Semantic integration by pattern priming: experiment and cortical network model

Frédéric Lavigne¹ · Dominique Longrée² · Damon Mayaffre³ · Sylvie Mellet³

Received: 13 April 2016 / Revised: 18 July 2016 / Accepted: 6 September 2016 / Published online: 17 September 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Neural network models describe semantic priming effects by way of mechanisms of activation of neurons coding for words that rely strongly on synaptic efficacies between pairs of neurons. Biologically inspired Hebbian learning defines efficacy values as a function of the activity of pre- and post-synaptic neurons only. It generates only pair associations between words in the semantic network. However, the statistical analysis of large text databases points to the frequent occurrence not only of pairs of words (e.g., “the way”) but also of patterns of more than two words (e.g., “by the way”). The learning of these frequent patterns of words is not reducible to associations between pairs of words but must take into account the higher level of coding of three-word patterns. The processing and learning of pattern of words challenges classical Hebbian learning algorithms used in biologically inspired models of priming. The aim of the present study was to test the effects of patterns on the semantic processing of words and to investigate how an inter-synaptic learning algorithm succeeds at reproducing the experimental data. The experiment manipulates the frequency of occurrence of patterns of three words in a multiple-paradigm protocol. Results show for the first time that target words benefit more priming when embedded in a pattern with the two primes than when only associated with each

prime in pairs. A biologically inspired inter-synaptic learning algorithm is tested that potentiates synapses as a function of the activation of more than two pre- and post-synaptic neurons. Simulations show that the network can learn patterns of three words to reproduce the experimental results.

Keywords Context · Inter-synaptic learning · Word occurrence · Multiple priming · Prospective activity · Word meaning

Introduction

Biologically inspired models of the cerebral cortex have been able to account for contextual recall on the basis of matrices of synaptic connectivity, coding word associations in long-term memory (Lavigne and Denis 2001, 2002; Mongillo et al. 2003; Brunel and Lavigne 2009; Lavigne and Darmon 2008). Recall is related to semantic activation between words and can be experimentally estimated by the magnitude of semantic priming effects, corresponding to shorter response times for processing target words when targets are associated with preceding prime words than when they are not (Meyer et al. 1972; Schvaneveldt and Meyer 1973; see Neely 1991; Hutchison 2003).

Semantic activation is described as developing in semantic networks formed of multiple associations between pairs of words (e.g., Brunel and Lavigne 2009; Lavigne et al. 2011; see Lerner and Shriki 2014). However, the statistical analysis of large text databases by linguists points to the chance occurrence not only of pairs of words but also of groups of more than two words. Such repeated sequences (Salem 1986), called motifs, function as integrated textual units (Longrée et al. 2008; Mellet and

✉ Frédéric Lavigne
lavigne@unice.fr

¹ BCL, UMR 7320 CNRS et Université de Nice-Sophia Antipolis, Campus Saint Jean d’Angely - SJA3/MSHS Sud-Est/BCL, 24 Avenue des diables bleus, 06357 Nice Cedex 4, France

² Université de Liège, Liège, Belgique

³ Université Côte d’Azur, CNRS, BCL, Nice, France

Longrée 2012; Longrée and Mellet 2013). These motifs are frequent semantic patterns that correspond to regular sequences of words. The minimal pattern not reducible to pairs of words corresponds to a triplet of co-occurrent words such as, for example, “by the way”. Such a pattern is described by three pair associations (“by the”, “by way”, and “the way”), plus a higher-level representation of the triplet itself (“by the way”), relying on the idiomatic principle (Sinclair 1991). A consequence is that the processing of a pattern requires knowledge of the triplet *in addition to* knowledge of the pairs, whereas the processing of a triplet not in a pattern (e.g., “a”, “strange”, and “way”) would involve solely pair associations (e.g., “a strange”, “a way”, and “strange way”). The possibility for higher-level knowledge of patterns of words points to the questions of their online processing and of their learning in cortical network models.

The present study reports experimental results and computational modeling results. A multiple priming experiment shows that pattern priming effects involve higher-level representations of patterns of three words not reducible to activation between pairs. A computational model of learning and priming in the cerebral cortex shows that a biologically realistic inter-synaptic learning rule allows the network to reproduce the experimental results.

Pairs priming and Hebbian learning

Cortical network models assume that associations between words are learned according to biologically realistic mechanisms of Hebbian learning (Brunel 1996; Mongillo et al. 2003; Hebb 1949). In biologically inspired models of the cerebral cortex, the value of pair associations has been shown to be learned proportionally to the number of occurrences of the pairs of items according to a realistic Hebbian learning paradigm (Brunel et al. 1998; Mongillo et al. 2003). A main feature of the Hebbian learning rule is that it is local: it potentiates or depresses a synapse connecting two pre- and post-synaptic neurons solely on the basis of the activity of these two neurons. The probability to potentiate the synapse is proportional to the frequency of co-activation of the two neurons, that is, as a function of the number of co-occurrence of the two items (here, words or concepts) coded by the two neurons. Hence, Hebbian learning generates associations between pairs of neurons (and of populations of neurons), coding for items in memory (e.g., Amit et al. 1994; Brunel 1996; Brunel et al. 1998; Amit and Brunel 1997; Mongillo et al. 2003). The magnitude of priming effects is commonly reported by experimentalists as predicted by the strength of the prime-target association (Abernethy and Coney 1993; Frishkoff 2007; Hutchinson et al. 2003; Coney 2002; Hutchison

2003). This association strength between words is estimated in free production norms by the percentage of production of a given target word in association with a given prime word (Ferrand and Alario 1998; Cree and McRae 2003; McRae et al. 2005; see Nelson et al. 1999). In line with the Hebbian learning rule, associations between words are predicted by the occurrence of the pairs of words in texts, measured as the number of times two words are close together within a same sentence or paragraph (Spence and Owens 1990; Landauer et al. 1998; Van Petten 2014; Luka and Van Petten 2014).

Multiple priming and the limits of Hebbian learning

The hypothesis of pattern priming requires an experimental paradigm involving sequences of more than two words. Multiple priming experiments investigate how two prime words activate a subsequent target word (McNamara and Altarriba 1988; McNamara 1992; Brodeur and Lupker 1994; Balota and Paul 1996; Lavigne and Vitu 1997). Lavigne et al.’s (2011) meta-analysis of the data has revealed that the rich phenomenology of multiple priming effects depends on the association between the target word (e.g., “tiger”) and each of the two preceding prime words (e.g., related primes “lion” and “stripes” and unrelated primes “fuel” and “shutter”). A main result is that targets associated with the two primes are more strongly activated than targets associated with only one of the two primes, suggesting that the stronger activation of associates that are *common* to the two primes could subtend the semantic integration of the meaning of the sequence of words (see Beeman et al. 1994). Further, recent experiments have reported that the precise level of association strength between each prime and the target is decisive for the activation of the target (Lavigne et al. 2012, 2013); targets strongly associated with one *or* the other prime are activated, but targets weakly associated with the primes require the *two* primes to be associated with the target.

Multiple priming experiments provide an understanding of the way activation of a target is generated by a sequence of primes on the sole basis of the pair associations between each prime and the target. However, during discourse comprehension, words’ meanings are dynamically integrated in the context of other words (Kamide et al. 2003; Hagoort et al. 2004; Hald et al. 2007). For example, in Kamide et al. (2003), participants fixated on objects in visual scenes while hearing sentences. They fixated more on a motorcycle when they heard “will ride” in the sentence beginning with “The man will ride...”, whereas they fixated more on a carousel if the sentence began with “The girl will ride...”. This indicates that the type of semantic information processed (e.g., the pictures of “motorcycle” or “carousel”) depends on the context of at least two other

words in the sentence (e.g., “boy” and “ride” or “girl” and “ride”, respectively). These results suggest that a given target word can be activated not only by each of two primes but by their combination, which seems to be in line with the identification of patterns of words by linguists. Two multiple-priming studies have investigated the possibility that two primes could activate a target on the sole basis of their combination and in the absence of pair associations in triplets (Chwilla and Kolk 2005; Khalkhali et al. 2012). Both studies report that a pair of primes (e.g., “check-in” and “boarding”) can activate a target (e.g., “flight”; examples from Khalkhali et al. 2012), even though targets were not produced as associates to the primes nor shared category relation or overlap with the primes (Chwilla and Kolk 2005) and even though each prime, presented alone, did not activate the target (Khalkhali et al. 2012). These results suggest that priming could occur in the absence of pair association, so that only knowledge of the triplet would have generated the activation of the target not generated by each prime individually. These studies base their interpretation on the observation of priming effects generated by the combination of two primes when each prime, taken alone, would not activate the target. However, accepting an absence of pair association between each prime and the target is difficult to prove. The methodological issues raised by these studies are developed in the “Experiment” section.

The processing of sequences of words not reducible to their pair associations is a case of representational learning of higher order combinations of multiple (i.e. more than two) features (see Fiser 2009). Interestingly, in experiments testing learning of groups of visual stimuli, humans are more sensitive to regular combinations of triplets (combo or pattern) made of regular pairs, than to random triplets made of regular pairs (Fiser and Aslin 2005). This is an important assessment of learning of higher-order representations of an input–output relation given a specific context, not reducible to pair associations. Statistical learning of such context-dependent representations is addressed by models based on tensor products of vectors by associating an output to the Kronecker product of an input and a context (e.g., Mizraji and Lin 2015; Mizraji et al. 2009). An important issue for cognitive neuroscience is to account for context-dependent processing in terms of biologically realistic mechanisms at the neural level. Input–output associations have been extensively studied in biologically realistic models of the cerebral cortex using Hebbian learning of multiple pairwise associations. However, learning of higher order representations of contextually related multiple elements is a challenge to classical Hebbian algorithms (Rigotti et al. 2010a, b, 2013; Bourjaily and Miller 2011a, b, 2012; Lavigne et al. 2014). To summarize, three arguments make the experimental testing

of multiple priming a central issue for our understanding of learning and processing of combinations of words: the statistical measure of combinations of more than two words in text databases, the experimental assessment of human learning of combinations of visual features, and the challenge of modeling the learning of higher order representations in biologically realistic models of the cerebral cortex.

Rationale of the experiment and model

The present study investigates the dynamic processing (experiment and modeling) and synaptic learning (modeling) of semantic patterns of word triplets:

- Semantic patterns challenge our understanding of their online processing: Does the activation of a target depend on its combination with two primes in a pattern for given pair associations? The behavioral experiment manipulates for the first time the level of occurrence of the patterns to test for priming effects as a function of the frequency of occurrence of the patterns.
- If the response of the experiment to the first question is positive, then the question of learning of patterns challenges biologically inspired models based on Hebbian learning: how the target can get activation from the higher-level representation of the pattern of three words not coded as pair associations? The model proposed will test a way of learning semantic patterns and its effects on the dynamic processing of semantic patterns.

Experiment

The experiment uses a multiple priming protocol to test whether patterns of words identified statistically in texts generate stronger priming effects than triplets of words not occurring in patterns. Multiple priming studies have reported priming by two primes while not reporting priming by a single prime alone (Khalkhali et al. 2012) or not reporting pair associations between each prime and the target (Chwilla and Kolk 2005). If there was indeed no association between the primes and the target, these cases of “unrelated” priming would be incompatible with the fundamental process of propagation of activation involved in semantic priming (see Khalkhali et al. 2012). However, we will see that the apparent absence of pair association and pair priming can be accounted for by a classical and simple mechanism reported at the neuronal level and used to describe activation in network models of priming (e.g., Anderson 1976, 1983a, b; Lavigne and Denis 2002; Mongillo et al. 2003; Brunel and Lavigne 2009; Lerner et al.

2012; Lerner and Shriki 2014; see McRae and Ross 2004; Randall et al. 2004; Cree et al. 1999; Becker et al. 1997; Moss et al. 1994; Masson et al. 1991; Masson 1995; Plaut 1995; Plaut and Booth 2000). Indeed, in computational models, activation obeys a nonlinear, current-to-frequency transfer function inspired by the one observed in cortical neurons (Ricciardi 1977; Tuckwell 1988; Ermentrout and Kopell 1986). The mathematical description of the transfer function of cortical pyramidal neurons was obtained analytically by Brunel and Latham (2003) for quadratic integrate-and-fire neurons in the presence of background noise (Fig. 1). This sigmoid-like current-to-rate transfer function (f - I curve) approximates the response of neurons coding for items in memory in response to a given input.

The nonlinear transfer function predicts that a single source of input activation can lead to a subthreshold output activation of the target neuron close to zero, whereas two sources of inputs of the same value can lead to a supra-threshold activation (Fig. 1A). This could account for the fact that very weakly associated words would generate neither association in free production norms nor priming effects in real-time priming experiments with a single prime. As a consequence, the observation of priming by two primes but not by one does not rule out the simple integration mechanism of neuronal inputs, according to which two inputs generating a subthreshold output, when presented alone, can generate supra-threshold output when presented together. The thresholding of the integration of inputs is a mechanism sufficient to account for such over-additive multiple priming effects. Network simulations predicted that when one prime is weakly associated with the target and the other is not associated, priming is almost null (10 ms priming effect; Fig. 5 B2 and C2 in Lavigne et al. 2011), whereas when the two primes are weakly associated with the target, priming is magnified (50 ms; Fig. 5 A2; Lavigne et al. 2011).

It has to be noted that, according to the I - f curve of Fig. 1, strictly additive multiple priming effects are possible, with the priming effect generated by the two primes equal to the sum of priming effects generated by each prime (see Balota and Paul 1996). Strictly additive effects (and even under-additive effects) are possible depending on the precise slope of the f - I curve so that two primes each generating an input current I_i (output frequency F_i) double, when presented together, the total input current that leads to an doubling in the output frequency ($2F_i$). This would correspond to a multiple priming effect equal to the sum of the effect generated by each prime in isolation, hence corresponding to additive effects (see Lavigne et al. 2011 for a meta-analysis of the pattern of additivity and model).

Turning to over-additive priming effects, recent experiments have shown that a target can be activated by two primes that are weakly associated to it, whereas it is not

activated by one prime weakly associated and another prime not associated (Lavigne et al. 2012, 2013). However, in these studies, the absence of priming when one prime is related could rely on interference effects generated by the other, unrelated prime (see Lavigne et al. 2011 for a model and discussion). In Khalkhali et al.'s (2012) study, priming was measured with isolated primes and compared to priming with two primes. This approach should allow evidence of the combined effects of two primes compared to one. In Khalkhali et al.'s (2012) study, the average pair association between the primes and the target was not null (.045 % on average, ranging from 0 to .08 %) but was

Fig. 1 Population f–I curve describing how the average firing rate of a population of excitatory neurons depends on the average synaptic input it receives. The *blue dashed line* (– –) corresponds to the threshold above which the neurons' spike rate generates a significant priming effect (compared to an absence of input). **A** Case of Hebbian pair learning only. The *red line* (—) corresponds to the sub-threshold output frequency F_i generated by a single input I_i . The *green line* (—) corresponds to the supra-threshold output frequency nF_i generated by two inputs $2I_i$. The current-to-rate transfer function (f–I curve) predicts that a single source of input activation of value I_i (i.e., a single input current) can lead to a subthreshold output activation F_i (i.e., output frequency) of the target neuron close to zero, whereas two sources of inputs of same value I_i can lead to a supra-threshold activation nF_i ($n > 2$) of the target neuron. **B** Case of inter-synaptic learning. The *red line* (—) corresponds to the subthreshold output frequency F_i generated by a single input I_i . The *green line* (—) corresponds to the subthreshold output frequency nF_i generated by two inputs $2I_i$ when not learned in a pattern. The *orange line* (—) corresponds to the supra-threshold output frequency $nF_i + n'F_i$ generated by two inputs *plus* their amplification $2I_i + xI_i$ when learned in a pattern. Here, inter-synaptic learning predicts that two sources of input activation of same value I_i can lead to a subthreshold output activation nF_i of the target neuron, whereas two sources of inputs of same value I_i *learned in a semantic pattern* can lead to a supra-threshold activation $nF_i + n'F_i$ of the target neuron. (Color figure online)

interpreted as an absence of relation. However, other studies have reported significant effects of a single prime associated with the target with strength ranging from 1 to 5 % (Hutchinson et al. 2003) and as low as 1 % in average (SD = 0.7) (Coney 2002). Therefore, it can't be ruled out that the non-significant priming by single primes reported by Khalkhali et al.'s (2012) corresponds to an absence of activation of the target. A small amount of activation of the target triggered by each prime could have generated non-significant activation by each prime alone and the significant multiple priming when the two primes were presented. This makes very difficult to assess that the absence of priming by a single prime is due to an actual absence of association. A nonsignificant priming effect at the behavioral level is no proof of an absence of activation at the network level. In Chwilla and Kolk's (2005) experiment, associations measured from production norms were not weak but null. However, given that only one or some associates with the prime are asked to participants, all associates in memory are not recorded. As a consequence, the absence of recording of a target as an associate with a prime does not prove that the two words are not associated in memory. This does not rule out the null hypothesis of very weak associations generating weak activation insufficient for the target word to be produced as an associate with the prime. In that case, the association would not be detectable in free-production norms (e.g., McRae et al. 1997; Cree and McRae 2003; see Nelson et al. 1999). Further, there is a possibility for indirect associations

between the primes and the target through common associates not tested in the norms. Such indirect associations have been reported in several studies to generate indirect priming of smaller magnitude than direct priming (Kiefer et al. 2005; Hill et al. 2002; Chwilla et al. 2000; McNamara 1992; see Weisbrod et al. 1998) or even at nonsignificant levels (Arnott et al. 2003; Hill et al. 2002; Bennett and McEvoy 1999; Spitzer et al. 1993; Kiefer et al. 2005; Moritz et al. 2003). To summarize, very weak or indirect associations could generate supra-threshold priming effects when two primes are presented, even though the association is too weak to generate priming from a single prime (see Lavigne et al. 2011 for a review and model).

Calling for a new mechanism that combines two primes and target would first require experimental results not already handled by the classical neuronal integration mechanisms. The present experiment aims at testing for the existence of pattern priming effects generated by the combination of specific primes embedded in the same pattern as the target, in addition to classical priming effect generated by pair associations. To circumvent the methodological limit of assessing for the null hypothesis of an absence of association between the primes and target, we propose here to test for the effect of a higher-level representation of patterns of three words on the activation of a target in the case of non-null associations between each prime and the target, but for two different levels of occurrence of the patterns. The strategy here is to test whether two primes embedded in a frequent pattern with the target lead to stronger priming than two primes embedded in a rare pattern, even though each prime can occur with the target (Fig. 1B). Reporting such pattern priming effects does not require accepting the null hypothesis of the absence of pair associations. Pattern priming effects would point to an additional mechanism of amplification of the activation of the target when a pattern is processed. The hypothesis of learning and coding patterns of three words in memory predicts that even though the target is not activated by a pair of primes that rarely occur with the target, it should be activated by a pair of primes that occur frequently in a pattern with the target. Regarding the notion of frequency of occurrence of pairs and patterns, Fiser and Aslin's (2001) frequency-balanced test has shown that humans learn pairs of visual elements that always occur together better than pairs of elements that could also occur without each other. This important result suggests that it is not the frequency of occurrence of the pair that is learned, but the frequency of occurrence of the pair as a function of the overall frequency of each item alone. As a consequence, the level of learning of a pair AB *of a given frequency* would be better if A and B do not occur with other items (always AB) than if they do (AB, AC, BD, etc.). The AB pair is then learned not as a function

of its *absolute frequency* but as a function of its *relative frequency* given the absolute frequencies of A and of B. This phenomenon is in accordance with Hebbian learning at the synaptic level, which specifies that the synapse potentiates as a function of the frequency of co-activity of the two neurons (e.g., coding for A and B) and depresses as a function of the activity of only one of the two neurons (i.e., when this neuron is co-active with other neurons: AC, BD, etc.). The interplay of synaptic potentiation and synaptic depression depends on the probability of co-activity of two neurons (e.g., activating B knowing A, with A potentially activating other neurons than B) and not only on the frequency of their co-activity (activating B by A whatever the other neurons activated by A). Turning to patterns of three items, Fiser and Aslin's (2001) results strongly suggest that the relevant information is not the absolute frequency of the pattern but its relative frequency given the absolute frequencies of the pairs in the pattern. The ABC pattern is then learned not as a function of its *absolute frequency* but as a function of its *relative frequency* given the absolute frequencies of AB, BC and AC. However, the corresponding pattern priming effects would challenge models based on classical Hebbian learning. The hypothesis of pattern priming was tested by cross-manipulating the relative frequency of occurrence of the prime-target pairs and the relative frequency of occurrence of patterns of three words.

Method of the experiment

Material

The experimental material was made of triplets of two prime words and a target word selected for which the frequency of occurrence of the single words, pairs and patterns was measured in order to calculate the relative frequencies. To maximize the relationship between the occurrence of the pairs and patterns in text databases and the expected priming effects generated in memory, frequencies were calculated on a standard database of Latin texts all read and known by the participants. The Latin texts have the advantage of constituting a well identified corpus of texts well-studied and learned by participants. The text database comprises clearly identified patterns of words organized in ordered sequences expected to be associated in memory (e.g., Longrée and Mellet in press). In this way relative frequencies of pairs and triplets of words that were calculated in the database corresponded to the ones encountered by participants. Pairs and pattern frequencies were calculated with the Hyperbase software over the 1.709.956 words of the database. All triplets were

selected on the basis of the relative frequency of the pattern of three words (rare vs. frequent “prime 1-prime 2-target” pattern). For each level of pattern frequency, a frequent and a rare level of pairs relative frequencies were calculated as the sum of each prime-target pair frequency (prime 1-target + prime 2-target; see Table 1) to estimate the existence of multiple priming effects in the material used (see Lavigne et al. 2011, for a review).

The two modalities of pattern relative frequency (rare vs. frequent) and of pairs relative frequency (rare vs. frequent) defined four experimental conditions: Latin examples (and their English translation) correspond to “prime 1 prime 2 target”:

- Rare pattern, rare pairs: “quid nunc dictum” (“what has now been”)
- Frequent pattern, rare pairs: “ab utroque latere” (“from both sides”)
- Rare pattern, frequent pairs: “his paratis rebus” (“these preparations”)
- Frequent pattern, frequent pairs: “nec multo post” (“soon after”)

Two experimental lists were created that involved 144 triplets. In half the triplets in a list, 72 targets were bona fide ancient Latin words. In the other half of the list, 72 pseudo-word targets were derived from a Latin word by replacing one or two letters but constructed in accordance with the phonological constraints of ancient Latin. The 72 word triplets included 36 frequent patterns and 36 rare patterns, each group with targets involved in frequent pairs and rare pairs. A target in a frequent pattern in a list was in a rare pattern in the other list by changing the two primes that preceded the target to create a possible but rare pattern. Each word therefore appeared only once in a list presented to a subject. Pattern priming effects were measured by changing primes for a same-word target so that targets were the same in the two conditions of pattern frequency (Table 1). The relative frequency of the pairs and pattern corresponds to the frequency of the pairs (or pattern) divided by the absolute frequency of the words (or pairs) alone. Given that the absolute frequencies of words alone are larger than absolute frequencies of pairs, the relative frequency of pairs is smaller than the one of patterns. ANOVAs were used to check that the manipulation of the relative frequencies of pairs and of patterns did not interact, that is that pattern frequency was not significantly larger for frequent pairs than for rare pairs. Pattern relative frequency did not significantly differ between frequent and rare pairs ($F[1, 140] = 2.91, ns$), and did not vary with pairs frequency neither for rare patterns ($F[1, 140] = 2.20, ns$) nor for frequent patterns ($F < 1, ns$). Pairs relative frequency did not significantly differ between frequent and rare patterns ($F[1, 140] = 1.88, ns$), and did not vary with

Table 1 Experiment: relative Frequency (SD) of the patterns (A) and of the pairs (B) in the triplets of words as a function of pattern frequency (Frequent vs. rare) and pairs frequency (Frequent vs. rare)

A. Patterns relative frequency		Pattern frequency (prime 1-prime 2-target)	
		Frequent	Rare
Pair frequency (Prime 1-target + prime 2-target)	Frequent	.325 (2.2)	.0001 (.007)
	Rare	.210 (.13)	.0001 (.003)
B. Pairs relative frequency		Pattern frequency (prime 1-prime 2-target)	
		Frequent	Rare
Pairs frequency (Prime 1-target + prime 2-target)	Frequent	.009 (.003)	.005 (.006)
	Rare	.0007 (.006)	.0002 (.003)

pattern frequency neither for rare pairs ($F[1, 140] = 3.20$, ns) nor for frequent pairs ($F[1, 140] < 1$, ns).

Participants

A total of 30 participants were involved in the experiment at the universities of Liège and Brussels, Belgium. All participants were master's degree students of ancient Latin who had studied standard Latin texts extensively. Participants had normal or corrected-to-normal visual acuity and had not taken part in other priming experiments.

Apparatus

Experiments were run on a computer to control stimulus presentation and to record participants' response times and errors. Participants were seated 60 cm from a 14-inch PC monitor where stimuli were displayed.

Design

A 2×2 factorial design was used, with pair frequency (frequent pairs vs. rare pairs) being manipulated as a within-participants and between-items variable and pattern frequency (frequent pattern vs. rare pattern) as a within-participants and within-items variable.

Task and procedure

Participants were tested individually. They were told that in each trial three words would be presented in a sequence at the center of the screen: first, two words in lowercase lettering and, second, a word or a pseudo-word in uppercase lettering. They were asked to perform a lexical decision by responding as quickly and as accurately as possible with their dominant hand to indicate whether the third letter string was a Latin word (left-click of the computer mouse) or not (right-click). Twenty practice trials were presented, followed by the experimental block of 144 trials. The order in which trials were presented was randomized and changed for each participant.

Protocol

Stimuli were displayed in 18-point Courier New in black font centered on a white background, with primes in lowercase lettering and targets in uppercase lettering. Each trial consisted of the following sequence of events: a blank screen for 800 ms, a visual warning signal made of three horizontally displayed asterisks for 400 ms, a blank screen for 200 ms; the presentation of the two prime words simultaneously one besides the other on the same line in the middle of the screen for 800 ms. Primes presentation was followed by the inter-stimuli interval for 250 ms; finally, the target was presented in the middle of the screen for 350 ms, followed by a blank screen until the participant's response. Response times and errors were recorded for each participant's response.

Results and discussion of the experiment

Data analyses

Analyses of variance (ANOVA) were performed on response times (RT) with pairs frequency (frequent vs. rare) and pattern frequency (frequent vs. rare) as within-participants. Data from three participants who made more than 15 % errors were excluded from the analyses. For each of the 27 remaining participants, trials in which participants made errors were excluded from the analyses, and a cut-off was set at ± 2.5 standard deviation (SD) units from each participant's mean RT on words (Table 2 and Fig. 4A).

Results of the experiment

Results showed a significant main effect of pairs frequency ($F[2, 26] = 12.71$, $MSE = 3113$, $p < 0.01$), with shorter response times on targets in frequent pairs (717 ms) than on targets in rare pairs (755 ms). This result is in accordance with studies showing that the frequency of pairs of words in text databases predicts priming between the two

words (Spence and Owens 1990; Landauer et al. 1998; Van Petten 2014; Luka and Van Petten 2014). The effect of pairs frequency on the activation of the target is accounted for by computational models in which priming relies on the strength of the pair association, itself learned by Hebbian mechanisms proportionally to frequency of the pair (e.g., Mongillo et al. 2003).

Interestingly, results show a significant interaction between pairs frequency and pattern frequency ($F[1, 26] = 4.78$, $MSE = 8947$, $p < 0.05$), indicating that the frequency of the pattern is a relevant variable that determines the amount of activation of the target by the two primes. Regarding high frequency pairs, the main effect of pattern priming is significant with responses times shorter on targets in a pattern than on targets not in a pattern (54 ms difference; $F[1, 26] = 4.96$, $MSE = 7963$, $p < 0.05$). Further, contrast analyses show that the effect of pairs frequency is significant when words are in a pattern (78 ms difference; $F[1, 26] = 17.56$, $MSE = 4686$, $p < 0.01$) but not when they are not in a pattern (−2 ms difference; $F[1, 26] < 1$), the effect of pattern priming being not significant for rare pairs (non-significant −24 ms difference; $F[1, 26] = 1.73$, $MSE = 5065$, ns). This indicates that in the present experiment, the two primes activate the target when the pattern of three words is frequent and for the most frequent primes-target pairs. The absence of pairs priming in the case of rare pattern is in line with non-significant pair priming for rare pairs (Van Petten, 2014). In the present experiment, the relative frequencies of the rare and frequent pairs were low (.0004 and .007 respectively). Nevertheless, when pairs were embedded in a pattern with the target, priming was significant. Therefore, the effect of pattern priming for frequent pairs, combined to the effect of pairs priming in the case of frequent patterns, can be attributed to a supplementary activation of the target by the pattern that is not due to the activation received through the two prime-target pairs. The activation generated by the pattern is therefore not reducible the sum of the activations generated by the primes, and probably relies on a higher-level of representation than the representation of pair associations between words.

This is the first experimental report of pattern priming of a target word by a combination of two other specific words when the three words occur in a frequent pattern. Note that the effect of pattern frequency arises for equivalent pairs frequencies in rare and frequent patterns. The new phenomenon of pattern priming contributes to the main

priming effect in addition to the classical priming by pair association and cannot be explained solely by pairs priming. Given that pattern frequency was manipulated in the present experiment, pattern priming is assessed for primes–target pair associations that are the most frequent in the text database. Hence, from a methodological point of view, pair associations need not be null, as in Khalkali et al. (2012) and Chwilla and Kolk (2005) to ascertain for pattern priming, but patterns frequency need to be manipulated according to measures from the database.

The present result on pattern priming enriches our understanding of cognitive processing of word sequences, in which pattern priming is an important mechanism of online semantic integration involving higher-levels of semantic coding of pluri-lexical ‘abstract’ concepts in semantic memory. A direct consequence is that current biologically inspired models of priming lack some essential mechanism to account for pattern priming (Brunel 1996; Mongillo et al. 2003; Lavigne and Denis 2001, 2002; Lavigne 2004). Pattern priming therefore raises the question of learning of higher-level representations of more than two words not accounted for by Hebbian learning. Further, results of the experiment challenge models of online processing of patterns in addition to pairs.

Cortical network model

Computational modeling allows linking behavioral data on semantic priming to biologically inspired properties of cortical networks (Mongillo et al. 2003; Brunel and Lavigne 2009; Lavigne et al. 2011; Lerner et al. 2012; Lerner and Shriki 2014; see Bernacchia et al. 2014). Results of the present experiment show that the processing of semantic patterns is content-specific at the level of triplets of words and not only of pairs of words. Thus, pattern priming generates activation of the target that is not reducible to the converging activations from the two primes. This phenomenon cannot be handled by classical models of the cerebral cortex. Up to now, the learning and processing of patterns of three words has been left aside mainly because realistic Hebbian learning was involved, which modifies synaptic efficacy solely on the basis of the local pre- and post-synaptic neurons. On this basis, Hebbian learning can account for learning of triplets (e.g., ABC) *only* as pair associations between the three pairs of the triplet (e.g., AB,

Table 2 Experiment: mean response times in ms (SD) on target words as a function of pattern frequency (Frequent vs. rare) and pairs frequency (Frequent vs. rare)

		Pattern frequency (prime 1–prime 2–target)	
		Frequent	Rare
Pairs frequency	Frequent	690 (131)	744 (206)
(Prime 1–target + prime 2–target)	Rare	768 (193)	742 (152)

AC, and BC). The consequence is that Hebbian learning fails at potentiating synapses as a function of the frequency of the pattern of three words and, hence, fails at reproducing the dynamic processing of patterns involving more than two items (see Rigotti et al. 2010a, b; Lavigne et al. 2014 for discussions).

The neurons' transfer function gives an output that depends on the amount of input but that is independent of the source of the input. An important consequence is that, at the computational level, *any* source of activation can generate an above-threshold response as long as they are of sufficient magnitude. This implies that any pair of primes can activate the target as long as they are associated strongly enough with this target. This mechanism is considered in such models of priming as accounting for semantic activation as a function of the strength of the prime-target association (Lavigne et al. 2011, 2012, 2013). However, the limit of this mechanism is that it does allow for a given target to be activated by a given pair of specific primes learned in a pattern (and not by any pair of associated primes). This contradicts the fact that semantic patterns are frequent occurrences of triplets of *specific* words.

Recent models show that processing content-specific combinations of more than two items cannot be generated solely by nonlinear integration of inputs; they would require additional computational mechanisms at the neuronal level that amplify inputs from the same combination (i.e., a pattern) compared to inputs not learned in that combination. The question of a possible underlying learning mechanism has been recently addressed by modelers and answered by adding a stage of nonlinear integration of the *specific* inputs (e.g., the primes in a pattern with the target) and not of *any* input (e.g., any prime associated with the target). This function can be modeled through additional neurons that handle the nonlinear integration of specific inputs to activate the outputs selectively as a function of specific combinations of inputs. These neurons are present as hidden neurons in hierarchical multi-layers models of object recognition can amplify outputs to specific combinations of inputs (e.g., Riesenhuber and Poggio 1999; Poggio and Edelman 1990) or as mixed-coding neurons in cortical-like architectures (Rigotti et al. 2010a, b; 2011a, b, 2012; Bourjaily and Miller 2011a, b, 2012). A new mechanism has been recently proposed that does need neither additional neurons nor a priori wired multi-layer architectures. This mechanism relies on the joint effects of amplification of synaptic potentiation and of nonlinear integration of specific inputs within the dendritic branches of the neurons coding for the target, this mechanism even improving the efficiency of mixed-coding neurons (Lavigne et al. 2014). Dendrites are nonlinear integrators of the inputs depending on the co-localization of the synapses (e.g. London and Hausser, 2005; Spruston 2008; Lavzin et al. 2012; see Xiumin 2014 for a model). This

mechanism could potentially allow for inputs from two given neurons (e.g., coding for two primes) to be amplified and activate a given target. However, the question of how the corresponding synapses could be potentiated at the adequate localization along the dendritic tree to have their inputs amplified was a pending problem. Thanks to recent neurophysiological studies, experiments have shown that when several synapses are activated at the same time within the same dendritic branch, the potentiation of these synapses is larger than when they are in different dendritic branches (Govindarajan et al. 2011; Fig. 6). An inter-synaptic learning mechanism of synaptic potentiation has been formalized and extended to inter-synaptic depression of synapses (Lavigne et al. 2014). The question addressed here regards the possibility that inter-synaptic learning can amplify potentiation of synapses between three words learned in a pattern compared to groups of three words learned only in pairs and, in turn, that the learned values of efficacy can generate dynamic activation of neurons coding for the words and, hence, pattern priming. The modeling approach proposed here aims at linking learning of patterns of words—as a function of the relative frequency of occurrence of the pairs and patterns—and the neural dynamics necessary to generate pattern priming.

Method of the model

We describe here the architecture of a cortical network of excitatory neurons that is regulated by inhibitory feedback (Fig. 2). In this network, words are coded by populations of excitatory neurons that exhibit a selective activity for each item (e.g., word) presented individually (Brunel 1996; Amit and Brunel 1997; Mongillo et al. 2003; Curti et al. 2004; Romani et al. 2006). The neural spiking dynamics accounting for the activity of the items in memory are explained in realistic biophysics terms by reverberating activation between neurons of the same population, or of different populations connected with potentiated synapses (Amit and Brunel 1997; Mongillo et al. 2003). Each neuron has a constant number of dendrites, each of them having a constant number of synapses. Synaptic connections between pre-synaptic neurons and the dendrites of post-synaptic neurons are random (Fig. 2). This new architecture makes possible inter-synaptic (IS) learning at the level of dendrites. IS learning takes into account not only of the activity of the pre- and post- synaptic neurons, but also of the other active synapses within the same dendrite (Lavigne et al. 2014). The main idea is that, in each individual dendrite, the potentiation or depression of a given synapse between a post-synaptic and a pre-synaptic neuron is amplified as a function of the number of other synapses from other active pre-synaptic neurons (Table 3).

Fig. 2 Schematic architecture and synaptic connectivity of the cortical network model embedding inter-synaptic learning. For clarity, connections to one dendrite of one neuron coding for the target (pink dot) learned in a pattern and in pairs with the primes (dark and light blue dots) are displayed. Values of connectivity are given in Table 3. Excitatory neurons are selective for distinct stimuli (color

dots), and inhibitory neurons provide unselective inhibition (black dots). According to the inter-synaptic learning algorithm, efficacy values within a same dendrite (thickness of the arrows) is larger between neurons coding for primes and target learned in patterns (between blue and pink neurons) than for words learned in pairs only (between purple and pink neurons). (Color figure online)

Network architecture and synaptic connectivity

The model includes a biophysically realistic cortical network of N_E excitatory glutamatergic pyramidal cells whose

activity is regulated by $N_I = 0.25 N_E$ inhibitory GABAergic inter-neurons (Abeles 1991; Braitenberg and Schütz 1991), with a probability $C = 0.2$ of having a synapse from any pre-synaptic neuron to any post-synaptic

Table 3 Model: mean response times (see text) in ms (SD) on target words as a function of pattern frequency (Frequent vs. rare) and pairs frequency (Frequent vs. rare)

		Pattern frequency (prime 1-prime 2-target)	
		Frequent	Rare
Pairs frequency	Frequent	691 (3)	742 (8)
(Prime 1-target + prime 2-target)	Rare	760 (17)	763 (13)

Table 4 Parameters of the model

N_E	Number of excitatory neurons	4000
N_I	Number of inhibitory neurons	1000
C	Connectivity	0.2
C_E	Number of recurrent excitatory connections per neuron	800
C_{ext}	Number of external excitatory connections per neuron	2200
C_I	Number of recurrent inhibitory connections per neuron	200
g	Number of populations	6
fN_E	Number of neurons per population	400
f	Coding level	$N_{Ep}/N_E = 0.1$
N_s	Number of synapses per dendrite	16
N_d	Number of dendrites per neuron	50
δ_E	Latency (transmission delay), excitatory neurons	15–30 ms
δ_I	Latency (transmission delay), inhibitory neurons	0.5 ms
τ_{AMPA}	Synaptic decay type, AMPA-R	2 ms
τ_{NMDA}	Synaptic decay type, NMDA-R	100 ms
τ_{GABA}	Synaptic decay type, GABA-R	5 ms
X	Fraction of NMDA currents	0.3
τ_{ml}	Membrane time constant, inhibitory neurons	10 ms
Θ	Firing threshold, both types	20 mV
$V_{\tau E}$	Reset membrane potential, excitatory neurons	10 mV
$V_{\tau I}$	Reset membrane potential, inhibitory neurons	15 mV
τ_{RP}	Refractory period, both types	2 ms
$q+$	Intrinsic probability of potentiation	0.01
$q-$	Intrinsic probability of depression	0.01
J_{EE}	Average E $\rightarrow$ E efficacy	0.047 mV
J_{IE}	E $\rightarrow$ I efficacy	0.09 mV
J_{EI}	I $\rightarrow$ E efficacy	−0.27 mV
J_{II}	I $\rightarrow$ I efficacy	−0.5 mV
J_{Eext}	External E $\rightarrow$ E efficacy	0.052 mV
J_{Iext}	External E $\rightarrow$ I efficacy	0.1 mV
J_I	Potentiated E $\rightarrow$ E efficacy between neurons coding for the same item	0.098 mV
J_a	Potentiated E $\rightarrow$ E efficacy between associated items	$J_0 + a(J_I - J_0)$
J_0	Depressed E $\rightarrow$ E efficacy between non-associated items	$(J_{EE} - fJ_I)/(1 - f)$
v_{ext}	External Poisson noise	15 Hz
Λ	Contrast of external input	0.08
I_{min}	Minimum value of dendritic current	0
I_{max}	Maximum value of dendritic current	0.115

neuron (see Brunel and Wang 2001). Different populations of excitatory neurons code for different words, and 40 % of the excitatory neurons do not encode any particular single word. Neurons in the network are connected through four types of synapses. Synaptic efficacies between excitatory neurons (EE) are subject to variations due to learning. In contrast, synaptic efficacies involving inhibitory neurons, that is, excitatory to inhibitory (IE), inhibitory to excitatory (EI), or inhibitory to inhibitory (II) are not subject to learning. The negative retroaction by inhibitory interneurons prevents the runaway propagation of activation and regulates population dynamics in the network. Every

excitatory post-synaptic neuron i has a set of dendrites, each having several excitatory (and inhibitory) synapses connecting the post-synaptic neuron i to pre-synaptic neurons. Given that connections between neurons are random, each neuron has specific dendrites with specific combinations of synapses from different pre-synaptic neurons (see Table 4 for parameters).

Inter-synaptic learning

Physiological experiments report that dendritic compartmentalization influences the pairing of excitatory post-

potentials generated in dendrites (EPSP) and action potentials (see Spruston 2008). This can in turn amplify the induction of long-term potentiation (LTP) for synapses that are close within the same dendritic branch rather than distant across branches (Govindarajan et al. 2011; Harvey and Svoboda 2007), particularly when several synapses are stimulated within a branch (see Fig. 6 in Govindarajan et al. 2011).

Here we consider binary plastic synapses with two discrete states: a potentiated state and a depressed state. The formalism of classical Hebbian learning generates potentiation or depression of synapses as a function of the activity of the two pre- and post-synaptic neurons (e.g., Brunel et al. 1998). For simplicity, we will consider here neurons whose current state $v_i \in [0; 1]$ is driven by the presence or absence of the word it codes for (prime 1, prime 2, or target), described as a binary string $\xi_i \in \{0; 1\}$. Classic Hebbian learning describes LTP when the two neurons are activated (Hebb 1949; Bliss and Lomo 1973; Bliss and Collingridge 1993). The synapse in the Down state has an instant probability q^+ to be switched to the Up state. Long-term depression (LTD) occurs when only one of the two neurons is active (e.g., Kirkwood and Bear 1994)—the synapse in the Up state has an instant probability q^- of being switched to the Down state. As a result, a synapse ij between pre- and post-synaptic neurons i and j has a probability $a_{ij} = q^+ \xi_i \xi_j$ to potentiate, a probability $b_{ij} = q^- [\xi_i(1 - \xi_j) + \xi_j(1 - \xi_i)]$ to depress, and probability $\lambda_{ij} = 1 - a_{ij} - b_{ij}$ that no change occurs (Brunel et al. 1998).

Inter-synaptic learning describes how LTP and LTD of a given synapse ij are amplified proportionally to the number of synapses ik , il , etc. that are co-transmitting within the same dendritic branch (Lavigne et al. 2014). The potentiation of a synapse between a post-synaptic and a pre-synaptic neuron increases with the increasing number of synapses also active within the same dendrite due to the activity of other neurons pre-synaptic within this dendrite (Govindarajan et al. 2011). The probability $a_{ij}(D)$ of potentiating a synapse connecting two active neurons i and j within a dendrite D is:

$$a_{ij}(D) = q^+ \xi_i \xi_j + q^+ \xi_i \xi_j (n_j - 1) + q^+ \xi_i \xi_j \left(\sum_{u=1, u \neq j}^g \xi_{Pu} n_u \right) \quad (1)$$

With $q^+ \xi_i \xi_j$ representing the synaptic potentiation due to classic Hebbian learning, $q^+ \xi_i \xi_j (n_j - 1)$ representing the inter-synaptic amplification of the potentiation due to co-active synapses connecting neurons from the same population P_j , and $q^+ \xi_i \xi_j \left(\sum_{u=1, u \neq j}^g \xi_{Pu} n_u \right)$ representing the inter-synaptic amplification of potentiation due to co-active synapses connecting neurons from different populations P_u .

As in the case of amplification of potentiation, depression of a synapse connecting two neurons i (active) and j (inactive) within a dendrite D depends on the number of other active synapses in D (see Lavigne et al. 2014). Based on the probabilities of potentiating $a_{ij}(D)$ or depressing $b_{ij}(D)$ synapses at each learning step, the inter-synaptic learning rule provides a mathematically tractable description of the average values of potentiation of the synapses J_{ij} after learning of either pairs or of combinations of three or more items. Learning of prime-target pairs only generates association between each prime and the target through the Hebbian component of Eq. 1 ($q^+ \xi_i \xi_j$), whereas learning of patterns of two primes and a target generates a stronger association between the primes and the target through the IS component of Eq. 1 ($q^+ \xi_i \xi_j \left(\sum_{u=1, u \neq j}^g \xi_{Pu} n_u \right)$) between neurons coding for each prime and neurons coding for the target. This IS component makes possible stronger associations between primes and target when they are learned in patterns than when they are learned in pairs only.

Learning is considered here as independent of the precise order of presentation of the pairs and patterns of words. To this aim, according to Brunel et al. (1998), we consider the case of complete and slow learning, corresponding to low values of instant probabilities of potentiation q^+ and of depression q^- of synapses. Considering a synapse between neurons coding for a given prime and neurons coding for a given target, its average probability J_{ij} to be potentiated after learning of all combinations of items in a set of pairs and triplets of items converges to $J_{ij}(D) = \frac{\langle a_{ij}(D) \rangle}{\langle a_{ij}(D) + b_{ij}(D) \rangle}$ (see Lavigne et al. 2014). For the simulations, the average probabilities to potentiate $\langle a_{ij}(D) \rangle$ and to depress $\langle b_{ij}(D) \rangle$ each synapse in each dendrite of each neuron are calculated by using the frequencies of pairs and patterns measured in the corpus for the items used in the experiment (Table 1a, b).

According to the Hebbian component, and in line with Fiser and Aslin's (2001) results, when learning the set of items the synapse J_{ij} potentiates each time the prime and target are present in a pair, and depresses each time the prime or target is present without the other (when occurring with other words but not together). Synaptic efficacy therefore increases according to the number of occurrences of the prime and target together in the learning set, and decreases according to the number of occurrences of the prime or the target without the other item. Accordingly, we take the frequency of occurrence of the prime and target pair as the number of cases of potentiation to calculate the probability a_{ij} to potentiate the synapse, and the frequency of occurrence of the prime or target to calculate the probability b_{ij} to depress the synapse. Synaptic efficacy therefore depends on the relative frequency of the prime-

target pair given the absolute frequency of each item taken alone.

According to the IS component, the synapse J_{ij} potentiates each time the two primes *and* target are presented together in a pattern, and depresses each time two of the items are present without the third one. Synaptic efficacy therefore increases according to the number of occurrences of the two primes *and* the target together, and decreases according to the number of occurrences of two of the items without the other. Accordingly, the values of a_{ij} and b_{ij} are also updated as a function of the frequency of the primes-target patterns. We take the frequency of occurrence of the two primes *and* target pattern as the number of cases of potentiation to calculate the probability a_{ij} to potentiate the synapse, and the frequency of occurrence of two items (e.g., the two primes) *or* the third one (e.g., the target) (when, e.g., the two primes occur with other words than the target) to calculate the probability b_{ij} to depress the synapse. Synaptic efficacy therefore depends on the relative frequency of the primes-target pattern given the absolute frequency of the each pair of the pattern taken alone.

The resulting precise value of synaptic efficacy depends on the frequency of occurrence of the items, but also on the number of synapses from different neurons in each dendrite of each neuron in the network. It is therefore very variable from dendrite to dendrite (see Lavigne et al. for a discussion).

Dendritic and neuronal dynamics

Upon the emission of a pre-synaptic spike, an epsp/ipsp is generated within the dendrite D . The total dendritic current I_D is calculated in each dendrite D as a function of the linear integration of the voltage-independent AMPA and GABA synaptic currents, generated by post-synaptic receptors GABA and AMPA, and of the nonlinear integration of voltage-dependent NMDA currents, generated by NMDA receptors (see Lavigne et al. 2014). Each neuron i of the network is a leaky integrate-and-fire neuron (Tuckwell 1988) whose state is described by its total depolarization V (mV) and is calculated as follows:

$$\tau_m \frac{dV(t)}{dt} = -V(t) + \sum_D I_D(t) \quad (2)$$

when V reaches a threshold V^θ , the neuron emits a spike, and V^T is reset to V_τ , following a refractory period τ_{RP} . Note that the nonlinear integration of dendritic current within each dendrite guarantees that the same synapses (i.e., within a given dendrite) that have benefited amplification of potentiation during learning will also benefit amplification of their inputs during processing of the items. This way

two input primes in a pattern with the target will have stronger efficacies within a given dendrite *and* their combined inputs will be amplified within this dendrite. Inputs from different dendrites, although with stronger efficacies, will not be amplified. This guarantees that inputs from primes learned in patterns will be amplified through nonlinear dendritic integration. Hence the two mechanisms together—of amplification of efficacies during learning and of amplification of the inputs during processing—make dendrites receptive to learned combinations of specific inputs (i.e., patterns of primes and target).

After the two stages of nonlinear dendritic and somatic integration, spike rates ν of populations coding for the primes and target are calculated (10 ms bins) during a priming trial as a function of the values of synaptic efficacy generated by learning pairs only or of a pattern of three items.

Results of the simulations

Rate dynamics

Numerical simulations in the model used the same multiple priming protocol as in the experiment, with two primes preceding a target. According to the conditions of pairs only vs. pattern tested in the experiment, simulations tested for triplets of words learned as independent pairs—involving the Hebbian component of potentiation of synapses between neurons coding for the primes and target—or learned as patterns of three words involving amplification of the potentiation of synapses. A simulation of an experimental trial is displayed in Fig. 3.

In the simulations, a trial begins with the network in a state of low-level spontaneous activity due to external noise on excitatory and inhibitory neurons obeying a Poisson process of rate $\nu_{\text{ext}} = 15$ Hz (see Brunel and Lavigne 2009). At the presentation of the primes, the corresponding populations reach an elevated level of activity (perceptive response; Brunel and Wang 2001) that decreases after the primes offset but remains above the level of spontaneous activity because of the strong excitatory feedback due to potentiated synapses between neurons coding for each prime. This behavior reproduces the elevated firing rates of neurons coding for a stimulus following its presentation, as reported in nonhuman primates (Fuster and Alexander 1971; Miyashita 1988; Miyashita and Chang 1988). Such retrospective activity is considered subtending the activation of items in working memory (Amit and Brunel 1997; Brunel and Wang 2001; Haarmann and Usher 2001; Renart et al. 2001; Amit et al. 2003). Retrospective activity of the primes leads in turn to the activation of the population of neurons coding for their associated target, due to potentiated synapses between

Fig. 3 Behavior of the cortical network model. The light-gray area corresponds to time of presentation of the two primes (50 ms to 250 ms). Spike rates of the populations coding for the primes (light and dark blue lines) and four different targets: a target learned in frequent pairs and in a pattern with the primes (pink line); a target learned in frequent pairs but not in a pattern with the primes (light green line); a target learned in rare pairs and in a pattern with the primes (purple line); a target learned in rare pairs and not in a pattern with the primes (dark green line). The spike rates of the primes are in perceptive response during presentation of the primes and are in retrospective activity after the offset of the primes. This generates prospective

activity of the targets at levels that depends on synaptic efficacies. Prospective activity of the target learned in frequent pairs and in a pattern with the primes (pink line) is higher than activity of the target learned in rare pairs and in a pattern (light green line), and of targets learned only in pairs with the primes (purple and dark green lines). Prospective activity of a target learned in very frequent pairs with the primes is displayed (grey line) to show that the model can generate prospective activity and priming without the need for a pattern pair under the condition that pairs associations are strong. (Color figure online)

neurons coding for the primes and neurons coding for the target. This behavior reproduces the increasing firing rates—above the level of spontaneous activity—of neurons coding for the associates to the stimulus presented, also as reported in nonhuman primates (Miyashita 1988; Miyashita and Chang 1988; Erickson and Desimone 1999; Rainer et al. 1999; Sakai and Miyashita 1991; Naya et al. 2001, 2003; see Fuster 2001). Such prospective activity is considered subtending the contextual recall of knowledge in memory (Brunel 1996; Lavigne and Denis 2001, 2002; Lavigne 2004; Mongillo et al. 2003) and associative priming effects in accordance with behavioral experiments in nonhuman primates (Erickson and Desimone 1999). Retrospective and prospective activities have been shown to reproduce various dynamics of priming in humans, involving one (Brunel and Lavigne 2009; Lavigne and Darmon 2008) or two primes (Lavigne et al. 2011). In the current protocol, prospective activity is generated by the joint retrospective activities of the two primes, and its magnitude depends on the potentiation of synapses between neurons coding for the primes and neurons coding for the target (see Fusi et al. 2007; Rigotti et al. 2010a, b; Lavigne et al. 2011, 2012, 2013, 2014).

Within the framework of the multiple priming protocols used in the literature and in the present experiment, levels of prospective activity of neurons coding for the target are compared after pairs learning and after pairs and pattern learning (Fig. 3). After pairs learning only, the low values

of synaptic efficacy between neurons coding for the primes and neurons coding for the target lead to a low level of prospective activity of the populations coding for the target (rare pairs, dark green line, frequent pairs, light green line). After pattern learning, the amplification of potentiation of synapses between neurons coding for the primes and neurons coding for the target—due to inter-synaptic learning—combined to the nonlinear integration of the inputs of these same primes, leads to an increased level of prospective activity of the target in a pattern and in frequent pairs (pink line) compared to the targets not in a pattern (light green line) and to the target in a pattern and in rare pairs (purple line). The IS learning algorithm therefore allows the network to generate different levels of prospective activity, depending on the type of learning of the primes and target (occurrences of pairs only or occurrences of patterns), with the higher level for the target learned in frequent pairs and in a frequent pattern.

Pattern priming

Priming effects can be calculated as a function of the activity of the population of neurons coding for the target at target onset. Indeed, electrophysiological studies in non-human primates report a negative correlation between spike rates and response times (Roitman and Shadlen 2002) as well as shorter reaction times on targets and increased prospective activity of the corresponding neurons, when

Fig. 4 **A** Experiment: response times and mean square errors on targets learned in frequent versus rare pairs, and in frequent versus rare patterns with the primes. **B** Model: response times and mean

square errors on targets using values of synaptic efficacies calculated according to the frequencies of the pairs and patterns used in the experiment (see the [Model](#) section and Table 3)

the target is preceded by an associated prime (Erickson and Desimone 1999). On this basis, we calculate the reaction time T (ms) on the target as inversely proportional to the firing rate v (Hz) of the corresponding neuron population at target onset ($T = a + \frac{b}{v}$, as usually done in models (see Brunel and Lavigne 2009; Lavigne and Darmon 2008; Wong and Wang 2006; Wang 2002). T therefore decreases with the increasing level of prospective activity of the neuron population coding for the target, itself depending on the efficacy values of synapses between neurons coding for the primes and neurons coding for the target after learning pairs or patterns. Simulations of the four experimental conditions of relative frequencies led to specific reaction times. The relative frequencies of patterns and pairs are learned as a function of their absolute frequencies and of the frequencies of the primes and primes pairs with other words than the target. Therefore a target can be in pairs of low relative frequencies (when each primes occurs with many other words) but in a pattern of high relative frequency (when the two primes together occur mainly with the target and not with other words). The four experimental conditions of primes-target relation that are tested in the model (2 pairs vs. 2 pattern relative frequencies) then lead to four reaction times: $T_{+pairs}^{+pattern}$ (frequent pattern and frequent pairs), $T_{-pairs}^{+pattern}$ (frequent pattern and rare pairs), $T_{+pairs}^{-pattern}$ (rare pattern and frequent pairs) and $T_{-pairs}^{-pattern}$ (rare pattern and rare pairs). Here response times of the model are matched to the ones of participants to the experiment with $a = 680$ and $b = 1000$. The difference in reaction times between the four experimental conditions allows us to calculate the effects of pattern priming and of pairs priming as in the experiment (Fig. 4B).

Response times of the model vary stochastically with variation in the prospective activity of the neurons coding for the target, due to the stochastic noise. Data were therefore analyzed on a 12 ms time window before presentation of the target for 10 runs (120 ms of recording for each of the four types of target). The Analysis of the response times of the model involved the same four conditions as in the behavioral experiment (Fig. 4B). Results show a significant main effect of pairs frequency ($F[1, 120] = 2850$, $MSE = 2850$, $p < 0.001$), in line with results of the experiments and with classical priming effects between prime and target pairs. Results also show a significant effect of pattern frequency ($F[1, 120] = 615$, $MSE = 145$, $p < 0.001$), in line with results of the experiment. As in the behavioral experiment, pattern priming is significant for frequent pairs ($F[1, 120] = 7719$, $MSE = 20$, $p < 0.001$) but not for rare pairs ($F[1, 120] = 2.04$, $MSE = 342$, ns), with a significant interaction between the effects of pairs frequency and pattern frequency ($F[1, 120] = 314$, $MSE = 218$, $p < 0.001$). This indicates that, for the pairs frequencies considered in the experiment and model, a target that gains activation from each prime in a pair and from the pattern is more activated than a target only in pairs or only in a pattern with the primes. Due to the higher spike rates of populations of neurons coding for a target in pairs and in a pattern with the primes, the model reproduces the effects of pairs priming and of pattern priming found on participants' response times in the behavioral experiment.

This is the first modeling of the effects of pattern priming that depend on learned associations between words in patterns of three co-occurrent words in addition to associations between words in pairs. The synergistic effects

of the activation of the target by each prime and by the pattern are modeled by the combination of two mechanisms (pink line, Fig. 3, see Fig. 4B). First, during learning, inter-synaptic amplification of the potentiation of synapses generates stronger efficacy between neurons coding for a prime and target when another prime is present (pattern) compared to pairs; second, during online processing of the two primes, the nonlinear integration of the NMDA input currents at the level of dendrites of neurons coding for the target generates higher levels of prospective activity of the target when learned in a pattern with the primes. In the model response times inversely correlate with the average spike rate of the population coding for the target after presentation of the primes (prospective activity of the target). The level of prospective activity of the target is itself related to the synaptic efficacies learned in pairs and in patterns accordingly to the frequencies measured in the text database for the experimental items. As a consequence, pairs frequencies different from the one used in the experiment may lead to different synaptic efficacies, to different levels of prospective activity, and then to different response times and priming effects (see the higher prospective activity for very frequent pairs in the absence of patterns; Fig. 3, grey line).

Discussion

The present experiment shows that multiple priming effects generated by two primes on a target depend on the relative frequency of occurrence of the prime-target pairs in text databases. When pairs only are involved, each prime activates the target as a function of the strength of its association with the target, in addition to the activation of the target by the other prime (see Lavigne et al. 2011 for a meta-analysis). Further, results show for the first time that priming effects are larger when the primes and target occur in databases in patterns with high relative frequency. When a pattern is involved, each prime activates the target *and* the combination of the two primes increases the activation of the target through increased synaptic efficacies for learned patterns. Whereas previous experiments compared priming from two primes to an absence of priming by each prime in isolation (Khalkhali et al. 2012) or to an absence of association of the target to each prime (Chwilla and Kolk 2005), the present experiment compared the effects of two levels of relative frequency of the patterns. This manipulation permitted the identification of an interaction between the effects of pair priming and of pattern priming, with two primes always presented in each condition of pattern frequency. In addition to the classical effect of pairs priming, a target learned in a pattern benefits more activation than a target not learned in a pattern (Fig. 4).

In the present experiment, the two modalities of pattern frequency were manipulated within subjects and within a same protocol that involved two primes. The measure of two levels of pattern frequency and the multiple priming protocol has allowed to rule out the possibility that the primes may be too weakly associated to the target to generate priming in isolation but could generate priming when presented together.

With the material used here, due to the low frequency of occurrence of the pairs, pairs-only priming is not significant for low-frequency patterns, although occurrences of the pairs are not null. This result is coherent with previous results on non-significant pairs priming for low pairs frequencies (e.g., Van Petten 2014). However pairs priming effects for higher values of pairs frequencies classically reported in the literature (e.g., Lavigne et al. 2011 for a meta-analysis) had to be also reproduced by the model. Pairs priming was tested in the model after learning of prime-target pairs of very high values of relative frequency. Simulations show that the prospective activity of a target in very frequent pairs (Fig. 3, grey line) is increased compared to rare (dark green line) and frequent (light green line) pairs. Pairs priming occurs here in the absence of learning of patterns. This is a model's prediction that for values of prime-target pairs frequency higher than the ones used in the experiment, the effect of pairs priming would be stronger than in the experiment even though the target is not in a pattern with the prime.

Pattern priming and nonlocal learning

The results of the experiment indicate that during language comprehension, sequences of three words are minimal sequences sufficient to involve higher-order pattern priming effects that are *quantitatively* different from pair priming effects. The learning model suggests that patterns priming effects are based on learning mechanisms of patterns that are *qualitatively* different from Hebbian pair learning. A direct consequence of the experimental results for computational models is that pattern priming cannot be attributed solely to the supra-threshold integration of two inputs that would otherwise be too weak to generate priming in isolation. The performance of the network in pattern priming relies on its ability to activate the target not only as a function of its association to each prime but also as a function of its association with the combination of primes. This is made possible by the dendritic amplification of synaptic inputs in cases of patterns compared to pairs, itself depending on the inter-synaptic amplification of synaptic efficacies between neurons coding for words learned frequently in patterns. Pattern priming is due to the non-linear integration of the inputs by the two primes *and* to the amplified synaptic efficacies when learned in a

pattern. This non-linear integration of two primes in a pattern with the target exceeds that of a single prime, and also exceeds that of two primes learned only in pairs with the target.

Pattern priming indicates that learning of patterns is not coded as synaptic potentiation only between pairs of neurons, but involves higher-order combinations of items. The inter-synaptic learning rule used here is biologically rooted and cognitively relevant to model pattern priming after learning patterns of words. Interestingly, the inter-synaptic learning rule potentiates (or depresses) synapses as a function of the activity of the post-synaptic neuron and of several pre-synaptic neurons. It is therefore a nonlocal though biologically realistic learning rule that suits the cognitive view of context-dependent activation.

Context-dependent activation and combinatorial meaning

Experimental evidence indicates that the comprehension of language relies on the online prediction of upcoming words based on the associations between words in memory (DeLong et al. 2005; Willems et al. 2016; see Kutas, DeLong, and Smith 2011). Recent studies have suggested that the magnitude of priming effects depends on contextual representations activated by specific tasks given to participants (Gollwitzer and Kinney 1989; Bermeitinger, Wentura, and Frings 2011; Kiefer 2007; Spruyt, De Houwer, and Hermans 2009; Kiefer and Martens 2010). Multiple priming experiments show that priming of a target by a prime can be increased by another prime also associated with the target (McNamara 1992; Balota and Paul 1996; see Faust and Kahana 2002; Faust and Lavidor 2003) at each processing time (Lavigne and Vitu 1997; Lavigne et al. 2012, 2013; see Lavigne et al. 2011 for a meta-analysis and model). From a general point of view, contextual activation may depend on the fact that, among the associates of a prime that are activated by this prime, some can benefit from additional activation from a context or another prime. In that case, context can be seen as biasing the way activation propagates from a prime to its associates, activating some associates more than others when they are also associated with its context (context-relevant). In such view of contextual activation, the activation of associates only depends on the presence of a context that activates them or not in addition to the prime. However, pattern priming indicates that the level—and even the possibility—of activation of associates to a prime depends on another specific prime (i.e., context), not only because they are also associated with this other prime but because they are combined with the two primes together. It is a particular case of context-dependent activation in which

the activation of the target is magnified—and even made possible—by a combination of specific primes but not by other primes even though they are each individually associated with it. Here, each prime word plays the role of a context to the processing of the other prime. This view of context-dependent activation between words suggests a new type of meaning activated during processing of sequences of words that have previously been learned as patterns.

In the present experiment, the presentation of the two primes simultaneously probably improved processing of the patterns that included the two primes. In this case participants could predict targets in a pattern with the primes as fast as possible. A sequential presentation of the primes could have incited participants to anticipate the second prime and the target on the basis of pairs associations only (given that only one prime is available at first), and hence to under-estimate the information on patterns when the second prime was presented. Indeed, studies on multiple priming show that priming effects generated by the second prime can be perturbed by processing of the first prime. In multiple priming experiments, each prime can generate interference on the processing of the associates to the other prime, probably through unselective retroactive inhibition (see Lavigne et al. 2011). Interference mechanisms could have perturbed processing of patterns if only one prime was presented first. Further experiments are needed to investigate the precise time-course of pattern priming effects as it has been done on pairs priming effects.

Combinatorial meaning and semantic integration

The experimental results and modeled behavior give new insights on the activation of a word's meaning and semantic integration during processing of word sequences. Previous studies on multiple priming have shown that the meaning activated in real time by the processing of two primes can range from a “single meaning” in which associates of one prime are changed to a “double meaning” in which associates of each of the two primes are activated simultaneously, with an intermediate case of “integrated meaning” in which common associates of the two primes are activated (Lavigne et al. 2012, 2013; see Beeman et al. 1994; Whitney et al. 2010; Whitney et al. 2009). In the latter case, pattern priming could play a central role in amplifying the activation of associates that are in patterns with the two primes compared to associates that are not in patterns with the two primes. Pattern priming would make possible the stronger activation of associates that are the most relevant, given the combination of primes compared to associates of both primes but not in a pattern. Semantic integration would then take into account the combination of primes and not only of their mutual presence. The

cognitive system activates a “combinatorial meaning” to integrate the meaning of combinations of words processed. Within such combinatorial meaning, associates of the combination would be activated more strongly than those associated with each of the two primes but with their combination, and more strongly than associates of only one of the two primes. These variable levels of activation of associates of the primes could be a solution to the selection of word meaning in the limited capacity of working memory (Vergauwe and Cowan 2015). Pattern priming is a good candidate process by which to select associates that are most relevant to the combination of primes, suggesting that combinations of words are more important than their simple addition. The model of inter-synaptic learning and pattern priming presented here could therefore describe a central mechanism involved in combinatorial learning and semantic integration in language comprehension.

Acknowledgments We thank Benoît Morimont for his assistance in the data collection, and the students of the University of Liège for their participation. This project was supported by a Programme Hubert Curien granted by Wallonie-Bruxelles International and the French Ministry of Foreign Affairs. Frédéric Lavigne, Dominique Longrée, and Damon Mayaffre were supported by a grant from PEPS CNRS and the Université de Nice Sophia Antipolis.

References

- Abeles M (1991) *Corticonics*. Cambridge University Press, New York, NY. doi:10.1017/CBO9780511574566
- Abernethy M, Coney J (1993) Associative priming in the hemispheres as a function of SOA. *Neuropsychologia* 31(12):1397–1409
- Amit DJ, Brunel N, Tsodyks MV (1994) Correlations of cortical hebbian reverberations: theory versus experiment. *J Neurosci* 14(11):6435–6445
- Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex* 7(3):237–252
- Amit DJ, Bernacchia A, Yakovlev V (2003) Multiple-object working memory—a model for behavioral performance. *Cereb Cortex* 13(5):435–443
- Anderson JR (1976) *Language, memory, and thought*. Lawrence Erlbaum, Oxford
- Anderson JR (1983a) Retrieval of information from long-term memory. *Science* 220(4592):25–30
- Anderson JR (1983b) A spreading activation theory of memory. *J Verbal Learn Verbal Behav* 22(3):261–295
- Arnott WL, Chenery HJ, Copland DA, Murdoch BE, Silburn PA (2003) The effect of backward masking on direct and indirect semantic priming in patients with Parkinson’s disease and in healthy individuals. *Brain Lang* 87:101–102
- Balota DA, Paul ST (1996) Summation of activation: evidence from multiple primes that converge and diverge within semantic memory. *J Exp Psychol Learn Mem Cogn* 22(4):827–845
- Becker S, Moscovitch M, Behrmann M, Joordens S (1997) Long-term semantic priming: a computational account and empirical evidence. *J Exp Psychol Learn Mem Cogn* 23(5):1059–1082
- Beeman M, Friedman RB, Grafman J, Perez E (1994) Summation priming and coarse semantic coding in the right hemisphere. *J Cogn Neurosci* 6(1):26–45
- Bennett DJ, McEvoy CL (1999) Mediated priming in younger and older adults mediated priming in younger and older adults. *Exp Aging Res* 25:141–159
- Bermeitinger C, Wentura D, Frings C (2011) How to switch on and switch off semantic priming effects for natural and artifactual categories: activation processes in category memory depend on focusing specific feature dimensions. *Psychon Bull Rev* 8:579–585
- Bernacchia A, La Camera G, Lavigne F (2014) A latch on priming. *Front Psychol* 5:869. doi:10.3389/fpsyg.2014.00869
- Bliss TV, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361(6407):31–39
- Bliss TV, Lomo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol* 232(2):331–356
- Bourjaily M, Miller P (2011a) Excitatory, inhibitory, and structural plasticity produce correlated connectivity in random networks trained to solve paired-stimulus tasks. *Front Comput Neurosci* 5(37). doi:10.3389/fncom.2011.00037
- Bourjaily M, Miller P (2011b) Synaptic plasticity and connectivity requirements to produce stimulus-pair specific responses in recurrent networks of spiking neurons. *PLoS Comput Biol* 7(2):e1001091. doi:10.1371/journal.pcbi.1001091
- Bourjaily M, Miller P (2012) Dynamic afferent synapses to decision-making networks improve performance in tasks requiring stimulus association and discrimination. *J Neurophysiol* 108:513–527
- Braitenberg V, Schütz A (1991) *Anatomy of the Cortex*. Springer-Verlag, Berlin. doi:10.1007/978-3-662-02728-8
- Brodeur DA, Lupker SJ (1994) Investigating the effects of multiple primes: an analysis of theoretical mechanisms. *Psychol Res* 57(1):1–14
- Brunel N (1996) Hebbian learning of context in recurrent neural networks. *Neural Comput* 8:1677–1710
- Brunel N, Latham PE (2003) Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural Comput* 15(10):2281–2306
- Brunel N, Lavigne F (2009) Semantic priming in a cortical network model. *J Cogn Neurosci* 21(12):2300–2319
- Brunel N, Wang X-J (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11(1):63–85
- Brunel N, Carusi F, Fusi S (1998) Slow stochastic Hebbian learning of classes of stimuli in a recurrent neural network. *Network* 9:123–152
- Chwilla DJ, Kolk HHJ (2005) Accessing world knowledge: evidence from N400 and reaction time priming. *Cogn Brain Res* 25:589–606
- Chwilla DJ, Kolk HHJ, Mulder G (2000) Mediated priming in the lexical decision task: evidence from event-related potentials and reaction time. *J Mem Lang* 42:314–341
- Coney J (2002) The effect of associative strength on priming in the cerebral hemispheres. *Brain Cogn* 50(2):234–241
- Cree GS, McRae K (2003) Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J Exp Psychol Gen* 132:163–201
- Cree GS, McRae K, McNorgan C (1999) An attractor model of lexical conceptual processing: simulating semantic priming. *Cogn Sci* 23(3):371–414
- Curti E, Mongillo G, La Camera G, Amit DJ (2004) Mean field and capacity in realistic networks of spiking neurons storing sparsely coded random memories. *Neural Comput* 16:2597–2637
- DeLong KA, Urbach TP, Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat Neurosci* 8(8):1117–1121

- Erickson CA, Desimone R (1999) Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci* 19(23):10404–10416
- Ermentrout GB, Kopell N (1986) Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM J Appl Math* 46:233–253
- Faust M, Kahana A (2002) Priming summation in the cerebral hemispheres: evidence from semantically convergent and semantically divergent primes. *Neuropsychologia* 40(7):892–901
- Faust M, Lavidor M (2003) Semantically convergent and semantically divergent priming in the cerebral hemispheres: lexical decision and semantic judgment. *Cogn Brain Res* 17(3):585–597
- Ferrand L, Alario F-X (1998) Normes d'associations verbales pour 366 noms d'objets concrets. *L'Année Psychologique* 98:659–709
- Fiser J (2009) Perceptual learning and representational learning in humans and animals. *Learn Behav* 37(2):141–153. doi:[10.3758/LB.37.2](https://doi.org/10.3758/LB.37.2)
- Fiser J, Aslin RN (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci* 12:499–504
- Fiser J, Aslin RN (2005) Encoding multielement scenes: statistical learning of visual feature hierarchies. *J Exp Psychol Gen* 134:521–537
- Frishkoff GA (2007) Hemispheric differences in strong versus weak semantic priming: evidence from event-related brain potentials. *Brain Lang* 100(1):23–43
- Fusi S, Asaad WF, Miller EK, Wang XJ (2007) A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* 54:319–333. doi:[10.1016/j.neuron.2007.03.017](https://doi.org/10.1016/j.neuron.2007.03.017)
- Fuster JM (2001) The prefrontal cortex—an update: time is of the essence. *Neuron* 30(2):319–333
- Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652–654
- Gollwitzer PM, Kinney RF (1989) Effects of deliberative and implemental mind-sets on illusion of control. *J Pers Soc Psychol* 56:531–542
- Govindarajan A, Israely I, Huang SY, Tonegawa S (2011) The dendritic branch is the preferred integrative unit for protein synthesis-dependent LTP. *Neuron* 69(1):132–146
- Haarmann H, Usher M (2001) Maintenance of semantic information in capacity-limited item short-term memory. *Psychon Bull Rev* 8(3):568–578
- Hagoort P, Hald L, Bastiaansen M, Petersson KM (2004) Integration of word meaning and world knowledge in language comprehension. *Science* 304(5669):438–441
- Hald LA, Steenbeek-Planting EG, Hagoort P (2007) The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Res* 1146:210–218
- Hebb DO (1949) *The organization of behavior: a neuropsychological theory*. Wiley, New York
- Hill H, Strube M, Roesch-Ely D, Weisbrod M (2002) Automatic vs. controlled processes in semantic priming—differentiation by event-related potentials. *Int J Psychophysiol* 44:197–218
- Hutchison KA (2003) Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychon Bull Rev* 10:785–813
- Hutchinson A, Whitman RD, Abeare C, Raiter J (2003) The unification of mind: integration of hemispheric semantic processing. *Brain Lang* 87(3):361–368
- Kamide Y, Altmann G, Haywood L (2003) The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements. *J Mem Lang* 49(1):133–156
- Khalkhali S, Wammes J, McRae K (2012) Integrating words that refer to typical sequences of events. *Can J Exp Psychol* 66(2):106–114. doi:[10.1037/a0027369](https://doi.org/10.1037/a0027369)
- Kiefer M (2007) Top-down modulation of unconscious “automatic” processes: a gating framework. *Adv Cogn Psychol* 3:289–306
- Kiefer M, Martens U (2010) Attentional sensitization of unconscious cognition: task sets modulate subsequent masked semantic priming. *J Exp Psychol General* 139:464–489
- Kiefer M, Ahlegian M, Spitzer M (2005) Working memory capacity, indirect semantic priming, and stroop interference: pattern of interindividual prefrontal performance differences in healthy volunteers. *Neuropsychology* 19(3):332–344
- Kutas M, DeLong KA, Smith NJ (2011) A look around at what lies ahead: prediction and predictability in language processing. In: Bar M (ed) *Predictions in the brain: using our past to generate a future*. Oxford Scholarship Online, Oxford, pp 190–207. doi:[10.1093/acprof:oso/9780195395518.003.0065](https://doi.org/10.1093/acprof:oso/9780195395518.003.0065)
- Landauer TK, Foltz PW, Laham D (1998) An introduction to latent semantic analysis. *Discourse Process* 25:259–284
- Lavigne F (2004) AIM networks : autoincursive memory networks for anticipation toward learned goals. *Int J Comput Anticip Syst* 8:74–95
- Lavigne F, Darmon N (2008) Dopaminergic neuromodulation of semantic priming in a cortical network model. *Neuropsychologia* 46(13):3074–3087
- Lavigne F, Denis S (2001) Attentional and semantic anticipations in recurrent neural networks. *Int J Comput Anticip Syst* 14:196–214
- Lavigne F, Denis S (2002) Neural network modeling of learning of contextual constraints on adaptive anticipations. *Int J Comput Anticip Syst* 12:253–268
- Lavigne F, Vitu F (1997) Time course of activatory and inhibitory semantic priming effects in visual word recognition. *Int J Psycholinguist* 13(3):311–349
- Lavigne F, Dumercy L, Darmon N (2011) Determinants of multiple semantic priming: a meta-analysis and spike frequency adaptive model of a cortical network. *J Cogn Neurosci* 23(6):1447–1474
- Lavigne F, Dumercy L, Chanquoy L, Mercier B, Vitu-Thibault F (2012) Dynamics of the semantic priming shift: behavioral experiments and cortical network model. *Cogn Neurodyn* 6(6):467–483
- Lavigne F, Chanquoy L, Dumercy L, Vitu F (2013) Early dynamics of the semantic priming shift. *Adv Cogn Psychol* 9(1):1–14
- Lavigne F, Avnaïm F, Dumercy L (2014) Inter-synaptic learning of combination rules in a cortical network model. *Front Psychol* 5:842. doi:[10.3389/fpsyg.2014.00842](https://doi.org/10.3389/fpsyg.2014.00842)
- Lavzin M, Rapoport S, Polsky A, Garion L, Schiller J (2012) Nonlinear dendritic processing determines angular tuning of barrel cortex neurons in vivo. *Nature* 490(7420):397–401. doi:[10.1038/nature11451](https://doi.org/10.1038/nature11451)
- Lerner I, Shriki O (2014) Internally and externally driven network transitions as a basis for automatic and strategic processes in semantic priming: theory and experimental validation. *Front Psychol* 5:314. doi:[10.3389/fpsyg.2014.00314](https://doi.org/10.3389/fpsyg.2014.00314)
- Lerner I, Bentin S, Shriki O (2012) Spreading activation in an attractor network with latching dynamics: automatic semantic priming revisited. *Cogn Sci* 36(8):1339–1382. doi:[10.1111/cogs](https://doi.org/10.1111/cogs)
- London M, Hausser M (2005) Dendritic computation. *Annu Rev Neurosci* 28:503–532
- Longrée D, Mellet S (2013) Le motif : une unité phraséologique englobante? Etendre le champ de la phraséologie de la langue au discours. *Langages* 189:65–79
- Longrée D, Luong X, Mellet S (2008) Les motifs: un outil pour la caractérisation topologique des textes. In: *JADT 2008, Actes des 9èmes Journées internationales d'Analyse statistique des Données Textuelles*. Lyon: Presses de l'ENS, vol 2, pp 733–744

- Luka BJ, Van Petten C (2014) Prospective and retrospective semantic processing: prediction, time, and relationship strength in event-related potentials. *Brain Lang* 135:115–129
- Masson MEJ (1995) A distributed memory model of semantic priming. *J Exp Psychol Learn Mem Cogn* 21(1):3–23
- Masson MEJ, Besner D, Humphreys GW (1991) A distributed memory model of context effects in word identification. Lawrence Erlbaum Associates Inc, Hillsdale
- McNamara TP (1992) Theories of priming I: associative distance and Lag. *J Exp Psychol Learn Mem Cogn* 8(6):1173–1190
- McNamara TP, Altarriba J (1988) Depth of spreading activation revisited: semantic mediated priming occurs in lexical decisions. *J Mem Lang* 27(5):545–559
- McRae K, de Sa VR, Seidenberg MS (1997) On the nature and scope of featural representations of word meaning. *J Exp Psychol: Gen* 126:99–130
- McRae K, Ross BH (2004) Semantic memory: some insights from feature-based connectionist attractor networks. Elsevier Academic Press, San Diego
- McRae K, Cree GS, Seidenberg MS, McNorgan C (2005) Semantic feature production norms for a large set of living and nonliving things. *Behav Res Methods* 37:547–559
- Mellet S, Longrée D (2012) Légitimité d'une unité textométrique: le motif. In: JADT 2012, 11èmes Journées internationales d'Analyse statistique des Données Textuelles
- Meyer DE, Schvaneveldt RW, Rudy MG (1972) Activation of lexical memory. Meeting of the Psychonomic Society, St Louis
- Miyashita Y (1988) Neural discharges related to pictorial memory in the primate temporal cortex. *Neurosci Res Suppl* 7:S16
- Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331(6151):68–70
- Mizraji E, Lin J (2015) Modeling spatial-temporal operations with context-dependent associative memories. *Cogn Neurodyn* 9(5):523–534. doi:10.1007/s11571-015-9343-3
- Mizraji E, Pomi A, Valle-Lisboa JC (2009) Dynamic searching in the brain. *Cogn Neurodyn* 3(401):414
- Mongillo G, Amit DJ, Brunel N (2003) Retrospective and prospective persistent activity induced by Hebbian learning in a recurrent cortical network. *Eur J Neurosci* 18(7):2011–2024
- Moritz S, Woodward TS, Kupperts D, Lausen A, Schickel M (2003) Increased automatic spreading of activation in thought-disordered schizophrenic patients. *Schizophr Res* 59:181–186
- Moss HE, Hare ML, Day P, Tyler LK (1994) A distributed memory model of the associative boost in semantic priming. *Connect Sci* 6(4):413–427
- Naya Y, Yoshida M, Miyashita Y (2001) Backward spreading of memory-retrieval signal in the primate temporal cortex. *Science* 291(5504):661–664
- Naya Y, Yoshida M, Miyashita Y (2003) Forward processing of long-term associative memory in monkey inferotemporal cortex. *J Neurosci* 23(7):2861–2871
- Neely JH (1991) Semantic priming effects in visual word recognition: a selective review of current findings and theories. In: Besner D, Humphreys GW (eds) Basic processes in reading: visual word recognition. Lawrence Erlbaum Associates Inc, Oxford, pp 264–336
- Nelson DL, McEvoy CL, Schreiber T (1999) University of South Florida word association, rhyme and word fragment norms. Retrieved 8 Nov 2008, from <http://cyber.acomp.usf.edu/FreeAssociation/>
- Plaut DC (1995) Semantic and associative priming in a distributed attractor network. Paper presented at the 17th annual conference of the cognitive science society, Pittsburgh
- Plaut DC, Booth JR (2000) Individual and developmental differences in semantic priming: empirical and computational support for a single-mechanism account of Lexical processing. *Psychol Rev* 107(4):786–823
- Poggio T, Edelman S (1990) A network that learns to recognize 3D objects. *Nature* 343:263–266
- Rainer G, Rao SC, Miller EK (1999) Prospective coding for objects in primate prefrontal cortex. *J Neurosci* 19(13):5493–5505
- Randall B, Moss HE, Rodd JM, Greer M, Tyler LK (2004) Distinctiveness and correlation in conceptual structure: behavioral and computational studies. *J Exp Psychol Learn Mem Cogn* 30(2):393–406
- Renart A, Moreno R, de la Rocha J, Parga N, Rolls ET (2001) A model of the IT-PF network in object working memory which includes balanced persistent activity and tuned inhibition. *Neurocomputing* 38–40:1525–1531
- Ricciardi LM (1977) Diffusion processes and related topics on biology. Springer, Berlin
- Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019–1025. doi:10.1038/14819
- Rigotti M, Ben Dayan Rubin D, Wang XJ, Fusi S (2010a) Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front Comput Neurosci* 4:24
- Rigotti M, Ben Dayan Rubin D, Morrison SE, Salzman CD, Fusi S (2010b) Attractor concretion as a mechanism for the formation of context representations. *Neuroimage* 52(3):833–847
- Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497(7451):585–590
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22(21):9475–9489
- Romani S, Amit DJ, Mongillo G (2006) Mean-field analysis of selective persistent activity in presence of short-term synaptic depression. *J Comput Neurosci* 20:201–217
- Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. *Nature* 354(6349):152–155
- Salem A (1986) Segments répétés et analyse statistique des données textuelles. *Histoire et mesure* 1(2):5–28
- Schvaneveldt RW, Meyer DE (1973) Retrieval and comparison processes in semantic memory. In: Kornblum S (ed) Attention and performance IV. Academic Press, New York
- Sinclair J (1991) Corpus, concordance, collocation. Oxford University Press, Oxford
- Spence DP, Owens KC (1990) Lexical co-occurrence and association strength. *J Psycholing Res* 19:317–330
- Spitzer M, Braun U, Maier S, Hermle L, Maher BA (1993) Indirect semantic priming in schizophrenic patients. *Schizophr Res* 11:71–80
- Spruston N (2008) Pyramidal neurons: dendritic structure and synaptic integration. *Nat Rev Neurosci* 9:206–221
- Spruyt A, De Houwer J, Hermans D (2009) Modulation of automatic semantic priming by feature-specific attention allocation. *J Memory Lang* 61:37–54
- Tuckwell HC (1988) Introduction to theoretical neurobiology. Cambridge University Press, Cambridge
- Van Petten C (2014) Examining the N400 semantic context effect item-by-item: relationship to corpus-based measures of word co-occurrence. *Int J Psychophysiol* 94:407–419
- Vergauwe E, Cowan N (2015) A common short-term memory retrieval rate may describe many cognitive procedures. *Front Hum Neurosci* 7(8):126. doi:10.3389/fnhum.2014.00126
- Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36(5):955–968
- Weisbrod M, Maier S, Harig S, Himmelsbach U, Spitzer M (1998) Lateralised semantic and indirect semantic priming effects in people with schizophrenia. *Br J Psychiatry* 172:142–146

- Whitney C, Grossman M, Kircher TT (2009) The influence of multiple primes on bottom-up and top-down regulation during meaning retrieval: evidence for 2 distinct neural networks. *Cereb Cortex* 19(11):2548–2560
- Whitney C, Jefferies E, Kircher T (2010) Heterogeneity of the left temporal lobe in semantic representation and control: priming multiple versus single meanings of ambiguous words. *Cereb Cortex* 21(4):831–844
- Willems RM, Frank SL, Nijhof AD, Hagoort P, van den Bosch A (2016) Prediction during natural language comprehension. *Cereb Cortex* 26(6):2506–2516
- Wong KF, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* 26(4):1314–1328
- Xiumin L (2014) Signal integration on the dendrites of a pyramidal neuron model. *Cogn Neurodyn* 8(1):81–85. doi:[10.1007/s11571-013-9252-2](https://doi.org/10.1007/s11571-013-9252-2)