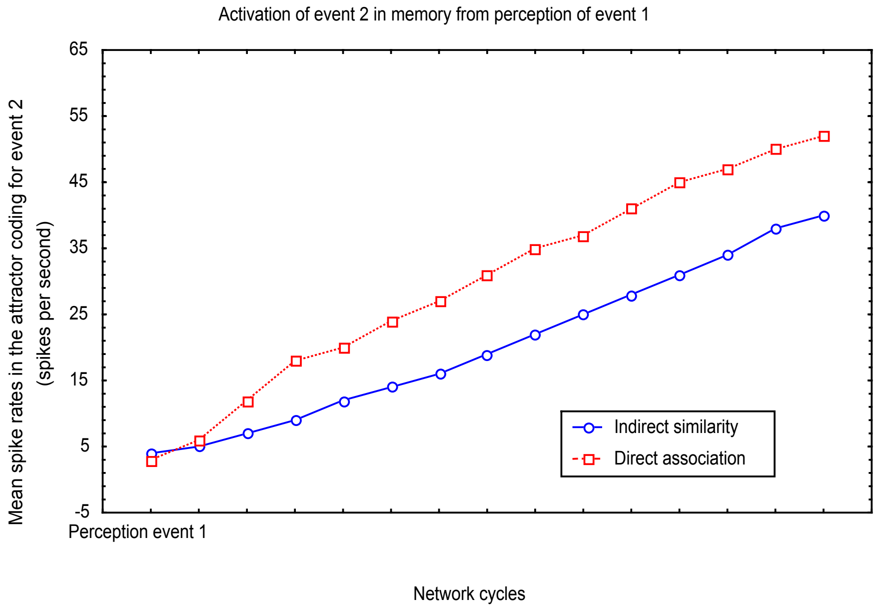

Figure 1: Activatory anticipations of event 2 from perception of event 1, as a function of associative learning between two events from direct co-occurrences (events 1 and 2 are presented in a sequence); and of similarity learning from indirect co-occurrences (events 1 and 2 are each co-occurrent with a third event 3).

(ii) Attentional factors can modulate associative learning based on direct binary co-occurrences as well as on context-based similarity. Indeed, these two modes of learning can be depicted as dependent, directly or indirectly, only on co-occurrences encountered in the environment (Landauer, Foltz & Laham, 1998; Foltz, Landauer & Dumais, 1997; Perfetti, 1998). However, to run adapted anticipations, a cognitive system must be able not only to store sequences of events occurring in its environment, but also to store them as a function of the event's relevance, that is on the basis of the attention given to the encountered events. Indeed, the acquisition of a new knowledge through associative learning processes can benefit from cognitive behavioral features such as attentional processing. For example, two co-occurring events (e. g., 'cherimoya' and 'good flavor') can be learned differentially as a function of the attention given to one or to both events. A simple hypothesis would be that (supra-threshold, possibly conscious) attention given to an event in memory is defined in the network as greater activation of the corresponding attractor's neurons compared to (infra-threshold, possibly unconscious) semantic anticipations. The more an event is learned (frequency of occurrence and perception time), the more its corresponding attractor would be, attentionally, activated in memory during a further perception. Then, the more attention is given to an event, the more its attractor can be activated (in intensity and time duration), and the stronger it can be associated to a co-occurring event through associative hebbian learning. Then, from the perception of one co-occurrence, attentional learning can modulate associations in memory from nearly zero to a maximum, which is a function of the intensity (equation 4) and time duration (equations 3.1. and 3.2.) of the activated attractors. Internal states of the network are as important as sequences of events perceived in the environment to determine the types and degrees of binary associative and similarity learning.

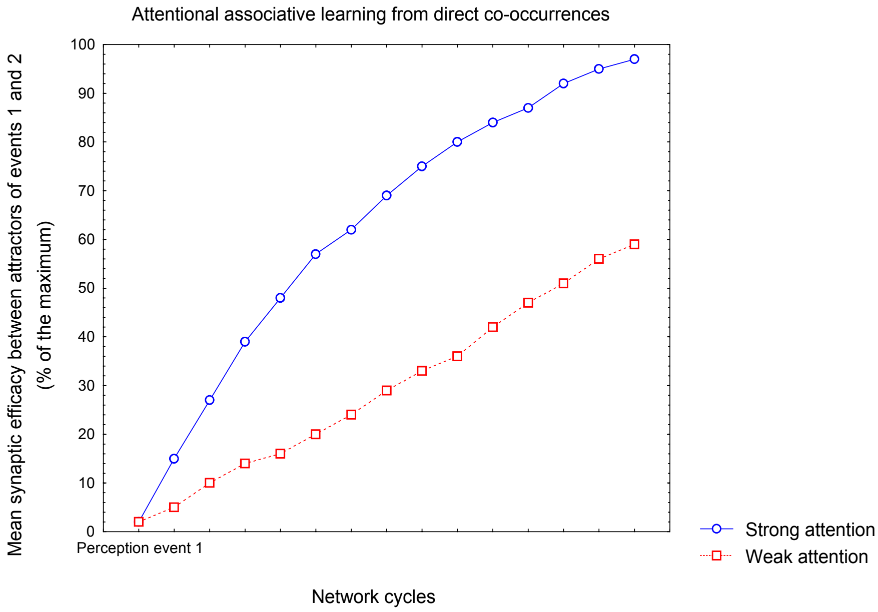

Figure 2: Activatory anticipations of event 2 from perception of event 1, as a function of the amount of attention given to event 1 during direct associative learning (values of $\tau_{(t)}Ii_{(\mu)}$ in equation 2).

4.2. Attentional and semantic anticipations

A common basis is given by the neural network model for semantic and attentional anticipations, in terms of degree of activation of attractors as a function of external and internal semantic and attentional properties of the perceived events. Semantic and attentional learning are based on common neural dynamics, modifying synaptic weights between event's attractors semantically associated in long term memory as a function of their attentional level of activation. Memorized knowledge can be differentially activated by perceived events and internal states of the system. The reverberating activations of neurons associated in attractors of delayed activity then correspond to knowledge activated in short term memory (see Amit et al., 1994). The attractor activated by the perception of the corresponding event activates in short term memory associated attractors corresponding to anticipated events not yet perceived in the environment (see Lavigne & Lavigne, 2000). Particular spike rate transmission through synaptic associations and slow network dynamics allows the model to vary attentionally the amount of (infra- or supra-threshold) activation of (the attractor's neurons of) semantic anticipations, as a function of the external and internal relevance of the perceived events. This type of model offers a unique opportunity to account for both anticipations and attention in unified terms of neural dynamics, associative semantic being coded in the synaptic weights between neurons and attention being represented as the level of activation of the event's attractors.

(iii) The model accounts for the rapid anticipations (2-3 network cycles) by automatic spreading of activation from a perceived event to an associated one (see Anderson, 1983; Balota, 1983; Greenwald, 1996; Keefe & Neely, 1990; Neely, 1991; Neely & Keefe, 1989; Neely, Keefe & Ross, 1989; Neely, 1976, 1977; Collins & Loftus, 1975; Collins & Quillian, 1969; Thompson-Schill, Kurtz, & Gabrieli, 1998;

VanVoorhis, & Dark, 1995). Furthermore, anticipations are sustained longer through time when more attention is given to the perceived event (Fuentes, Carmona, Agis, & Catena, 1994; Fuentes, & Ortells, 1993; Fuentes, & Tudela, 1992; Neely, 1991).

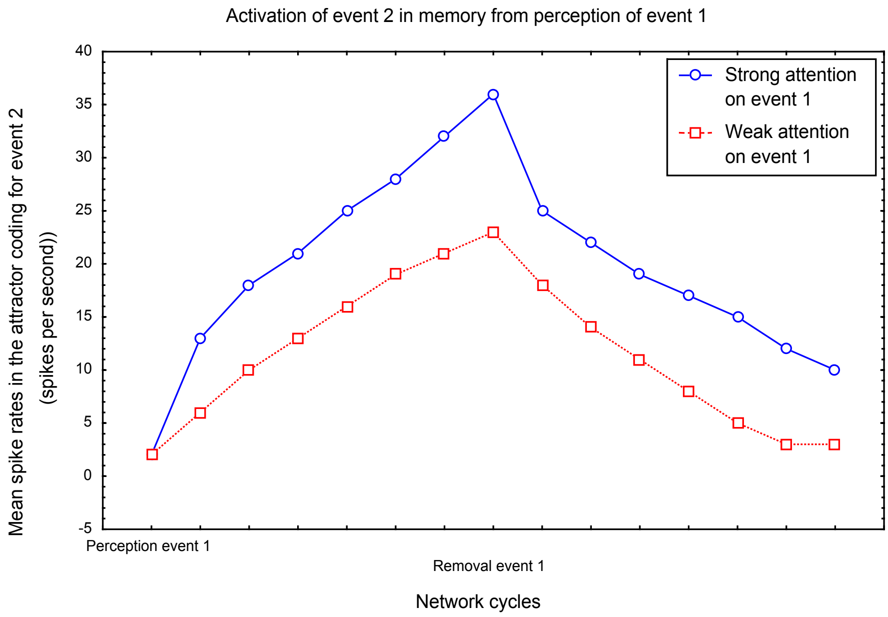

Figure 3: Activatory anticipations of event 2 from perception of event 1, as a function of the amount of attention given to event 1 during perception (values of $\tau_{(t)}Ii_{(\mu)}$ in equation 2).

(iv) The model explains how two perceived events triggering the same anticipations activate more an associated event in memory than a single perceived event (Balota & Paul, 1996; Brodeur & Lupker, 1994; Lavigne & Vitu, 1997), by increasing the amount of activation of the attractor coding for the anticipated event and received from perceived events in an additive way.

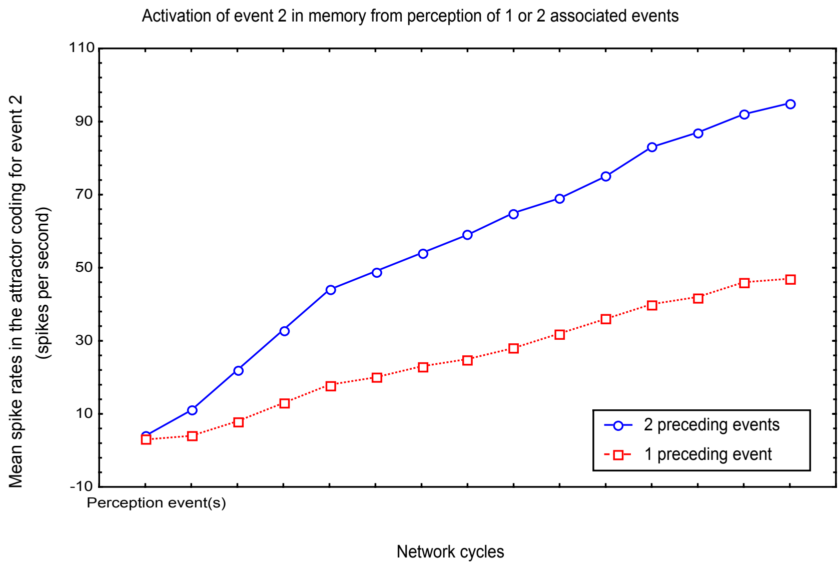

Figure 4: Activatory anticipations of event 2 from perception of one or two associated events (additive activatory effects on anticipations).

(v) When two events are perceived in a sequence, the model rapidly activates anticipations in parallel in memory, which are associated to the each events (rapid activation of anticipations and resistance to local incoherence between perceived events). The model also slowly inhibits anticipations associated to only one event (change of anticipations when perceived events are not coherent together: Lavigne & vitu, 1997; Neely, 1991; see Glenberg, 1997; see Berthoz, 1996; Dubois, 1996, 1998b).

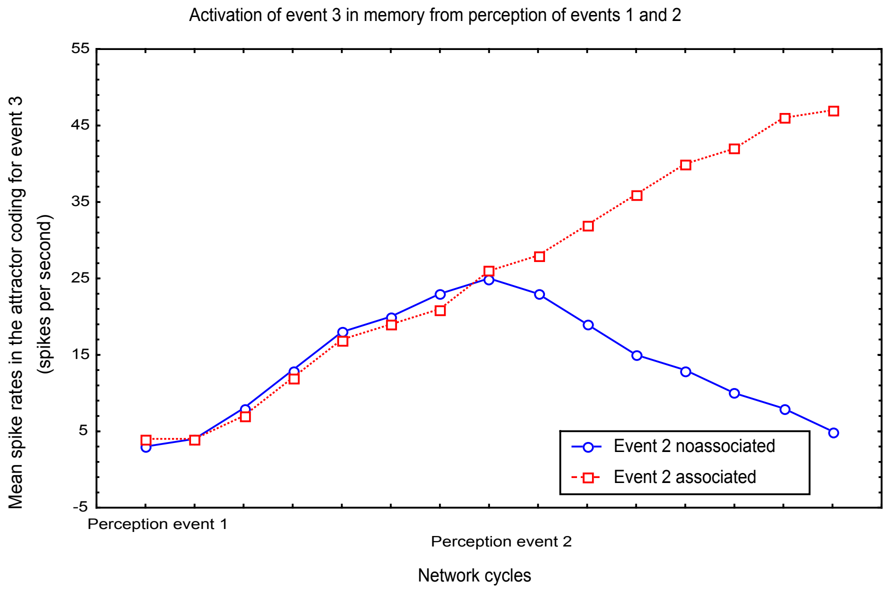

Figure 5: Fast activation and slow inhibition of anticipation of event 3 from perception of events 1 and 2 (inhibition of anticipations).

(vi) The model account for the effect of perception duration of an event on the strength of the anticipations driven in memory (Greenwald et al., 1996; Lorch, 1982; McNamara, 1994; Ratcliff & McKoon, 1981). The longer an event is perceived the more it activates an anticipated associated event in memory.

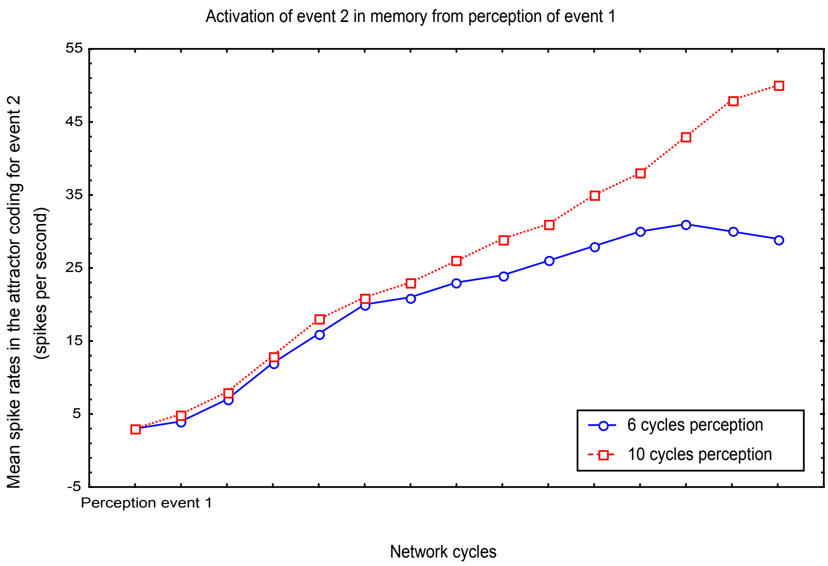

Figure 6: Activatory anticipations of event 2 from perception of event 1 as a function of the perception duration of event 1.

(vii) In the model, the perceptibility of an event corresponds to the number of neurons coding the event which are actually activated by the presentation of the event. The more neurons coding the event are activated during perception (e. g. all the neurons) compared to a noisy stimulus (e. g. part of the neurons and a background noise to the whole network), the more an anticipated associated event is activated in the network (see Holender, 1986; Fuentes et al., 1992, 1993, 1994; Lavigne & Dubois, 2000; Lavigne, Vitu, & d'Ydewalle, 2000).

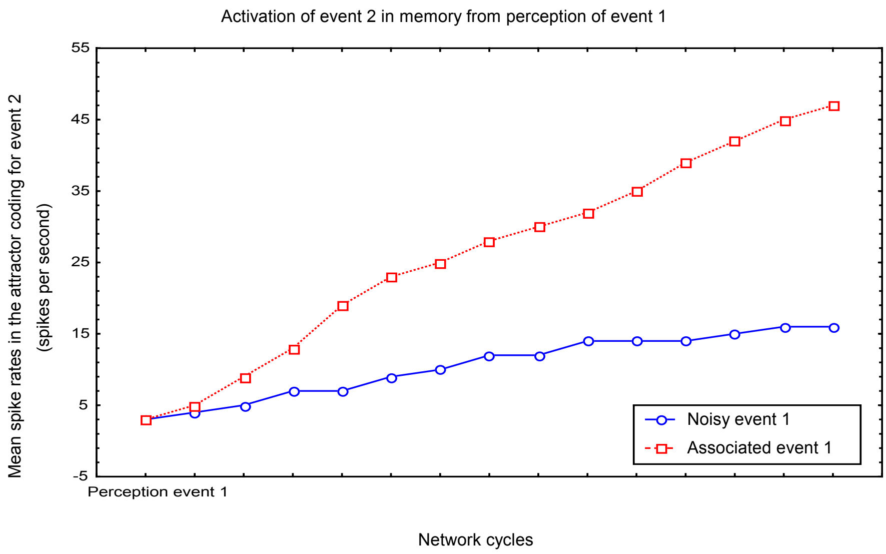

Figure 7: Activatory anticipations of event 2 from perception of event 1, as a function of the perceptibility of event 1 (learned event 1 *vs.* noisy event 1).

(viii) In the model low frequency events need more time to reach neuron's activation thresholds and activate associated anticipations in the network (Rayner & Balota, 1989; Vitu, 1991; Lavigne, Vitu & d'Ydewalle, 2000; see Monsell, 1991 for a review). When activated longer, the attractor of the perceived event activates longer associated anticipations, which reach higher activation levels. This accounts for the fact that more attention given to less frequently encountered events (Tipper, Bourque, Anderson & Brehaut, 1989).

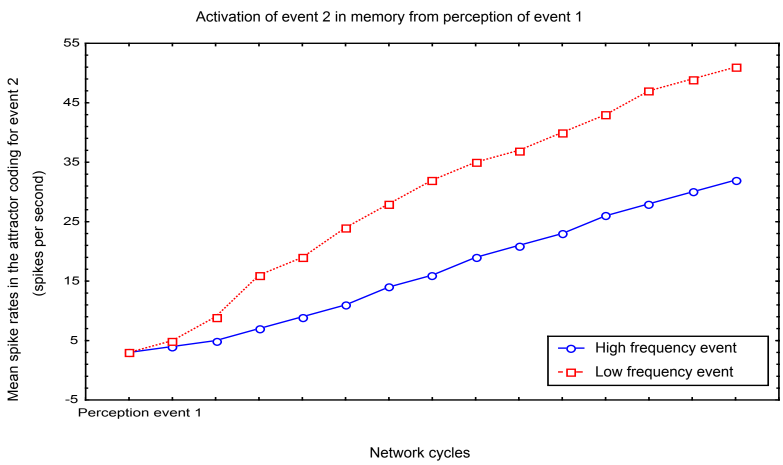

Figure 8: Activatory anticipations of event 2 from perception of event 1, as a function of the frequency of occurrence of event 1 (in numbers of cycles of presentation of event 1 during learning).

(ix) In the model the frequency of co-occurrence of two events during learning leads to stronger associations between their corresponding attractor neurons (see Conrad, 1972; Freedman & Loftus, 1971; Foltz, Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998; Perlmutter et al., 1976; Spence & Kimberly 1990). The more two events are associated the more a perceived one can trigger strong anticipation of the other in memory (see Becker, 1979; Lorch, 1982), which is then given more attentional relevance (see Jones & Yee, 1993).

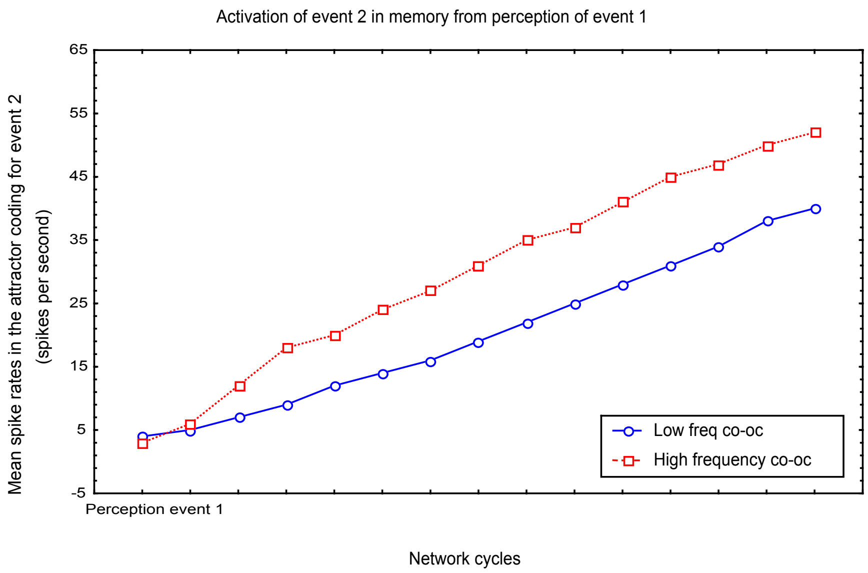

Figure 9: Activatory anticipations of event 2 from perception of event 1, as a function of the frequency of co-occurrence of events 1 and 2 (in numbers of cycles of presentation of events 1 and 2 in a sequence during learning).

(viii) In the network the more a perceived has associates, the more activation can add on an anticipated event through all the associated (i. e., the semantic field), given that many associates transmit their activation to a given associate in the field (Lavigne et al., in preparation). Perceived events for which one has the more knowledge (i. e., which have large semantic fields) are more relevant for attentional processes and lead to stronger anticipations.

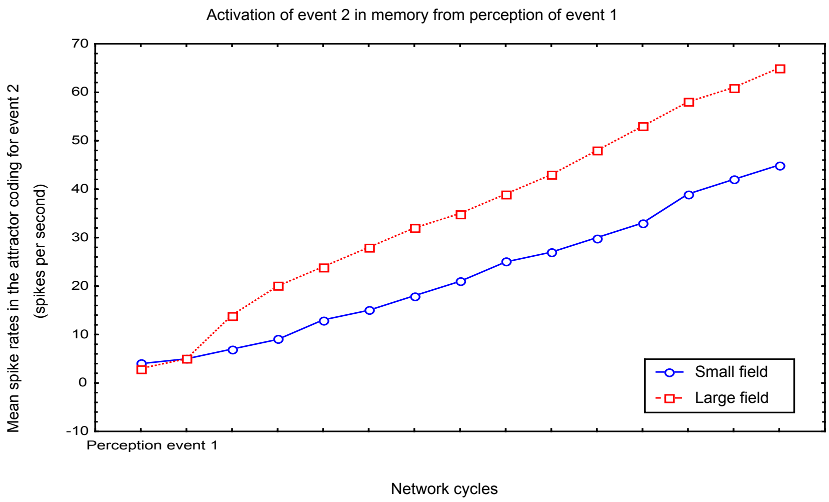

Figure 10: Activatory anticipations of event 2 from perception of event 1, as a function of the number of events associated to events 1 and 2 (semantic field = 0 vs. 3).

5. Conclusion

Attentional processing of events occurring in the environment is a fundamental cognitive ability to drive semantic anticipations (see Lavigne & Lavigne, 2000). Attentional drive of anticipations appear to be an important process in memory to finely adapt behavior to complex sequences of perceived events. As a function of both events external properties and learned semantic internal properties, attentional processing allows to evaluate events relevance in order to orient anticipations toward behaviors adapted to the most relevant anticipated events.

The ability to drive anticipations, through attentional processes, as a function of learned semantic knowledge about events in the environment, guaranty the adaptation of behaviors adopted by the cognitive system. This fundamental cognitive ability can be handled by anticipatory attractor neural networks, which allow to understand the interactions between semantic and attentional anticipations on the basis of a common neural structure. To deal with attentional drive of semantic anticipations, further developments of the model will need tuning of the neuronal parameters to allow the network to learn more events and to be more powerful in dealing with the processes reported altogether.

Furthermore, semantic anticipations are central cognitive processes which interact with fundamental cognitive abilities such as attention (Laberge, 1995), emotion (Damasio, 1998) and goal direction (Levine, Leven & Prueitt, 1992; Thagard, 1998). In addition to the attentional properties presented in the model, a great challenge to anticipatory neural networks is to code emotions and goals that can drive anticipations (Lavigne, & al. In preparation). This would lead to a better understanding of the learning and processing of emotions and goals by a cognitive system which adaptively anticipates in its environment.

Acknowledgments

We express special thanks to Nicolas Brunel for stimulating discussions in the development of the model. We also thank Dirk Steiner for his reading and comments on a previous version of the article.

References

Amit, D. J. (1989). Modeling Brain Function: The world of attractor neural networks. Cambridge Univ. Press.

Amit, D. J. & Brunel, N. (1995). Learning internal representations in an attractor neural network, Network,6,359.

Amit, D. J., Brunel, N., & Tsodyks, M. V. (1994). Correlations of hebbian cortical reverberations: Experiment vs. theory. Journal of Neurosciences, 14, 6635.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Balota, D. A. (1983). Automatic semantic activation and episodic memory encoding. Journal of Verbal Learning and Verbal Behavior, 22, 88-104.

Balota, D. A., Black, S. R. & Cheney, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two-process model. Journal of Experimental Psychology: Human Perception and Performance, 18, 485-502.

Balota, D. A. & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. Journal of Experimental Psychology: Learning, Memory and Cognition, 22, 4, 827-845.

Balota, D. & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: the range of influence of lexical variables. In D. Besner & G. Humphreys (Eds.), Basic processes in reading: Visual word recognition (pp. 198-232). Hillsdale, Hove, London: Lawrence Erlbaum.

Becker, C. A. (1979). Semantic context and word frequency effect in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 5, 252-259.

Berthoz, A. (1996). Les sens du mouvement. Paris, Odile Jacob.

Broadbent, D. E. (1971). Decision and Stress, New York, Academic Press.

Broadbent, D. E. (1982). Task combination and selective intake of information, Acta Psychologica, 50, 253-290.

Brodeur, D. A. & Lupker, S. J. (1994). Investigating the effects of multiple primes: An analysis of theoretical mechanisms. Psychological Research, 57, 1-14.

Brunel, N. (1996). Hebbian learning of context in recurrent neural networks. Neural Computation, 8, 1677-1710.

Brunel, N. (1994). Dynamics of an attractor neural network converting temporal into spatial correlations. Network, 5, 449.

Collins, A. M. & Loftus, E. F. (1975). A Spreading Activation Theory of Semantic Processing. Psychological Review, 82, 407-28.

Collins, A. M. & Quillian, M.R. (1969). Retrieval Time from Semantic Memory. Journal of Verbal Learning and Verbal Behaviour, 8, 240-247.

Conrad, C. 1972. Cognitive economy in semantic memory. Journal of Experimental Psychology, 92, 149-154.

Dubois, D. M. (1996). Hyperincursive stack memory in chaotic automata. In A. C. Ehresman, G. L. Farre & J.-P. Vanbremeersh (Eds.): Proceedings of the ECHOS Symposium, 77-82. Amiens – France, Université de Picardie Jules vernes.

Dubois, D. M. (1998a). Introduction to computing anticipatory systems. International Journal of Computing Anticipatory systems, 2, 3-14. D. Dubois (Ed.), CHAOS (Publisher).

Dubois, D. M. (1998b). Computing anticipatory systems with incursion and hyperincursion. Computing Anticipatory Systems: CASYS – first international conference, D. Dubois (Ed.), American Institute of Physics (Publisher), AIP conference proceedings 437, 3-29.

Freedman, J. L. & Loftus, E. F. (1971). Retrieval of words from long term memory. Journal of Verbal Learning and Verbal Behavior, 10, 107-115.

Fuentes, L. J., Carmona, E., Agis, I. F. & Catena, A. (1994). The role of the anterior attention system in semantic processing of both foveal and parafoveal words. Journal of Cognitive Neuroscience, 6, 17-25.

Fuentes, L. J. & Ortells, J. J. (1993). Facilitation and interference effects in a Stroop-like task: Evidence in favor of semantic processing of parafoveally-presented stimuli. Acta Psychologica, 84, 213-229.

Fuentes, L. J. & Tudela, P. (1992). Semantic processing of foveally and parafoveally presented words in a lexical decision task. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45, 299-322.

Glenberg, A. M. (1997). What memory is for, Behavioral and Brain Sciences, 20 (1), 1-55.

Green, D. M. & Swets, J. A. (1966). Signal detection theory and psychophysics, New York, Wiley.

Greenwald, A. G., Draine, S. C. & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. Science, 273, 1699-1702.

Hollender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. Behavioral and Brain Sciences, 9, 1-66.

Jones, M. R. (1976). Time, our lost dimension: toxard a new theory of perception, attention and memory, Psychological Review, 83, 323-355.

Jones, M. R. & Boltz, M. (1989). Dynamic attending and response time, Psychological Review, 3, 459-491.

Jones, M. R. & Yee, W. (1993). Attending to auditory events: the role of temporal organization, in S. McAdams & E. Bignand (Eds.), Thinking in sound, Oxford, Clarendon Press.

Keefe, D. E. & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. Memory & Cognition, 18, 289-298.

Laberge, D. (1995). Attentional Processing, Cambridge, Mass., Harvard University Press.

Landauer, T. K., Foltz, P. W. & Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes, 25, 259-284.

Lavigne, F. & Dubois, D. (2000). Context effects and associative anaphora in reading, Journal of Pragmatics, 31, 399-415.

Lavigne, F. & Vitu, F. (1997). Time course of facilitatory and inhibitory semantic priming effects in visual word recognition. International Journal of Psycholinguistics, 13, 311-349.

Lavigne, F., Vitu, F & d'Ydewalle. (2000). Context effects and parafoveal preview in sentence reading, Acta Psychologica, 104, 291-214.

Lavigne, F., Vitu, F & d'Ydewalle. (submitted). Time-course of facilitatory and inhibitory context effects in sentence reading.

Lecas, J. C. (1992). L'attention Visuelle, Bruxelles, Mardaga.

Levine, D. S., Leuven, S. J. & Prueitt, P. S. (1992). Integration, disintegration and the frontal lobes. In Levine, D. S. & Leven, S. J. (Eds), Motivation, emotion and goal direction in neural networks (pp 301-334), Lawrence Erlbaum Associates, Hillsdale, N.J.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification, In D. Besner and G. Humphreys (Eds.), Basic processes in reading: Visual word recognition (pp. 233-263) Hillsdale, Hove, London, Lawrence Erlbaum.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. Journal of Experimental Psychology: Learning, Memory & Cognition, 21, 3-23.

McNamara, T. P. (1994). Theories of Priming: II. Types of primes. Journal of Experimental Psychology: Learning, Memory & Cognition, 3, 507-520.

Millis, M. L. and Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. Memory and Cognition, 17, 141-147.

Monsell, S. (1991). The nature and locus of word frequency effect in reading, In D. Besner and G. Humphreys (Eds.), Basic processes in reading: Visual word recognition (pp. 148-197). Hillsdale, Hove, London: Lawrence Erlbaum.

Neely, J.H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. Memory & Cognition, 4, 648-654.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. Journal of Experimental Psychology: General, 106, 226-254.

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of currents findings and theories. In D. Besner & G. Humphreys (Eds.), Basic processes in reading: Visual word recognition (pp. 264-336). Hillsdale, Hove, London: Lawrence Erlbaum.

Neely, J.H. & Keefe, D.E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower, (Ed.), The psychology of learning and motivation: Advances in research and theory, vol 24 (pp. 207-248). NewYork: Academic Press.

Neely, J.H., Keefe, D.E. & Ross, K. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 1003-1019.

Perlmutter, J., Sorce, P. & Myers, J. L. (1976). Retrieval processes in recall. Cognitive Psychology, 8, 32-63.

Posner, M. I. (1980). Orienting of attention, Quarterly Journal of Experimental Psychology, 32, 3-25.

Posner, M. I. & Cohen, Y. (1984). Components of visual orienting, in H. Bouma & D. Bouhuis (Eds.), Attention and Performance X, 531-556, London, Erlbaum.

Posner, M. I. & Snyder, C. R. R. (1975a). Facilitation and Inhibition in the processing of signals. In P. M. A. Rabitt & S. Dornic (Eds.), *Attention and Performance V* (pp. 669-698). New York: Academic Press.

Posner, M. I. and Snyder, C. R. R. (1975b). Attention and cognitive control. In R. L. Solso (Ed.), Information processing and cognition: The Loyola Symposium (pp. 55-85), Hillsdale, NJ: Erlbaum.

Rayner, K. &. Balota, D.A. (1989). Parafoveal preview and lexical access during eye fixations in reading. In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 261-290). Hillsdale, Hove, London: Lawrence Erlbaum.

Rosen, R. (1985). Anticipatory Systems – Philosophical, Mathematical and Methodological Foundations. Pergamon Press.

Rumelhart, D. E. & McClelland, J. L. (1981). An interactive activation model of context effects in letter perception : I. An account and basic findings, Psychological review, 88, 375-407.

Rumelhart, D. E. & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception : II. The contextual enhencement effect and some tests and extensions of the model. Psychological review, 89, 60-94.

Shiffrin, R. M. (1988). Attention, in R. L. Atkinson, R. J. Merrenstein, G. Lindsey & R. D. Luce (Eds.), Steven's Handbook of Experimental Psychology, Second edition (vol II) : Learning and Cognition, New York, Wiley & Son, 739-811.

Spence, D. P. & Kimberly, C. O. (1990). Lexical co-occurrence and association strength, Journal of Psycholinguistic Research, 19, 317-330.

Sperling, G. & Reeves, A. (1980). Measuring the reaction time of an unobserved response; a shift of visual attention, in R. Nickerson (Ed.), Attention and Performance VIII, Hillsdale, Lawrence Erlbaum.

Tanner, W. P. & Swets, J. A. (1954). A decision-making theory of visual detection, Psychological Review, 61,

Thompson-Schill, S., Kurtz, K.J. & Gabrieli, J.D.E. (1998). Effects of semantic and associative relatedness on automatic priming, Journal of Memory and Language, 38, 440-458.

Tipper, S.P., Bourque, T.A., Anderson, S.H. & Brehaut, J. (1989). Mechanisms of attention: a developmental study, Journal of Experimental Child Psychology, 48, 353-378.

VanVoorhis, B.A., & Dark, V. (1995). Semantic matching, response mode, and response mapping as contributors to retroactive and proactive priming, Journal of Experimental Psychology: Learning, Memory and Cognition, 21, 4, 913-932.

Vitu, F. (1991). The influence of parafoveal preprocessing and linguistic context on the optimal landing position effect. Perception & Psychophysics, 50, 58-75.

Weichselgartner, E. & Sperling, G. (1987). Dynamic of automatic and controlled visual attention, Science, 238, 778-780.