Available online at www.sciencedirect.com
SCIENCE DIRECT
Neural Networks 18 (2005) 863–877
Neural Networks
www.elsevier.com/locate/neunet

Elman topology with sigma–pi units: An application to the modeling of verbal hallucinations in schizophrenia

Juan C. Valle-Lisboa[a,*], Florencia Reali[a], Héctor Anastasía[b], Eduardo Mizraji[a]

[a]*Sección Biofísica, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay*
[b]*Servicio de Psiquiatría, Hospital Militar, Montevideo, Uruguay*

Received 5 April 2004; revised 2 March 2005; accepted 2 March 2005

Abstract

The development of neural network models has greatly enhanced the comprehension of cognitive phenomena. Here, we show that models using multiplicative processing of inputs are both powerful and simple to train and understand. We believe they are valuable tools for cognitive explorations. Our model can be viewed as a subclass of networks built on sigma–pi units and we show how to derive the Kronecker product representation from the classical sigma–pi unit. We also show how the connectivity requirements of the Kronecker product can be relaxed considering statistical arguments. We use the multiplicative network to implement what we call an Elman topology, that is, a simple recurrent network (SRN) that supports aspects of language processing. As an application, we model the appearance of hallucinated voices after network damage, and show that we can reproduce results previously obtained with SRNs concerning the pathology of schizophrenia.

Keywords: Sigma–pi unit; SRN; Kronecker product; Language processing; Schizophrenia

1. Introduction

One of the aims of the present-day neural network theory is to understand and represent the dynamics of cognitive processes. Surely, these dynamics arise as a consequence of complex interactions among diverse neural modules, each of them showing particular architectures suited for their tasks. In order to understand such systems, many approaches can be undertaken, ranging from detailed descriptions of neuronal properties to abstract symbolic models, including most of the connectionist neural network models.

The great variety of existing neural network models prompts us to analyze the possible similarities and differences between them with respect to both physiological and pathological functioning. Superficially distinct models might nevertheless share a common underlying principle. The present-day knowledge on mental pathologies provides exceptional factual data. Consequently, the consideration of pathology is central in order to refine or even discard inappropriate models. In return, models might point to a better understanding (and hopefully treatment) of pathologies.

In the present study, we explore the dynamical capabilities of a class of neural networks through simulations of language production and processing. We will focus on a particular recurrent network configuration, similar to the simple recurrent network (SRN) model developed by Elman (1990, 1995). Elman used an SRN as an illustration of the possibilities of the connectionist approach to language. In the Elman model of language, the input layer of the network receives phonetic input, the output layer gives the conceptual interpretation of the input, and there is a single hidden layer which receives its input both from the phonetic layer and from a working memory, which is just the output of the hidden layer itself in previous time. This model network, though simple, has inspired many discussions and applications specially to simulate linguistic processing (Elman, 1990, 1995), and has been particularly successful in human data fitting (Christiansen & Chater, 1999). A group of recent and relevant applications of SRNs comes from psychiatry, where Elman approach has

* Corresponding author. Tel.: +598 252586187139; fax: +598 25258629.
E-mail address: juancvl@fcien.edu.uy (J.C. Valle-Lisboa).

0893-6080/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2005.03.009

been used as a base for the integration of a variety of data (developmental, genetic and pharmacological) on the pathology of schizophrenia (Cohen & Servan-Schreiber, 1992; Hoffman, 1987; Hoffman & McGlashan, 1997, 1998; Hoffman et al., 1995; McGlashan & Hoffman, 2000). Thus, although the SRN has been used by many authors, we will refer to a network having an input module, an output module and a recursive working memory module as having *Elman topology*. In the original Elman model, the implementation of the network was based on the well-known theory of multilayer perceptrons, trained using the backpropagation algorithm (Rumelhart, Hinton, & Williams, 1986). The purpose of the present work is to show that some of the properties and applications of Elman model persist when we employ a network constructed with a particular type of sigma–pi unit (Rumelhart, Hinton, & McClelland, 1986). In the following, we will abbreviate sigma–pi Elman topology model as SPELT model.

The sigma–pi unit responds both to the weighted sum of its inputs and also to the weighted sum of the product of its inputs. There is neurobiological evidence for the presence of multiplicative synapses relevant for sensory processing (Peña & Konishi, 2001; see also Koch & Segev, 2000) which enhance the biological plausibility of models based on multiplication. Recently, the computational power of multiplicative processes has received much attention both from a computational complexity perspective (Schmitt, 2002) and to expand the generalization capabilities of traditional connectionist models (Neville & Elridge, 2002).

Here, we present a model built on a particular type of sigma–pi unit, using a simple network architecture that has a simple algebraic representation. We discuss the biological plausibility of this model as well as its computational properties. We show that the use of that type of units allows the network to be trained using very simple procedures, such as the Widrow–Hoff delta rule (Widrow & Hoff, 1960) in a very quick and efficient manner. In our experience, the use of the multiplicative model reduces the complexity of the learning phase, a finding that has been also put forth by other authors (Pao, 1989).

In the last part of the present study, we explore the applicability of the SPELT model. For that purpose we test the model's predictions on a linguistic task, similar to those successfully modeled using SRNs. Particularly, inspired by the simulation of psychiatric disorders performed by McGlashan and Hoffman (2000), Hoffman and McGlashan (1997), and Hoffman, Rapaport, Ameli, McGlashan, Harcherik, and Servan-Schreiber (1995), we show that it is possible to model hallucinated voices in schizophrenia using our alternative model. This work has to be regarded as a technical modification aimed to compare our network to an SRN-based simulation of pathology (Cohen & Servan-Schreiber, 1992; Hoffman et al., 1995). In this respect, we show that many of the conclusions reported using an SRN to study hallucinated voices persist in our SPELT model. This is noteworthy given the simpler overall architecture (since we use no hidden layers) and learning dynamics in SPELT. We also discuss some of the important differences that we found with previous works.

2. Sigma–pi units and the Kronecker product

2.1. Theory

In the traditional PDP approach (Rumelhart et al., 1986a, b), the most common type of unit uses a sigmoid activation function and a threshold output function. The net input to each unit is the weighted sum of the activation value of the input units. In contrast, a sigma–pi unit is a formal neuron having a real-valued activation a_i, which (in our case) is updated in discrete time, as follows

$$a_i(t+1) = \sigma_i \left[\sum_{j=1}^{N} M_{ij} \prod_{\alpha_{ij}} b(\alpha_{ij}; t) \right] \quad (1)$$

where σ_i is an activation function, and it is assumed that neuron i has N synaptic coefficients and j indexes a set of multiplicative inputs b (α_{ij}); $\alpha_{ij} \in Q_{ij}$, Q_{ij} being the set of sub-indexes that correspond to the inputs to block j of neuron i. As an example, with $N=4$ and $Q_{i1}=\{m, p, s\}$, $Q_{i2}=\{p, q\}$, $Q_{i3}=\{n\}$, $Q_{i4}=\phi$ (see Fig. 1).

For the sake of simplicity we are going to assume that σ_i is a linear function, and can be omitted; we also assume that the output is the activation value itself and we will not use the bias in the equations until the simulations of Section 4.

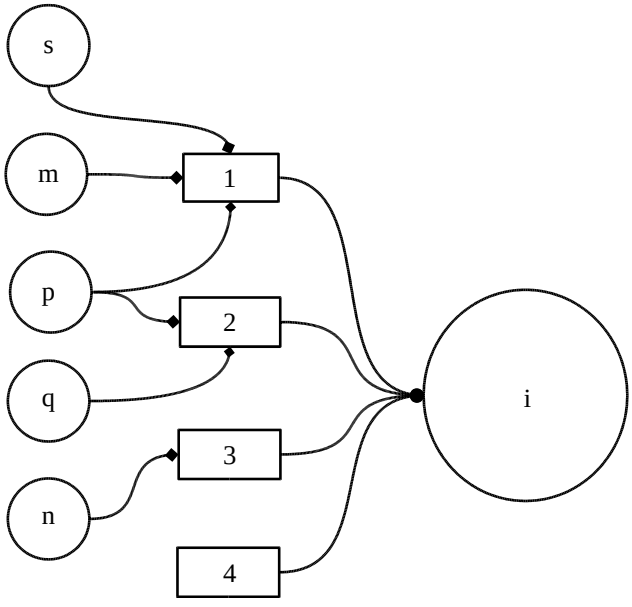

Fig. 1. A schematic view of the connectivity used as the example described in the text. Neuron 'i' evaluates the weighted sum of the results of multiplication performed by the blocks (rectangles) 1–4. In this particular example block 4 receives no input. The other blocks perform the multiplication of inputs coming from neurons s, m, p, q and n. A weight is assigned to each block.

These assumptions should be regarded as useful approximations only. We discuss their biological meaning in Section 4. Then, in the case of Fig. 1 the updating equation for neuron i is

$$a(i;t+1) = M_{i1}b(m;t)b(p;t)b(s;t) + M_{i2}b(q;t)b(p;t) + M_{i3}b(n;t)$$

Note that in above equation the term M_{i4} is absent due to the absence of the corresponding input to the synapse.

We restrict ourselves to second order sigma–pi units. Moreover, we assume that the information affecting multiplicative synapses comes from disjoint neural sets. One set of signals, represented by vectors f_i, comes from a network Net_1 with m units and another set of signals, p_j, proceeds from Net_2 with n units. The separation of the multiplying inputs into two sets allows for a simple and clear algebraic format, which we describe below. Although this separation is a simplifying assumption, note that it is only required at the level of the multiplying synapses; the input cells can still be affected by common top-down or recurrent projections. This assumption can be applied whenever there are interacting but more or less separate neural modules. For instance, each network Net_1 and Net_2 could process information from different sensory modalities. In any case, the vectors f and p represent the activity coming from different nets.

The second order sigma–pi unit associated to those data banks can be represented as follows

$$a(i;t+1) = \sum_j M_{ij} f(\alpha_{ij};t)p(\beta_{ij};t). \tag{2}$$

In this expression, α_{ij} indexes the input f coming from Net_1 and affecting M_{ij}, whereas β_{ij} refers to the input p from Net_2 which affects synapse M_{ij}. Let us illustrate Eq. (2) with an arbitrary example. Suppose that j goes from 1 to 3 and the networks Net_1 and Net_2 have dimensions $m=4$ and $n=6$, then one possible configuration of inputs is

$$\alpha_{i1}=2;\ \ \beta_{i1}=1;\ \ \alpha_{i2}=1;\ \ \beta_{i2}=5;\ \ \alpha_{i3}=3;\ \ \beta_{i3}=2.$$

Then the output of the neuron can be calculated as

$$a(i;t+1) = M_{i1}f(2;t)p(1;t) + M_{i2}f(1;t)p(5;t) + M_{i3}f(3;t)p(2;t).$$

The latter representation is readily generalizable to one in which all the pairs of possible products are included. In order to clearly indicate the different origins of the signals (i.e. the different banks where the signals come from) we replace the sub-index j of Eq. (2) with the corresponding pair $\alpha_{ij}\ \beta_{ij}$ in a way that assigns the correct α and β to each j. To represent the same data with both formulas (i.e. Eqs. (2) and (3)), we require that if a pair $\alpha\beta$ does not exist, its coefficient $M_{i\alpha\beta}$ equals 0. With these transformations the general case can be described as

$$a(i;t+1) = \sum_{\alpha=1}^{m}\sum_{\beta=1}^{n} M_{i\alpha\beta} f(\alpha;t)p(\beta;t). \tag{3}$$

The structure defined by Eq. (3) can be represented in terms of matrix algebra by the Kronecker product as we show below (in the following, we assume that inputs in time t produce outputs in time $t+1$, so we can drop the variable t).

Defining the vector of synaptic coefficients

$$M(i) = [\,M_{i11}\quad M_{i12}\quad \ldots\quad M_{imn}\,]^T,$$

and the signal vectors

$$\boldsymbol{f} = [f(1)\ \ f(2)\ \ \ldots\ \ f(m)]^{\mathrm{T}},\ \ \boldsymbol{p} = [p(1)\ \ p(2)\ \ \ldots\ \ p(n)]^{\mathrm{T}},$$

the output of neuron i evoked by the vectorial pattern f and the vectorial context p is given by

$$a(i) = M(i)^{\mathrm{T}}(f \otimes p), \tag{4}$$

where $\mathbf{f}\otimes\mathbf{p}$ denotes the Kronecker product (Bellman, 1960). The Kronecker product of arbitrary matrices $A=[a_{ij}]$ of size $m\times n$ and $B=[b_{ij}]$ of size $p\times q$ can be defined as

$$A\otimes B = [a_{ij}B],$$

the product being a matrix of size $mp\times qn$.

An important property of the Kronecker product for conformable matrices A, C and B, D is $(A\otimes B)(C\otimes D)=(AC)\otimes(BD)$. For vectors a, b, c and d this property implies that $(a^{\mathrm{T}}\otimes b^{\mathrm{T}})(c\otimes d) = a^{\mathrm{T}}c\otimes b^{\mathrm{T}}d = \langle a,c\rangle\langle b,d\rangle$.

The formal advantages of the Kronecker product when trying to represent the connections between input patterns and their contexts are illustrated in the following situation: an elementary associative memory M which associates orthonormal patterns and contexts f and p with outputs g has the following structure

$$M = \sum_s\sum_v g_{sv}(f_s\otimes p_{sv})^{\mathrm{T}}, \tag{5}$$

where g_{sv} and f_s are patterns associated under different contexts p_{sv}.

If this memory receives the input $f\otimes p$ the system produces a double filtering by scalar products:

$$M(f\otimes p) = \sum_s\sum_v g_{sv}(f_s\otimes p_{sv})^{\mathrm{T}}(f\otimes p) = \sum_s\sum_v g_{sv}\langle f_s,f\rangle\langle p_{sv},p\rangle. \tag{6}$$

The double filtering (Eq. (6)) significantly increases the computational capabilities of these memories. In particular, they allow for an immediate implementation of the XOR gate (Mizraji, 1989) not requiring hidden layers and, as a consequence, amenable to simpler learning paradigms.

2.2. The statistical realization of the Kronecker product under sparse connectivity

From a biological point of view, the Kronecker product in its detailed form is highly improbable, because it imposes a high requirement of neuroanatomical regularities. The purpose of this section is to show that a non-fully connected network can still compute the relevant context-dependent associations, provided that the network is relatively large This can be shown as follows.

According to Eq. (5) the output of a neuron α (that is $g'(\alpha)$) to input f and context p, can be written as

$$g'(\alpha) = g(\alpha)h^{\mathrm{T}}h + \sum_{i=1}^{K} g_i(\alpha)h_i^{\mathrm{T}}h, \tag{7}$$

where $h = f \otimes p$ for the particular input and $h_i = f_i \otimes p_i$ are the inputs associated to vector g_i in the memory (the memory depicted by Eq. (7) stores $K+1$ associations).

In what follows, the dimension of the space of f vectors is m and the dimension of the space of vectors p is n. We suppose that the network has been created with an imperfect synaptic connectivity, at least with respect to the Kronecker product, which means that what is really functioning is a subset of the 'ideal' network. To implement this idea, we use the rarefaction operator D_α that is a diagonal matrix, where a fraction ϕ of diagonal elements equals zero (see Mizraji, Pomi, & Alvarez, 1994). Since each neuron receives all the inputs from the two banks and performs 'its own' Kronecker product, we use a different rarefaction operator for each neuron α of the output bank (see Pomi & Mizraji, 1999). Using this operator, the output of neuron α in the incomplete network is

$$g^0(\alpha) = g(\alpha)h^{\mathrm{T}}D_\alpha h + \sum_{i=1}^{K} g_i(\alpha)h_i^{\mathrm{T}}D_\alpha h,$$

which implies the assumption that each neuron receives an incomplete Kronecker product of input and context vectors. In the last expression, we denote the output from the imperfect network as g^0.

To obtain the conditions under which g^0 is a good approximation to g we calculate the usual correlation between them as $r \equiv \langle g, g^0\rangle / |g||g^0||$.

To simplify the argument we shall restrict ourselves to orthonormal sets of inputs, contexts and outputs. With the orthonormality hypothesis the norm is $\|g\| = 1$ and can be omitted. To simplify the notation we use the following definitions:

$$u(\alpha) \equiv h^{\mathrm{T}}D_\alpha h, \quad v_i(\alpha) \equiv h_i^{\mathrm{T}}D_\alpha h,$$

We will assume that $v_i(\alpha) = v_i$ and $u(\alpha) = u$. The assumption that u and v are independent of the output neuron requires that each h that enters the output layer be filtered in a similar way by each neuron. For each different input, the detailed numbers will be different, but this assumption can be seen as a statistical approximation. This leads directly to the formula

$$\langle g, g^0\rangle = ||g||^2 u + \sum_{i=1}^{K} v_i \langle g_i, g\rangle = u,$$

and

$$||g^0|| = \sqrt{\langle g^0, g^0\rangle} = \sqrt{u^2 + \sum_{i=1}^{K} v_i^2}.$$

Finally

$$r = \frac{u}{\sqrt{u^2 + \sum_{i=1}^{K} v_i^2}} \tag{8}$$

To link the result of Eq. (8) to the fraction of terms deleted in the product (i.e. ϕ) we note that u is the square of the norm of h but calculated after the elimination of ϕJ terms each of them with expected value $J^{-1/2}$, being $J = mn$, the dimensions of vectors h. Thus

$$u = 1 - \phi.$$

The same type of argument yields

$$\sum_{i=1}^{K} v_i^2 \approx K\phi\frac{(1-\phi)}{J-1},$$

so finally

$$r = \frac{1-\phi}{\sqrt{(1-\phi)^2 + K\phi\frac{(1-\phi)}{J-1}}} = \frac{1}{\sqrt{1 + \frac{K}{J-1}\frac{\phi}{1-\phi}}}. \tag{9}$$

As an illustration, we plot Eq. (9) (r vs. ϕ) for different values of K and J in Fig. 2. A more detailed analysis of the applicability and consequences of Eq. (9) is presented (in a different context) in a previous work (Mizraji et al., 1994).

Thus, it can be seen that provided that the dimensions of the neural network are higher than the number of stored patterns, the rarefied network can still compute an approximate Kronecker product. Although the calculation can be regarded as being general, the performance can be sensitive to particular values of the parameters, and specially to the encoding employed in the input, output and context (see Pomi & Mizraji, 1999).

2.3. Learning and an application to the XOR problem

In order to compare the learning dynamics of a single-layered multiplicative network and a two-layered perceptron, we performed preliminary simulations of computation of the logical gate exclusive-or (XOR). In this section, we use part of the preliminary results obtained by Reali (2002). A more detailed analysis of the learning capabilities of the multiplicative model with respect to logical gates and vocabularies is in preparation.

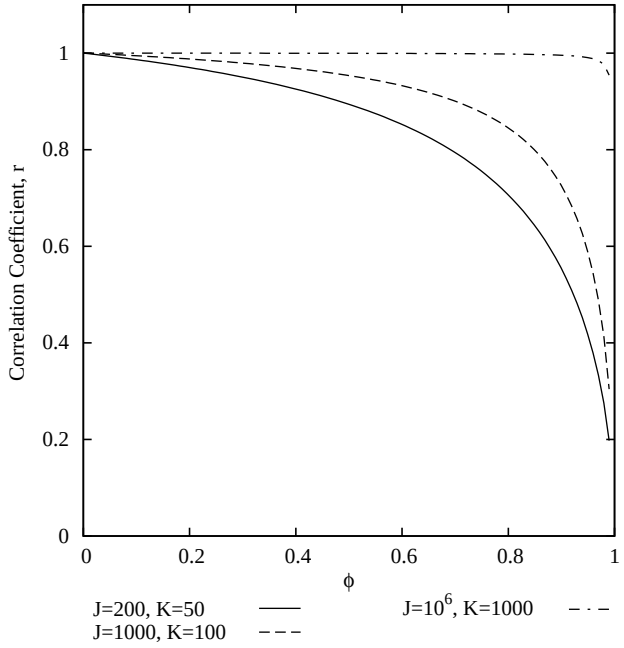

Fig. 2. Theoretical correlation coefficient between a fully connected multiplicative network and a rarefied one as a function of the fraction of disconnections (ϕ) as implied by Eq. (9).

We built a multiplicative feedforward network, which has two input layers with two neurons in each and an output layer made of two neurons. The truth-values were coded in two-dimensional vectors and the output neurons received the Kronecker product of both input vectors. This network was trained using the delta rule (Widrow & Hoff, 1960), which for the multiplicative model with inputs h and p is

$$\delta \mathbf{M}^i = 2\alpha e(i)(h \otimes p)^{\mathrm{T}}, \tag{10}$$

where $\mathbf{M}^i$ is the ith row of the weight matrix, $e(i)$ is the error in the output (expected minus obtained) of the ith neuron and α is a learning constant. During each epoch we presented the four patterns of possible truth-values, correcting the weights with the aid of Eq. (10) after each presentation. The mean square error for each epoch shown in Fig. 3 is an average of the results of the presentation of the four entries.

The two-layered network was built on traditional units (with linear summation of inputs and sigmoid activation functions). This network had four input units, four hidden layer units and two output units. The input was a four-dimensional vector. The training algorithm used was a variation of backpropagation (Rumelhart et al., 1986a,b) as described in Boers and Kuiper (1992) including a momentum term to improve convergence while minimizing oscillations

$$\mathbf{W}(t+1) = \mathbf{W}(t) - \left[\eta \frac{\partial E}{\partial \mathbf{W}} + \beta \Delta \mathbf{W}(t-1)\right] \tag{11}$$

where $\mathbf{W}$ refers to the weight vector of each neuron, η is the learning rate, E is the error and β is the momentum term that

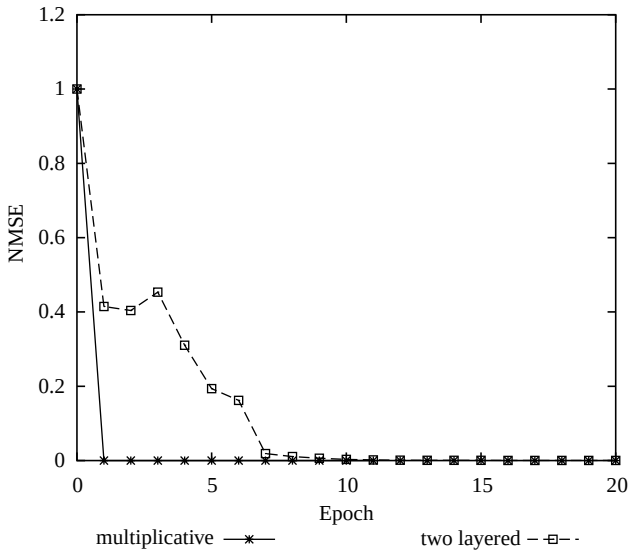

Fig. 3. Normalized mean square error (NMSE) vs. epoch for a multiplicative network and a two-layered network, learning the XOR function. Starting from random weights, the multiplicative network learns the task in one epoch, i.e. after the presentation of the four associations needed to compute the logical function. The backpropagation procedure takes more than 15 epochs to achieve the same NMSE. Parameters used: for the delta rule (Eq. (10)) learning rate $\alpha = 1/8$ Backpropagation procedure: we used Eq. (11) with the learning rate $\eta = 0.5$ and the momentum constant $\beta = 0.9$.

multiplies the change in weights from the previous time $(t-1)$ to their value in time t. We measured the squared error in each epoch in the output layer in the same way as we did with the multiplicative network (Reali, 2002). An example of the performance of both networks is shown in Fig. 3.

Fig. 3 shows that for the particular task the multiplicative model learns faster than the two-layered network. The same observation has been previously reported by other authors, using different multiplicative models (see for example Pao, 1989). Changing the encoding scheme of the truth-values or the dimensions of the vectors does not change the general conclusion, namely, that training the multiplicative network is faster than the backpropagation procedure (Reali, 2002; and see below for the learning of a vocabulary).

3. The SPELT and its application to the simulation of hallucinated 'voices' in schizophrenia

3.1. The sigma–pi Elman topology model

We turn now to the SPELT model. Our version of the SPELT model is shown in Fig. 4. The presence of multiplying units enables the implementation of a model with Elman topology without the need of hidden layers. Phonetic information is represented by the activity of the input layer, so a word enters the network as a 'phonetic' input activity. The output activity is considered as the 'concept' associated to that phonetic information (Elman, 1995; Hoffman & McGlashan, 1997; Hoffman et al., 1995).

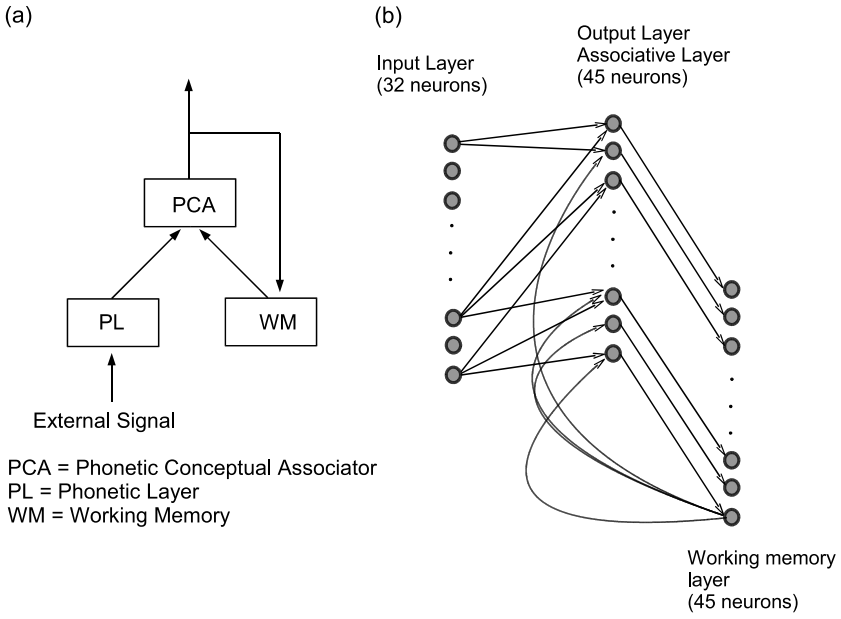

Fig. 4. (a) Topological diagram of the SPELT model. The working memory activity acts as a context to the arriving phonetic information. In the model analyzed here, the working memory activity at time t is the previous output (i.e. at time $t-1$) from the phonetic–conceptual associator (PCA). (b) Multiplicative network able to associate the *concept* pattern (output layer activity) to the *phonetic* input (input layer activity) and its context (working memory layer activity). We use 32 neurons in the input layer and 45 neurons in the output layer.

The output activity is projected to the working memory layer, which, in turn, produces an activity pattern in response to that activity. Each working memory activity pattern will represent the 'context' to the immediately following phonetic input arriving to the network. Each output layer neuron receives an input vector that results from the multiplicative preprocessing of a normalized phonetic input and its context, using the Kronecker product. In our simple formulation, the context for an input at time t equals the normalized output layer activity at time $t-1$ with the addition of a small bias, which introduces the possibility of small correlations between the conceptual output and the phonetic input. The need for a bias is imposed by the type of coding scheme we adopt at the output layer (following Hoffman et al., 1995) to perform the simulations to be shown below. We are aware that this is a very crude representation for a working memory, which is known to be able to store structured information (Baddeley, 1992) but its simplicity allows us to use the delta rule in training while keeping the properties of the original model, as shown in Section 2.3.

Since in this work we restrict ourselves to single-layered networks, they can be trained using just a simple gradient descent method, i.e. the delta rule. This is so because the network can be seen as a linear associator between the product of a pair of vectors ($\mathbf{f}\otimes\mathbf{p}$) and the output $\mathbf{g}$. In our experience, the SPELT network learned faster than the SRN trained using backpropagation (Reali, 2002). In the following paragraphs, we show that the SPELT model can be used to implement linguistic processing simulations and reproduce results of models applied to psychiatric disorders.

We considered it important to explore if previous reported results, from other computational models (Elman, 1990; McGlashan & Hoffman, 2000) were still present with the multiplicative neural architecture. In this section, we show that our model, based upon multiplicative procedures, is able to learn a definite vocabulary and use linguistic expectations to guide the recognition of words. We demonstrate the relevance of linguistic expectations by showing that word-recognizing abilities are stronger when the target word is preceded by a syntactically or semantically correct word (see below). We also show that the content of illusions depends on the previous inputs (Hoffman et al., 1995).

As a second objective, we tested the model's capacity to mimic schizophrenic disorders caused by the deterioration of network connectivity, as described by Hoffman et al. (1995). The importance of this test is twofold; in the first place it complements the considerations about sparse connectivity and the feasibility of the implementation of the Kronecker product made in Section 2.2 by showing the robustness of the SPELT under two 'anatomical manipulations' after learning (as opposed to the case before learning treated in Section 2.2); in the second place it shows comparatively how the SPELT model can be used as an alternative to Elman-type models of pathology and normal cognition.

3.2. Hallucinated voices as a result of excessive pruning of neural connections in the language recognition system

Schizophrenia is one of the most disabling psychiatric disorders and affects about 1% of the population

(Sawa & Snyder, 2002). Auditory hallucinations are among the most distressing symptoms of schizophrenia, reported in 50–80% of patients (Sartorius, Shapiro, & Jablensky, 1974; Shergill, Robin, Murray, & McGuire, 1998). There are many evidences suggesting that those hallucinations involve a deterioration of the neural structures responsible for speech processing (Shergill, Bullmore, Simmons, Murray, & McGuire, 2000). The neural network model developed by Hoffman & McGlashan (1997, 1998) and Hoffman et al. (1995) accounts for the generation of spontaneous speech percepts experienced as hallucinations arising from connectivity reductions. This reduction mimics the reduced synaptic connectedness on cortico-cortical neuronal circuits as a result of developmental disturbances of synaptogenesis and/or synaptic pruning during adolescence that has been postulated to occur in schizophrenia (McGlashan & Hoffman, 2000).

Similar to what Hoffman and co-workers did, we simulated a reduction in our model's connectivity and showed that hallucinated speech can appear. We used two different paradigms for network disconnection that show different patterns of behavior. We also simulated the effect of neuromodulatory disturbances on speech recognition.

3.3. Methods

The training and assessment methods used in this study were strongly based on those defined in the works of Hoffman and McGlashan (1998). As shown in Fig. 3, our SPELT model uses 32 units in the input layer and 45 in the output layer. In all the simulations described here we used a bias of 0.17. Recall that the bias is added to the output at time $t-1$ to form the context to the input arriving at time t.

Inspired upon the vocabulary defined by Ritter and Kohonen (1989), we used 28 words distributed in three classes: VERBS (in simple present, third person singular form: runs, walks, works, speaks, visits, phones, buy, sells, likes, hates, drinks, eats); NOUNS (Bob, Mary, dog, horse, beer, water, bread, meat); ADVERBS (much, little, fast, slowly, often, seldom, well, poorly). The vocabulary was superficially different from that used by Hoffman et al. (1995) but based upon the same ideas: (1) use a random binary coding in the input 'phonetic' layer and (2) employ a sparse coding at the output 'semantic and syntactic' layer. Thus, each word was represented as a particular binary random pattern at the input layer and was assigned a precise pattern within the output layer where three of the output neurons were turned on. These neurons coded for semantic and syntactic features. We choose the orthogonal coding at the output layer to be able to compare our model to SRNs. To allow the presence of correlation between different contexts we then introduced a bias in the context units.

We trained the network using 12 repetitions of 100 grammatically correct sentences, separated by blanks representing a silence. The correct grammatical structure is defined by the following rules: (1) each sentence is a string of three words; (2) the first word is a human or animal noun; (3) the second word is a verb allowed to follow the first noun (for example, the verb 'speaks' cannot follow the noun 'dog'); (4) the third word is an adverb or a noun which fits to the previous words. Due to the fact that we do not encode articles in the network the employed grammar is not English but a simpler one. For instance, two of the training phrases were 'Mary drinks water' and 'dog drinks water'. The set of phrases used for assessment were different from the training ones.

The training consists on the 'on-line' modification of the synaptic weights, using the delta rule (Widrow & Hoff, 1960).

Since the output activity is real-valued, to determine whether the network has perceived a word or not, we adopted the following criteria. First we normalized both output and target words. We then choose the word belonging to the training data set that had the smallest Euclidean distance to the output word as a candidate for recognition. If the differences between the Euclidean distance to the candidate word x and the distances to all other words are higher than a 'recognizing threshold' α, we consider that the network perceived the word x. In all other cases, we consider that there's no recognition (Hoffman & McGlashan, 1997; Hoffman et al., 1995). After some trial simulations, α was set to 0.3.

In order to classify the activity of the network, we followed the criteria established by Hoffman and McGlashan (1995, 1997). When the network recognizes a word that corresponds to the input word, we say that the network performs a 'successful identification'. If the recognized word does not correspond to the input word, the network performs a 'wrong identification', and if the output activation pattern demonstrates no clear-cut best fit, we say that the network performs 'no identification'. If the input pattern corresponds to the neutral vector defined as 'silence', and the network perceives a word, it is considered as a 'hallucinated word'. If the input pattern corresponds to some 'noisy input' and the network perceives a word it is defined as an 'illusion'. The noisy inputs are defined as random activation patterns, different from those coding the vocabulary words (see below).

In a first stage, we test the performance of the fully connected network. As a first test we present 200 sentences separated by five silences, and repeat the procedure 20 times with different seeds to the random number generator. We also perform a test with 900 words randomly presented (and 300 silences) and repeat it 20 times with different seeds. To assess the capacity of the network to rely upon linguistic expectations, we replace the correct words (nouns, verbs or adverbs) of a sentence by noisy blanks, following the procedures of Hoffman et al. (1995). Noisy blanks are obtained by setting each entry of the input vector to 1 or 0 with probability 0.5. We repeat this procedure changing the noisy blank position to test the expectations and their relation to the grammatical structure.

We determined the percentage of illusions, separating them in to three categories: (1) grammatically correct expectation, (2) grammatically wrong expectation with high presence during the training sentences (this category includes the nouns 'Bob', 'Mary', 'dog' and 'horse'), and (3) grammatically wrong expectation with low presence during the training sentences.

3.3.1. Neuroanatomical manipulation

In a second stage, we measured the performance of the network under conditions of connectivity reduction. We first simulated a pruning procedure guided by the concept of neurodevelopmental 'Darwinism', where the less robust synaptic connections are eliminated (Hoffman & McGlashan, 1997; Hoffman et al., 1995). This was accomplished by making zero all the weights in the associative memory whose absolute values were lower than a threshold. The percentage of pruning informed in Section 3.4 corresponds to the percentage of weights that are eliminated by this process. To obtain different levels of pruning we increased threshold systematically.

The second simulated situation was 'working memory functionality reduction' (WMFR). This was accomplished by making zero some randomly chosen (with fixed probability) entries of the context vector. The percentage of WMFR pruning corresponds to the percentage of the entries in the context vector that were made equal to zero.

3.3.2. Neuromodulatory simulation

The neuromodulatory action of antipsychotic drugs was simulated multiplying the components of the output vectors of the working memory by a numerical constant smaller than one (the bias remained unchanged). This procedure admits at least two different interpretations. On the one hand, it can be viewed as a relative reduction of the coded signals emitted by the working memory. On the other hand, this reduction can be interpreted as the result of a selective action at the level of the synaptic transmission that selectively affects the inputs coming from the working memory. We also perform a symmetric numerical experiment where the output of a working memory is enhanced.

We define a simulation set as the result of a training stage and a complete testing stage. The complete testing stage was performed by the presentation of 100 sentences for each level of neuroanatomical manipulation. We simulated a total of 20 simulation sets, using different random seeds during the training stage. Given the limited number of words and possible associations between them, the number of sentence presentations is limited, so the total evaluation described produced redundant information.

3.4. Results

We present here the results of our simulations, divided into three sections. In Section 3.4.1, we show that the network can use linguistic expectations to recognize words; in Section 3.4.2, we show that hallucinations appear as a result of synaptic destruction or working memory neuron elimination; finally, in Section 3.4.3, we show that the effects of neural damage can be ameliorated by parameter changes that mimic neuromodulatory actions.

3.4.1. Word recognition abilities and linguistic expectations

In the absence of neuroanatomic manipulations, we found that the SPELT network was able to successfully recognize the vocabulary when it was presented as novel sentences in the test set. As a mean result, considering the total number of word presentations, we found that the SPELT model was able to recognize 99.14% of the words presented within sentences, and there were no misidentifications. In line with the results of Hoffman and McGlashan (1997) when we presented the words in random order the recognition capabilities dropped but the intact network was still able to recognize 83.9% of the words. The difference between the recognition percentages of fixed and random order presentation is highly significant (with a test of proportions based on the normal distribution, $p<10^{-6}$).

It must be highlighted that such a successful performance was reached with only 12 presentations of 100 sentences during each training stage, which represents a high learning speed. That learning rate represents a significant improvement when compared with standard SRNs results (for comparison to an analogous set of simulations, see Hoffman & McGlashan, 1997).

To test the importance of linguistic expectations in word recognition by the SPELT network, we presented noisy blanks interspersed in 1000 test sentences and found the following results. When the expected word was a verb, the network produced 14% of illusions of verbs (i.e. when we presented the sentence 'Mary‖…noise…‖water' we get 'Mary‖ drinks ‖ water' or 'Mary‖ likes ‖ water'), 8% illusions of the more frequent words presented during the training (i.e. the nouns Mary, Bob, horse, dog), 75% of absences of recognition, and 3% of illusions of words not related to the expectations (and that where presented in a low frequency during training). Considering that in our database we used 43% of verbs, the latter results show that the difference was significant ($\chi^2=17.8$, $p<2.4\times10^{-5}$). When the expected word was an adverb or a noun but not a verb (i.e. when noise was presented in the third position), the network produced 17% of grammatically correct illusions, 82.5% of absences of recognition, and 0.5% of illusions of verbs ($\chi^2=114.3$, $p<10^{-6}$).

These results demonstrate the network's ability to guide its recognition by the use of linguistic expectations. We also found that for low percentages of connection's reduction, the network was able to successfully recognize a high percentage of words (Table 1, columns 1–4).

The network resistance to damage was different for the two kinds of simulated neuroanatomic manipulations.

Table 1
Simulation of pruning of the weakest connections

%Pruning	%Recognized words	%No identification	%Wrongly recognized words	No. of sets where hallucinations arose	Mean number of hallucinated words per set[a]
2.0	99.1	0.9	0.0	0	–
12.0	99.2	0.8	0.0	0	–
22.0	99.1	0.9	0.0	0	–
37.0	99.2	0.7	0.0	0	–
47.0	99.0	1.0	0.0	0	–
67.0	98.0	2.0	0.0	0	–
77.0	93.7	6.3	0.0	0	–
82.0	90.0	10.0	0.0	0	–
87.0	81.9	18.0	0.1	0	–
92.0	67.1	32.1	0.8	3	3
97.0	40.5	57.5	2.0	8	18

We eliminated the desired percentage (as shown in the first column of the table) of the smallest synaptic coefficients. Average results over 20 simulation sets. We presented 300 words within each set, corresponding to 100 sentences; the percentages are calculated over 300 words.

[a] This average number was calculated regarding only those simulation sets where hallucinated words appeared. Five hundred silences were presented to test the presence of hallucinations within each set.

The network presented a higher resistance to the selective pruning (shown in Table 1) than to the working memory functionality reduction (shown in Table 2). The number of mis-identifications increased when we simulated high levels of neuroanatomic manipulations, but remained marginal for low levels of simulated damage.

A graphical representation is shown in Fig. 5, where the robustness of the network for the two kinds of damage is compared. The trend for random order presentation is similar, with hallucinations appearing only when 92% of connections or more where selectively pruned or when 14% or more of working memory units where eliminated.

3.4.2. Hallucinated speech

When we simulated selective pruning (elimination of the weakest connections), hallucinated speech appeared in a fraction of the cases. Recall that each case is characterized by a particular random seed that defines the encoding of each word at the phonological level (i.e. at the input layer). The results are shown in Table 1 and Fig. 6. As can be seen, the hallucinated words appeared concomitantly to high levels of deterioration of the general performance.

For working memory functionality reduction we found the results shown in the fifth and sixth columns of Table 2. Fig. 6 shows the mean percentage of hallucinated words per silence averaged over the 20 simulation sets. Notice the different behavior of the model when confronted with the two disconnection paradigms. In both cases, for different simulation sets we found different hallucinated words, but within a particular simulation set the hallucinated word was the same, regardless of the particular place where the hallucination arose.

Only a fraction of the simulation sets produced hallucinated words, while others only suffered a reduction of their recognition capabilities but producing no hallucinations. The reason of this variability is that each set uses a different random seed, which determines the encoding of the phonetic input and the initial values of synaptic weights. The hallucinated words appeared in response to the second or third silence of the sentence, or following words that

Table 2
Working memory functionality reduction (WMFR)

%WMFR pruning	%Recognized words	%Not recognized words	%Wrong identifications	No. of sets where hallucinations arose	Mean number of hallucinations per set[a]
0.0	99.1	0.9	0.0	0	0
13.4	98.6	1.4	0.0	0	0
23.7	97.1	2.9	0.0	0	0
36.6	94.4	5.6	0.0	0	0
41.3	94.8	5.2	0.0	0	0
46.1	91.0	9.0	0.0	2	7.0
50.0	90.4	9.6	0.0	2	94.7
53.8	86.9	13.1	0.0	1	2.0
63.6	79.8	29.6	0.6	2	2.0
86.6	50.0	46.9	3.1	2	75.0

WMFR pruning stands for the mean fraction of cells in the working memory that have been destroyed. We used the same probability of destruction for each simulation set. Average results over 20 simulation sets (300 word presentations within each set). Test of hallucinations are given in columns 5 and 6.

[a] This average number was calculated regarding only those simulation sets where hallucinated words appeared. Five hundred silences were presented to test the presence of hallucinations within each set.

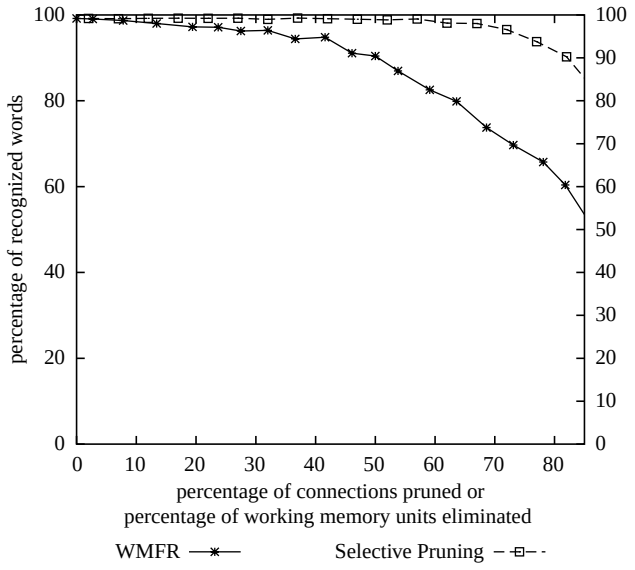

Fig. 5. Comparison of the robustness of the SPELT network under the two disconnection paradigms considered. In the case of Darwinist pruning, the abscissas represent the percentage of synaptic coefficients made 0 in the memory matrix; in the case of working memory functionality reduction (WMFR) abscissas represent the percentage of cells eliminated in the working memory layer.

preceded them during the training phase. In one of the simulations we found two different hallucinated words: 'hates' and 'seldom'. In this particular simulation, the hallucinated words always appeared following

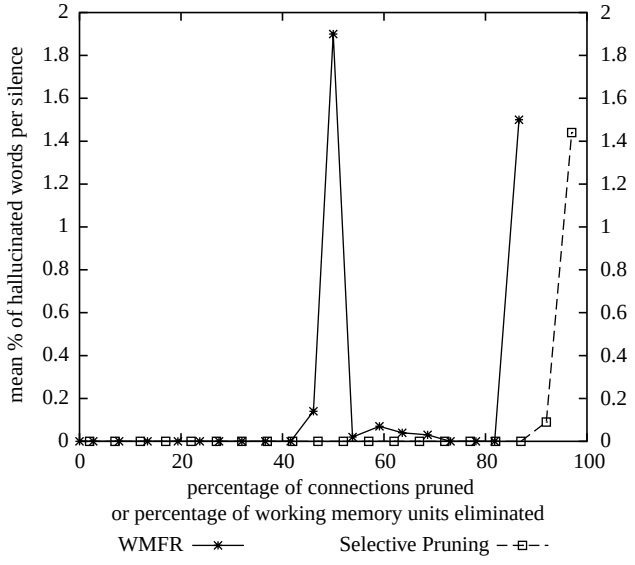

Fig. 6. The hallucinatory behavior was evaluated by testing the network with the presentation of silences after words and between sentences. This evaluation was repeated for each level of destruction. Within each of the 20 simulation sets 500 silences were presented. The ordinates show the mean percentage of hallucinations per silence (averaged over the 20 sets). Thus, for 50% WMFR the numbers presented imply that in average almost 2% of the 500 silent inputs to the network in each of the sets (i.e. 10 silences) caused a hallucination (but see Table 2). The peak in hallucinations seen in WMFR is mainly due to one of the simulation sets.

Table 3
Neuromodulation: simulation of the effect of reducing the activity of the working memory to eliminate hallucinations

%WMFR pruning	%Successfully recognized words	%Not recognized words	%Wrong identifications	Hallucinated words
13.4	98.4	1.6	0.0	0.0
23.7	98.3	1.7	0.0	0.0
36.6	97.5	2.5	0.0	0.0
41.3	97.3	2.7	0.0	0.0
46.1	96.1	3.8	0.1	0.0
50.0	95.6	4.3	0.1	0.0
53.8	94.3	5.5	0.2	0.0
68.6	88.2	10.9	0.9	0.0
86.6	80.5	19.3	0.2	0.0

The output of the working memory was multiplied by a constant scalar that in this particular example was 0.5, before entering the Kronecker product. Neuroanatomic manipulation procedure = Working memory functionality reduction. Average results over 20 simulation sets (300 words presentation, and 500 'silences' presented within each set).

the words that preceded them during the training phase. The bimodal profile of hallucinations in the WMFR stems from the fact that some of the simulation sets where hallucinations arouse in the first peak were totally disabled by further reductions in connectivity. After deterioration, they showed poor recognition capabilities and no hallucinations.

3.4.3. Neuromodulation and antipsychotic drugs

When output vectors of a deteriorated working memory were multiplied by a constant smaller than one, the model showed an improvement in its performance, in particular a reduction in the observed number of hallucinations. These results are shown in Table 3. Therefore, this particular modulation simulates the effect of an antipsychotic drug.

It is interesting to note that when we enhance the working memory output (multiplying by a constant higher than one) hallucinations appear even for normally connected networks. Hence, this last situation roughly mimics the induction of hallucinations in normal subjects by drugs acting as positive neuromodulators.

4. Discussion

The main purpose of this paper is to show that a neural network built on a subclass of sigma–pi model neurons is a powerful device, yet easy to train and comprehend. We showed how our model can be derived from the general sigma–pi unit and how that leads to the Kronecker product filtering of the inputs to a linear matrix associator. Since we use just one layer of neurons with modifiable weights, the network can be easily trained with the delta rule (Reali, 2002). In spite of its simplicity, our version of Elman topology is still able to learn a toy-model language. We showed that the model presents word recognition abilities based on linguistic expectations. Besides, our model

produced word 'hallucinations' after damage and repression of these hallucinations by parameter changes that simulate neuromodulatory actions, mimicking results obtained using SRNs (McGlashan & Hoffman, 2000).

Our model must be viewed as a gross approximation to real neuronal behavior with the purpose of showing the power of multiplicative interactions. We use formal neurons that are simple point unit models with linear output functions (see Koch & Segev, 2000). In some sensory systems, the linearity hypothesis is well supported by experimental findings (Brodie, Knight, & Ratliff, 1978). In associative memory networks, the hypothesis can be justified if the model neurons are assumed to be leaky integrators receiving a basal noisy input (Nass & Cooper, 1975). In our case, linearity should only be regarded as a simplifying assumption to keep the model as simple as possible (see also Cooper, 2000). Aside from linearity, the biological plausibility of our model rests on two strong assumptions: first, the existence of neurons able to multiply their inputs; second, the existence of a particular type of anatomy which permits the calculation of the Kronecker product. We have shown that the latter requirement can be relaxed provided that the network is large; in a network with fewer connections the incomplete Kronecker product can be statistically computed. Of course this depends on the particular encoding of information used (see Pomi & Mizraji, 1999). A detailed consideration of the robustness of different information representations within the network is currently under study (an example is contained in our studies of hallucinations, see below).

Regarding the first assumption—the existence of multiplicative synapses—let us mention some reports supporting the execution of multiplications by neural circuits involved in perception. Neurons of the monkey's posterior parietal lobe show 'gain fields' that can be explained by a multiplication of retinal and eye or head position signals (Andersen, Snyder, Bradley, & Xinget, 1997). It has been reported that in locust, the dendritic tree of a high-order visual neurons may function as a biophysical device that can carry out a multiplication of two independent inputs (Hatsopolous, Gabbiani, & Laurent, 1995). Recently, Peña and Konishi (2001) have reported that multiplication of separated postsynaptic potentials, rather than addition, can account for some responses of neurons in owl's auditory system. Multiplicative effects in neuronal processing have been increasingly looked for by neuroscientists (see specially Koch & Poggio, 1987, 1992; Poggio, 1990; Tal & Schwartz, 1997). Multiplication as a coincidence detector has also been explored with a variety of approaches going from signal analysis (Bialek & Zee, 1990) to integrate-and-fire neuron models (Bugmann, 1992). Support for the multiplicative capacities is said to be based upon properties of the NMDA receptor (Mel, 1992, 1993; Montagne & Sejnowski, 1994; Poggio, 1990). An extensive review on the computational properties of neurons including multiplication has recently been published (Koch & Segev, 2000).

Multiplication admits another interpretation if the sigma–pi units are regarded as phenomenological devices. In this case, besides synaptic mechanisms, multiplying effects can be obtained using higher-level circuits. In fact, under some conditions memory units can be regarded as neural networks themselves (Amari, 1977). Thus, multiplicative capacities can arise from shunting and gating mechanisms such as those exhibited by adaptive resonance theory (ART) networks (see for instance the appendix in Grossberg & Myers, 2000).

Advancing in the necessary refinement of neural cognitive theories requires the consideration of models, which have greater structural and dynamical complexity such as those, framed in the ART (Carpenter & Grossberg, 2003). In fact, ART-based models have recently been applied to study diverse aspects of neurological and psychiatric disorders (Grossberg, 1999). We plan to address this complexity in future works.

In parallel with the search for biological realism (which might be premature) it is important to evaluate the computational capabilities of this type of networks. In both, the previously explored applications of the Kronecker product model and the one examined here, we found that multiplicative context modulation is a very powerful strategy (Mizraji, 1989; Mizraji et al., 1994; Pomi & Mizraji, 1999, 2001). An important point to be made is that regardless of the detailed implementation, the attainment of a particular computational goal depends on the properties of the Kronecker product. The double scalar product filtering (see Eq. (6)) tends to make correlated patterns quasi-orthogonal, a fact which is at the roots of the model's computational powers. Our group has shown the suitability of associative memories modulated by multiplicative contexts for modeling classical tasks of psychology and computation. Of particular relevance here are interpolation (Mizraji et al., 1994), disambiguation of ambiguous perception (Pomi & Mizraji, 1999), logics and fuzzy logics (Mizraji, 1992).

In connection with these applications, in the present study we used our model to explore how the deterioration of a network responsible for some aspects of language processing can lead to hallucinations of the type seen in schizophrenia (McGlashan & Hoffman, 2000).

We found that the multiplicative network is able to recognize a vocabulary when it is presented as non-learned sentences having a correct grammar structure. The network is not only able to recognize single words but also relies on linguistic expectations (i.e. after a word has been recognized, the network expects to find a syntactically correct—and semantically suitable—word). In this sense, our model mimics the results obtained with SRNs. Linguistic expectations are demonstrated both by the presence of syntactically correct illusions when mild noise is inputted and by the fact that presenting words in random order reduces recognition capabilities (from more than 99% to less than 84%). This is to be compared to the results of Hoffman et al.

(1995) where an SRN was shown to reproduce word sensitivity rates of normal people. We did not attempt to fit precisely human recognition rates and it is clear that this would require scaling up the model as we discuss below in connection with the role of working memory.

In our model, the expectancy depends on the previous identified word, which imposes the context for recognition. Given the double filtering illustrated by Eq. (6) the output to a noisy input will be a word learned in a similar context. Raizada and Grossberg (2003) show how more complex versions of modulatory expectations might work in the functioning of cortex. Although their model is rather different from ours, we believe that they share the ability to enhance certain perceptions by contextual information. It remains to be seen whether the differences between the two models are or not reconcilable.

We also show that the network is resistant to damage when low levels of pruning were simulated. When we applied higher levels of pruning, the network progressively failed to perform recognition, and some spontaneous percepts analogous to hallucinated speech appeared. The simulation of reductions of connectedness led to hallucinated speech, as previous computational models predict (Hoffman & McGlashan, 1997; Hoffman et al., 1995), using a different neuron model and learning algorithm. We also used a superficially different vocabulary from that used by Hoffman, reinforcing the idea that many of the properties obtained by SRNs are insensitive to implementation details.

The fact that some phonetic codes are more robust to deterioration than others, as the hallucinatory behavior shown in Tables 1 and 2 demonstrates, is an example of the importance of information representation. From the biological point of view, this could mean that individual factors play an important role, i.e. individuals sharing the same general properties can be more or less tolerant to disruptions depending on the exact internal representation of information. Even if the real biological representation and networks are much more complex than ours, the model is enough to show this feature clearly.

In some important aspects our simulations differ from those of Hoffman and McGlashan. The most important difference in behavior is seen when performing an intermediate level of pruning. Hoffman and McGlashan (1997) show that moderately pruning a traditional SRN improves its recognition capabilities. In contrast, our model showed no consistent improvement, although some simulations sets displayed a minimal enhancing of their performance. This behavior is desired if we try to give an explanation for the persistence of schizophrenia in the population in spite of its maladaptive character (Crow, 2000). Our lack of improvement might be due to the fact that in all of our deterioration simulations, we disrupt both the working memory projections and—indirectly—the input projections. This can be readily seen from the type of product unit we use. Another possibility is that employing a single layer of linear output, sigma–pi unit prevents improvement by pruning. Moreover, the efficiency in learning for the small network does not leave room for improvement, something that will surely not be valid for larger networks. It is an interesting research problem to see if minor modifications can be made, either in the output function of the units or in their connectivity, to make pruning beneficial.

Another difference with the results obtained by Hoffman and McGlashan (1997) lies in the effect of damaging the working memory units. The connectivity lost in correlation with a definite working memory damage seems to have, according to our model, stronger and more negative consequences on linguistic functions than the pruning guided by the concept of neurodevelopmental ‘Darwinism’ (selective pruning) which does not target the working memory units. We found hallucinated speech when we reduced the working memory functionality by 46%, in contrast to the 92% needed to obtain them in the pruning method. This derives from the fact that the elimination of one context unit disturbs the input to the associative matrix in a very different way than what the pruning method does, making zero some entries that might have high weights.

When we use low levels of neural manipulations that are nevertheless sufficient to produce hallucinated speech, the multiplicative context-dependent network is still able to successfully recognize a high percentage of words (above 90%). This means that, according to the model, hallucinated speech can arise as a consequence of a deterioration of the capacities of normal linguistic expectations. This reinforces the previously explored hypothesis that working memory connectivity impairments (acting directly in the alteration of linguistic expectations) have injurious consequences in normal word association and recognition (Hoffman & McGlashan, 1997).

We are aware that our approach to working memory modeling is extremely crude. Since the main aim of the present paper is to understand the properties of the SPELT and compare it with traditional SRNs, we adopt very simple representations of important cognitive properties. It is clear that modeling working memory as in an SRN might have many drawbacks. In particular, to say that a network produced a hallucinated word we require that it perceives a word when a neutral vector representing silence is entered, reproducing the procedures of Hoffman et al. (1995). As has been analyzed and modeled by Grossberg and Myers (2000) the interpretation of silences is context-dependent and backward effects on perception show that the context is more complex than the word preceding the input word. Further developments in our SPELT model will include higher level processing modules with the possibility that the disambiguation of a given word could be helped, not only by the influence coming from the previous words, but also from the ones following the target word.

We also implemented ‘neuromodulatory simulations’ in a very simple way. We sought for simple modifications of parameters that could lead to a reduction of hallucinations.

If it is assumed that the entries in context vectors are the result of 'integrate and fire' neurons, working in the linear range (see for example Koch & Segev, 2000; Nass & Cooper, 1975) the same increase in the threshold of each neuron can reduce all the activities by the same fraction. It can be shown that in the multiplicative model, this modification reduces the chance of silent inputs being wrongly recognized as words. It should be remarked that this kind of neural modifications are opposite to what is needed in other network models of schizophrenia, where an increase in the strength of the output from a pathological working memory activity reduce hallucinations (e.g. Cohen & Servan-Schreiber, 1992). The action of dopamine has been included in some neural network models as an increase in the parameters controlling signal-to-noise ratio (Cohen & Servan-Schreiber, 1992; Spitzer, 1997). Hoffman et al. (1995) postulated a neuromodulatory disturbance on their model, considering that hyperdopaminergic activity causes the positive symptoms of schizophrenia. In contrast, in our network, multiplying working memory output vectors by a constant higher than one increases the number of hallucinations. This 'hallucinogenic' effect is expected according to our explanation of the neuroleptic-like behavior described above, but we know no evidence that supports working memory enhancements on human subjects who use hallucinogenic drugs.

Given the simplicity of SPELT particularly with regard to neuromodulation, we do not expect a precise matching of pharmacological data. If model units are to be taken as real neurons, we could modify the output of the working memory in more complex ways to achieve pharmacologically relevant properties. The fact that in order to reduce the hallucinations we require less influence of the working memory requires an explanation. Our simulations suggest that the pruned working memory not only fails to induce normal expectations but also enhances parasitic recognitions. Then, it would be necessary to 'turn the working memory off' in order to obtain a better functioning. This questions the role of working memory since recognition is minimally impaired, and robustness enhanced, when the working memory output is diminished (see Table 3). It is clear that these results deserve further consideration, but if the model is taken literally, the simulations of Table 3 imply that the pathophysiology of schizophrenia requires both an excessive pruning of connections and a higher excitability of working memory units. It would also suggest that the initial excitability level we have chosen would be a pathological rather than normal one (see Sawa & Snyder, 2002 for a review of the possible molecular and anatomical disturbances present in schizophrenia).

We think that these issues should be studied in connection to the scalability of our model. It is well known that many neural network models that work well in small problems deteriorate rapidly as the size of the task they accomplish grows (Minsky & Papert, 1988). The level of disconnection we need to impose to the network in order to have hallucinations (92% in the case of Darwinian pruning and 46% in the case of Working memory Functionality reduction) together with the neuromodulatory simulations commented above, highlights the need to study how these properties scale with increasing size of the network and of vocabulary to learn. In particular, notice that we do not expect the full connectivity implied by the Kronecker product to be operative (compare to Section 2.2) but given that we consider our SPELT as a gross approximation, we attempted no fine adjustments. Moreover in some preliminary simulations we employed low levels of random a priori disconnection (data not shown) with little effect in the general properties of the model. The high percentage of pruning needed to elicit hallucinations in our network should not be taken as meaning that the model is incompatible with Darwinist pruning being the cause of hallucinations. The connectivity of the SPELT gives it a high and artificial degree of robustness and we envision a realistic network having a much lower proportion of connections before training. For instance, as Section 2.2 shows, training the model with 60% of the connections would show almost no deterioration in performance and yet a lower proportion of pruning needed to achieve hallucinatory behavior.

The present work is a first attempt to apply the SPELT model to language recognition and the study of the possible role of working memory damage in producing hallucinated voices in schizophrenia. We have shown that many of the results obtained by using an Elman type of network are retained and that others deserve further considerations. We think that understanding the differences between our results and those of other models (notably those of Hoffman & McGlashan, 1997; Hoffman et al., 1995; see also Grossberg, 1999) is central to 'calibrating' the application of neural networks to study physiology and pathology and to devise experimental tests of the models.

Many works support the idea of schizophrenia as a pathology involving reduced connections (reviewed in McGlashan & Hoffman, 2000). Recent Diffusion Tensor Imaging studies implicate white matter anomalies in schizophrenia (Agartz, Andersson, & Skare, 2001; Buchsbaum et al., 1998; Foong et al., 2000; Hubl et al., 2004; Lim et al., 1999). Although most of these studies were done with small sample sizes they add valuable information about where disconnection are in the brain. In our model, a reduction of connectivity localized at the synaptic level or due to a broader 'white matter disruption', leads to spontaneous activity that can be classified as hallucinations. In the present work our aim was to explore the ability of the multiplicative model not only to simulate a simple linguistic task but also to show that it converges to similar psychological conclusions as previously reached using SRNs. The results presented above indicate that the multiplicative model presents, at least for the tasks explored, properties that make it suitable to represent cognitive functions as complex as those shown in

psychiatric studies. Moreover, these results suggest that conclusions obtained by modeling linguistic performances using SRNs may depend more on the recursive character of Elman topology, than on the detailed properties of the units used. We believe that the model's success in computing the present task provides evidence for its potential for further utilization, especially in the cognitive science field.

It has been proved that the SRNs of Elman are (at least) Turing equivalent (Siegelmann & Sontag, 1995). In a recent article comparing the performance of several recurrent neural network models, Lawrence, Giles, and Fong (2000) concluded that the Elman network learns to distinguish grammatical from ungrammatical sentences better than other recurrent networks. We have shown in this paper that the multiplicative model can reproduce results obtained by traditional SRNs and that the former has some advantages over the latter.

One important reason for using the multiplicative model is that during training the local correcting information only propagates forward and not backward, relying on synaptic-like mechanisms that have shown to be somehow present in neurons, thus, increasing its biological plausibility. Clearly, in a higher timescale and to reproduce complex behavior, learning has to be modulated by information arriving from diverse cortical areas (Raizada & Grossberg, 2003). In our model, this modulation in learning depends on the influence of the working memory. A second reason to use the SPELT is that our version of the multiplicative architecture is simple, relaying on fewer neuron layers than a traditional SRN to produce the same performance and it is very easy to train. Moreover, its multiplicative nature confers it with additional computational and analytical properties at the price of little extra complexity. Given these advantages, one may ask whether the SPELT model, as we presented it, has the full potential of SRNs, but this is a matter that requires further investigation.

Acknowledgements

We acknowledge the three anonymous referees for their helpful comments. This work was partly supported by PEDECIBA, Uruguay.

References

Agartz, I., Andersson, J. L., & Skare, S. (2001). Abnormal brain white matter in schizophrenia: A diffusion tensor imaging study. *Neuroreport*, *12*, 2251–2254.

Amari, S.-I. (1977). Neural theory of association and concept-formation. *Biological Cybernetics*, *26*, 175–185.

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, *20*, 303–330.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.

Bellman, R. (1960). *Introduction to matrix analysis*. New York: McGraw-Hill.

Bialek, W., & Zee, A. (1990). Coding and computation with neural spike trains. *Journal of Statistical Physics*, *59*, 103–115.

Boers, J. W., & Kuiper, H. (1992). Biological metaphors and the design of modular artificial neural networks. Master's Thesis, Department of Computer Science and Experimental Psychology, Leiden University, The Netherlands.

Brodie, S., Knight, B. W., & Ratliff, F. (1978). The response of the *Limulus* retina to moving stimuli: A prediction by Fourier synthesis. *Journal of General Physiology*, *72*, 129–166.

Buchsbaum, M. S., Tang, C. Y., Peled, S., Gudbjartsson, H., Lu, D., Hazlett, E. A., et al. (1998). MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *Neuroreport*, *9*, 425–430.

Bugmann, G. (1992). Multiplying with neurons: Compensation for irregular input spike trains using time-dependent synaptic efficiencies. *Biological Cybernetics*, *68*, 103–115.

Carpenter, G. A., & Grossberg, S. (2003). Adaptive resonance theory. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks*2nd ed. (pp. 87–90). Cambridge, MA: MIT Press, 87–90.

Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, *23*, 157–205.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Reviews*, *90*, 45–77.

Cooper, L. N. (2000). Memories and memory: A physicist's approach to the brain. *International Journal of Modern Physics A*, *15*, 4069–4082.

Crow, T. (2000). Schizophrenia as the price that *Homo sapiens* pays for language: A resolution of the central paradox in the origin of the species. *Brain Research Reviews*, *31*, 118–129.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Elman, J. L. (1995). Language as a dynamical system. In R. F. Port, & T. van Gelder (Eds.), *Mind as motion. Explorations in the dynamics of cognition* (pp. 195–225). Cambridge, MA: MIT Press, 195–225.

Foong, J., Maier, M., Clark, C. A., Barker, G. J., Miller, D. H., & Ron, M. A. (2000). Neuropathological abnormalities of the corpus callosum in schizophrenia: A diffusion tensor imaging study. *Journal of Neurology, Neurosurgery and Psychiatry*, *55*, 242–244.

Grossberg, S. (1999). Neural models of normal and abnormal behavior: What do schizophrenia, Parkinsonism, attention deficit disorder and depression have in common?. In J. Reggia, E. Rupin, & D. L. Glanzman (Eds.), *Disorders of brain behavior and cognition: The neurocomputational perspective* (pp. 375–406). Amsterdam: Elsevier, 375–406.

Grossberg, S., & Myers, C. (2000). The resonant dynamics of conscious speech: Interword integration and duration-dependent backward effects. *Psychological Review*, *107*, 735–767.

Hatsopolous, N., Gabbiani, F., & Laurent, G. (1995). Elementary computation of object approach by a wide-field visual neuron. *Science*, *270*, 1000–1003.

Hoffman, R. E. (1987). Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Archives of General Psychiatry*, *44*, 178–185.

Hoffman, R. E., & McGlashan, T. H. (1997). Synaptic elimination, neurodevelopment, and the mechanism of hallucinated voices in schizophrenia. *American Journal of Psychiatry*, *154*, 1683–1689.

Hoffman, R. E., & McGlashan, T. H. (1998). Reduced corticocortical connectivity can induce speech perception pathology and hallucinated 'voices'. *Schizophrenia Research*, *30*, 137–141.

Hoffman, R. E., Rapaport, J., Ameli, R., McGlashan, T. H., Harcherik, D., & Servan-Schreiber, D. (1995). A neural network simulation of hallucinated voices and associated speech perception impairments in schizophrenia patients. *Journal of Cognitive Neuroscience*, *7*, 479–497.

Hubl, D., Koenig, T., Strik, W., Federspiel, A., Kreis, R., Boesch, C., et al. (2004). Pathways that make voices: White matter changes in auditory hallucinations. *Archives of General Psychiatry*, *61*, 658–668.

Koch, C., & Poggio, T. (1987). Biophysics of computation: Neurons, synapses and membranes. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Synaptic function* (pp. 637–697). New York: Wiley, 637–697.

Koch, C., & Poggio, T. (1992). Multiplying with synapses and neurons. In T. McKenna, J. Davis, & S. F. Zornerster (Eds.), *Single neuron computation* (pp. 315–345). New York: Academic Press, 315–345.

Koch, C., & Segev, I. (2000). The role of single neurons in information processing. *Nature Neuroscience*, *3*, 1171–1177.

Lawrence, S., Giles, C. L., & Fong, S. (2000). Natural language grammatical inference with recurrent neural networks. *IEEE Transactions on Knowledge and Data Engineering*, *12*, 126–140.

Lim, K. O., Hedehus, M., Moseley, M., de Crespigny, A., Sullivan, E. V., & Pfefferbaum, A. (1999). Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Archives of General Psychiatry*, *56*, 367–374.

McGlashan, T. H., & Hoffman, R. E. (2000). Schizophrenia as a disorder of developmentally reduced synaptic connectivity. *Archives of General Psychiatry*, *57*, 637–648.

Mel, B. W. (1992). NMDA-based pattern discrimination in a modeled cortical neuron. *Neural Computation*, *4*, 502–517.

Mel, B. W. (1993). Synaptic integration in an excitable dendritic tree. *Journal of Neuroscience*, *70*, 1086–1101.

Minsky, M., & Papert, S. (1988). *Perceptrons*. Cambridge, MA: MIT Press.

Mizraji, E. (1989). Context-dependent associations in linear distributed memories. *Bulletin of Mathematical Biology*, *51*, 195–205.

Mizraji, E. (1992). Vector logics: The matrix-vector representation of logical calculus. *Fuzzy Sets and Systems*, *50*, 179–185.

Mizraji, E., Pomi, A., & Alvarez, F. (1994). Multiplicative contexts in associative memories. *Biosystems*, *32*, 145–161.

Montagne, P. R., & Sejnowski, T. J. (1994). The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms. *Learning and Memory*, *1*, 1–33.

Nass, M. M., & Cooper, L. N. (1975). A theory for the development of feature detecting cells in visual cortex. *Biological Cybernetics*, *19*, 1–18.

Neville, R. S., & Elridge, S. (2002). Transformation of sigma–pi nets: Obtaining reflected functions by reflecting weight matrices. *Neural Networks*, *15*, 375–393.

Pao, Y. H. (1989). *Adaptive pattern recognition and neural networks*. Reading, MA: Addison-Wesley.

Peña, J. L., & Konishi, M. (2001). Auditory spatial receptive field created by multiplication. *Science*, *292*, 249–252.

Poggio, T. (1990). A theory of how the brain might work. *The Brain. Cold Spring Harbor Simposia on Quantitative Biology* (Vol. LV). New York: The Cold Spring Harbor Laboratory Press pp. 390–431.

Pomi, A., & Mizraji, E. (1999). Memories in context. *Biosystems*, *50*, 173–188.

Pomi, A., & Mizraji, E. (2001). A cognitive architecture that solves a problem stated by Minsky. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, *31*, 729–734.

Raizada, R. D. S., & Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, *13*, 100–113.

Reali, F. (2002). Interacciones multiplicativas en modelos de redes neuronales: algunas aplicaciones en redes de procesamiento del lenguaje. Tesis de Maestria. PEDECIBA—Facultad de Ciencias, Uruguay.

Ritter, H., & Kohonen, T. (1989). Self-organizing semantic maps. *Biological Cybernetics*, *61*, 241–254.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, & J. L. McClelland (Eds.), *Parallel distributed processing*. Cambridge, MA: MIT Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representation by error propagation. In D. E. Rumelhart, & J. L. McClelland (Eds.), *Parallel distributed processing*. Cambridge, MA: MIT Press.

Sartorius, N., Shapiro, R., & Jablenski, A. (1974). The International Pilot Study of schizophrenia. *Schizophrenia Bulletin*, *1*, 21–35.

Sawa, A., & Snyder, S. H. (2002). Schizophrenia: Diverse approaches to a complex disease. *Science*, *296*, 692–695.

Schmitt, M. (2002). On the complexity of computing and learning with multiplicative neural networks. *Neural Computation*, *14*, 241–301.

Shergill, S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenics patients with auditory hallucinations. *American Journal of Psychiatry*, *157*, 1691–1693.

Shergill, S. S., Robin, M., Murray, R. M., & McGuire, P. K. (1998). Auditory hallucinations: A review of psychological treatments. *Schizophrenia Research*, *32*, 137–150.

Siegelmann, H. T., & Sontag, E. D. (1995). On the computational power of neural nets. *Journal of Computer and System Sciences*, *50*, 132–150.

Spitzer, M. (1997). A cognitive neuroscience view of schizophrenic thought disorder. *Schizophrenia Bulletin*, *23*, 29–46.

Tal, D., & Schwartz, E. L. (1997). Computing with the leaky integrate and fire neuron: Logarithmic computation and multiplication. *Neural Computation*, *9*, 305–318.

Widrow, B., & Hoff, M. E. (1960). Adaptative switching circuits. In J. A. Anderson, & E. Rosenfeld (Eds.), *Neurocomputing*. Cambridge, MA: MIT Press.