Cogn Neurodyn (2008) 2:79–88
DOI 10.1007/s11571-008-9042-4

RESEARCH ARTICLE

Language processing with dynamic fields

Peter beim Graben · Dimitris Pinotsis · Douglas Saddy · Roland Potthast

Received: 15 January 2008 / Accepted: 24 March 2008 / Published online: 31 May 2008

Abstract We construct a mapping from complex recursive linguistic data structures to spherical wave functions using Smolensky's filler/role bindings and tensor product representations. Syntactic language processing is then described by the transient evolution of these spherical patterns whose amplitudes are governed by nonlinear order parameter equations. Implications of the model in terms of brain wave dynamics are indicated.

Keywords Computational psycholinguistics · Language processing · Fock space · Dynamic fields

Introduction

Human language processing is accompanied by modulations of the ongoing electrophysiological brain waves. If these are evaluated in a stimulus-locked manner (cf. the contributions of Fründ et al. and Kiebel et al. in this special issue), one speaks about event-related brain potentials that reflect syntactic (Osterhout and Holcomb 1992; Friederici 1995), semantic (Kutas and Hillyard 1980, 1984) and also pragmatic (Noveck and Posada 2003; Drenhaus et al. 2006) processing problems.

Modeling human language processing has previously relied mostly upon computational approaches from automata theory and cognitive architectures (Hopcroft and Ullman 1979; Lewis and Vasishth 2006), while dynamical system models that could also be able to account for brain wave dynamics are still in their infancy (beim Graben et al. 2008; Vosse and Kempen 2000; Garagnani et al. 2007). The contentious issues of the former approach regarding the computational viability of grammars intended to capture properties of human language have been with us since Chomsky (1957). The nature of this long debate centers around whether the models of language and language processing proposed are, in principle, computable; computability being a minimal requirement for any attempt to formally describe complex behavior exhibited by a biological system. As our understanding of the biology and physiology of the brain has increased, similar issues have guided the development of computational models of brain physiology. There are now interesting competing models for both language processes (Elman 1995; Tabor et al. 1997; Christiansen and Chater 1999; Vosse and Kempen 2000; beim Graben et al. 2008; Hagoort 2005; Lewis and Vasishth 2006; van der Velde and de Kamps 2006; Smolensky and Legendre 2006) and brain functions (Wilson and Cowan 1973; Amari 1977; Jirsa and Haken 1996; Coombes et al. 2003; Jirsa 2004; Wright et al. 2004; Garagnani et al. 2007) (see also the contributions in this special issue).

Although language understanding takes place in the human brain, computational modeling of language processes and computational modeling of brain physiology are not the same. The models address quite different levels of description and reflect different assumptions regarding the operational primitives and desired final states. While the former generally refer

Electronic supplementary material The online version of this article (doi: 10.1007/s11571-008-9042-4) contains supplementary material, which is available to authorized users.

P. beim Graben (✉) · D. Saddy
School of Psychology and Clinical Language Sciences, University of Reading, Whiteknights, P.O. Box 217, Reading RG6 6AH, UK
e-mail: p.r.beimgraben@reading.ac.uk

D. Pinotsis · R. Potthast
Department of Mathematics, University of Reading, Whiteknights, P.O. Box 217, Reading RG6 6AH, UK

to abstract *feature spaces* such as "stack tapes", "neural blackboards" or "sketch pads" (Hopcroft and Ullman 1979; Anderson 1995; van der Velde and de Kamps 2006), the latter aim at describing membranes, neurons, or mass potentials (beim Graben 2008). Nevertheless, a successful model of language processing should be interpretable in terms of a model of brain function. This represents a new kind of evaluative paradigm on models of language or more generally cognitive processes: (1) Is the model computationally tractable? (2) Is the model interpretable in terms of descriptions of the supporting physiology?

Following this approach, computational models of language processing must take very seriously the properties inherent in computational models of the brain. In doing so we are led to address the issues and assumptions raised by the differing levels of description targeted by the models. We are also provided with a metric for comparison of competing models.

Here we address the question: Can the basic assumptions and mechanisms underpinning a language processing model be expressed in terms that are compatible with viable models of the brain? Our answer will be yes, and furthermore, we argue that this result has direct bearing on important debates regarding the viability of certain classes of language processing models.

An outstanding controversial issue is whether grammars and processing mechanisms of human languages are *recursive* or not (Hauser et al. 2002; Everett 2005) and whether neural network models should implement this property either *faithfully* or rather by means of *graceful saturation* (Christiansen 1992; Christiansen and Chater 1999; Smolensky 1990; Smolensky and Legendre 2006). Especially Smolensky's *Integrated Connectionist/Symbolic architecture* (Smolensky and Legendre 2006; Smolensky 2006) represents symbolically meaningful states by very few very sparse patterns in very high, yet finite, dimensional activation vector spaces (beim Graben et al. 2007, 2008).

In order to avoid such sparse representations, we suggest employing infinite-dimensional function spaces in this paper. This approach is in line with related attempts by Smolensky (1990), Moore and Crutchfield (2000), Maye and Werning (2004), Werning and Maye (2007) and compatible with neural and dynamical field theories of cognition (Wilson and Cowan 1973; Amari 1977; Jirsa and Haken 1996; Coombes et al. 2003; Jirsa 2004; Wright et al. 2004; Erlhagen and Schöner 2002; Schöner and Thelen 2006; Thelen et al. 2001).

We shall construct a mapping from the dynamics of language processing into a field dynamics in several steps. First, in Section "Dynamic parsing", we describe a simplified parsing dynamics based upon a toy-grammar, to be introduced in Section "Grammars". Second, in Section "Fock space representations", we shall study a particular vector space representation of phrase structure trees as suggested by Smolensky (1990) and Smolensky and Legendre (2006). We also embed the time-discrete representation of the parsing process into a time-continuous dynamics. In Section "Order parameter dynamics" we derive neurally motivated order parameter equations (Haken 1983) of the parser. Third, in Section "Spherical wave functions", we map the vectorial representation of the parsing states into the function space of spherical harmonics defined on an abstract feature space. Here, the crucial point is to reduce the dimension of the vector space by a separation of time scales. Finally, the different parts of the model are integrated into a field-theoretic representation with transient dynamics in Section "Dynamic fields". Section "Simulations" presents results of numerical simulations of the parsing example discussed throughout the paper. We conclude with a discussion about a tentative relation between our model and electrophysiological findings on language-related brain waves.

Grammars

Sentences are hierarchically structured objects, commonly described by *phrase structure trees* in linguistics (Chomsky 1957; Hopcroft and Ullman 1979). Contemporary linguistic and parsing theories have elaborated considerably on these early approaches (cf. Stabler 1997 for one particular account). For our purposes, we investigate a toy-grammar that simplifies our task but is nevertheless representative of the basic operations required of a natural language parser.

Consider e.g. the sentence

Example 1 `Susan ate grass`.

This simple sentence (`S`) consists of a subject, the noun phrase `Susan`, and a predicate, the verbal phrase (`VP`), `ate grass`. The latter in turn is construed from the verb `ate` and another noun phrase, the direct object `grass`. Therefore, the sentence from example 1 can be described by the tree depicted in Fig. 1.

From the phrase structure tree in Fig. 1, a context-free grammar (CFG) can be easily derived by taking the node `S` as the *start symbol* and `VP` as another *nonterminal* such that every branching of the tree corresponds to a *production* of the grammar $G = (\mathbf{T}, \mathbf{N}, P, \mathtt{S})$, with

$$
\begin{aligned}
\mathbf{T} &= \{\mathtt{Susan}, \mathtt{ate}, \mathtt{grass}\} \\
\mathbf{N} &= \{\mathtt{S}, \mathtt{VP}\} \\
P &= \{(1) \quad \mathtt{S} \to \mathtt{Susan} \quad \mathtt{VP}, \\
&\qquad (2)\, \mathtt{VP} \to \mathtt{ate} \quad \mathtt{grass} \\
&\qquad \},
\end{aligned} \tag{1}
$$

where $\mathbf{T}$ is a finite *terminal alphabet*, $\mathbf{N}$ is a finite set of nonterminal *categories*, $P \subset \mathbf{N} \times (\mathbf{T} \cup \mathbf{N})^*$ comprises the

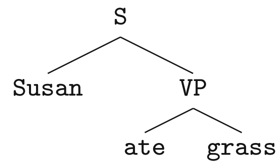

Fig. 1 Phrase structure tree of the sentence `Susan ate grass`

production rules, and $\mathtt{S} \in \mathbf{N}$ is the distinguished *start symbol*.

The productions $p \in P$ are usually drawn as *rule expansions*

$$p : A \rightarrow \gamma \tag{2}$$

as in (1), where $A \in \mathbf{N}$ and γ is a finite *word* of the Kleene hull $(\mathbf{T} \cup \mathbf{N})^*$ (Hopcroft and Ullman 1979). We call a production p *binary branching*, if $\gamma \in (\mathbf{T} \cup \mathbf{N})^2$, i.e. γ in (2) is a word of length 2, $\gamma = v_1 v_2$, with $v_1, v_2 \in \mathbf{T} \cup \mathbf{N}$. Accordingly, we call a CFG G binary branching, if all rules $p \in P$ are binary branching. Obviously, our example grammar G is binary branching.

Dynamic parsing

In order to understand a sentence the brain has to recognize at least "who is doing what to whom?" (Bornkessel et al. 2005), i.e, it has to reconstruct the phrase structure tree from the sequence of words. This mapping from a sentence to a tree is called *parsing*. Context-free languages can be parsed through push-down automata (Hopcroft and Ullman 1979). The simplest of these devices, the top-down recognizer, emulates the so-called *left-derivation* of a phrase structure tree, where always the leftmost not yet expanded nonterminal is expanded according to the rules of the grammar (1). Starting with the start symbol `S` we can thus derive the following strings:

$$\begin{aligned} \mathtt{S} &\rightarrow \mathtt{Susan} \quad \mathtt{VP} \\ &\rightarrow \mathtt{Susan} \quad \mathtt{ate} \quad \mathtt{grass}. \end{aligned} \tag{3}$$

Binary branching CFGs give rise to *labeled binary phrase structure trees* by successively expanding rules from P through left-derivations as in Fig. 1. Figure 2 shows the evolution of the phrase structure trees for the sentence `Susan ate grass` according to grammar (1).

Figure 2 reveals parsing as a *dynamics* in the space of phrase structure trees (Kempson et al. 2001). Formally, we define: Let T be the set of binary labeled phrase structure trees consistent with a given binary branching context-free grammar G. Clearly, T contains the "tree" `S` (the start symbol at the root) and at least one tree s for each well-formed sentence. A mapping

$$\pi : T \rightarrow T, \quad x \mapsto y, \quad x, y \in T \tag{4}$$

is a *sequential dynamic top-down parser* if the following holds: There is a constant $L \in \mathbb{N}$ such that for all $l \in \mathbb{N}, 0 \leq l \leq L$:

1. $\pi^l(\mathtt{S})$ is a subtree of s of height l,
2. π expands the tree s as a left-derivation,
3. $\pi^L(\mathtt{S}) = s$.

We call L the *duration* of the parsing process. The *parse* of s generated by π is the trajectory

$$U = (\pi^l(\mathtt{S}))_{0 \leq l \leq L}, \tag{5}$$

which is a "word" of length L in the Kleen hull T^*.

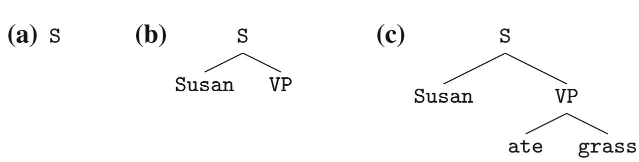

Fig. 2 Left-derivation (3) of the sentence `Susan ate grass` according to grammar (1)

Fock space representations

In this section, we present a mathematically rigorous reconstruction of the tensor product representations that have been introduced by Smolensky (1990, 2006), Smolensky and Legendre (2006) and further supported by Mizraji (1989, 1992).

Let S be a set of symbolic structures, e.g., of feature lists (Stabler 1997) or of phrase structure trees (i.e. $S = T$) (Chomsky 1957; Hopcroft and Ullman 1979). How can we represent a structured expression $s \in S$ by a vector $\mathbf{s}$ in some vector space? We follow Smolensky by introducing two finite sets F and R of *elementary fillers* and *elementary roles*, respectively.

Consider e.g. the set of binary labeled trees T for the CFG G (1) as an example. Let $s \in T$ be the tree in Fig. 2c. First, we have to chose fillers and roles. A suitable choice for the elementary fillers are the variables of G, i.e. $F = \mathbf{T} \cup \mathbf{N}$. The elementary roles are the three positions $r_1 = \textsc{Parent}$, $r_2 = \textsc{LeftChild}$, $r_3 = \textsc{RightChild}$, $R = \{r_1, r_2, r_3\}$ as indicated in Fig. 3.

A filler/role binding for a basic building block of s, denoted f/r, is an ordered pair $(f, r) \in F \times R$. Decomposing s into a set of (elementary) filler/role bindings yields a subset $f' = \{(f_i, r_j) \mid i \in I, j \in J\}$ (I, J are particular index sets) of $F \times R$. The set f' is therefore an element of the power set

$$f' \in F_1 = \wp(F \times R), \tag{6}$$

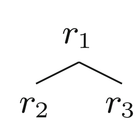

Fig. 3 Elementary role positions of a labeled binary tree

where $\wp(X)$ denotes the set of all subsets of a set X.

In our example, we first bind the elementary fillers VP to r_1, ate to r_2 and grass to r_3, obtaining the *complex filler*

$$f' = \{(\mathtt{VP}, r_1), (\mathtt{ate}, r_2), (\mathtt{grass}, r_3)\} \in F_1.$$

Next, recursion comes into the game. The subsets f' of $F \times R$ are complex fillers that can in turn bind to other roles: f'/r. This again is an ordered pair $(f', r) \in \wp(F \times R) \times R$, belonging to the next-level filler/role binding $f'' = \{(f'_i, r_j) \mid i \in I', j \in J'\}$. Therefore

$$f'' \in F_2 = \wp(\wp(F \times R) \times R). \tag{7}$$

Looking at the tree Fig. 2c in our example, reveals that the complex filler $f' = \{(\mathtt{VP}, r_1), (\mathtt{ate}, r_2), (\mathtt{grass}, r_3)\}$ is recursively bound to tree position r_3 at the higher level, whereas the elementary fillers S and Susan are attached to r_1 and r_2, respectively. Thus, the tree $s \in T$ is mapped onto its filler/role binding

$$f'' = \{(\mathtt{S}, r_1), (\mathtt{Susan}, r_2), (\{(\mathtt{VP}, r_1), (\mathtt{ate}, r_2), (\mathtt{grass}, r_3)\}, r_3)\} \in F_2. \tag{8}$$

For even higher trees, we have to repeat this construction recursively, entailing a hierarchy of filler/role bindings

$$\begin{aligned} F_0 &= F \\ F_{n+1} &= \wp(F_n \times R). \end{aligned} \tag{9}$$

In this way, any finite structure $s \in S$ becomes decomposed into its filler/role bindings by a map

$$\beta : S \to F_N, \quad \beta(s) \in F_N, \quad s \in S \tag{10}$$

for a particular $N \in \mathbb{N}$. In order to deal with recursion properly, we further define the collection

$$F_\infty = R \cup \left(\bigcup_{n=1}^{\infty} F_n \right). \tag{11}$$

After decomposing the structure s into its filler/role bindings, $\beta(s)$, we map s onto a vector $\mathbf{s}$ from a vector space $\mathcal{F}$ by the *tensor product representation* ψ, obeying

1. $\psi : F_\infty \to \mathcal{F}$,
2. $\psi(F_n)$ is a subspace of $\mathcal{F}$, for all $n \in \mathbb{N}$, in particular for $F_0 = F$ is $\psi(F) = \mathcal{V}_F$ a subspace of $\mathcal{F}$,
3. $\psi(R) = \mathcal{V}_R$ is a subspace of $\mathcal{F}$,
4. $\psi((f, r)) = \psi(f) \otimes \psi(r)$, for all $f \in F_n$, $r \in R$,
5. $\psi(\bigcup_{i \in I} \beta(s_i)) = \bigoplus_{i \in I} \psi(\beta(s_i))$ for all substructures s_i of $s \in S$.

Taken together, these properties yield that $\mathcal{F}$ is the Fock space

$$\mathcal{F} = \bigoplus_{n=1}^{\infty} \mathcal{V}_F \otimes \bigotimes_{k=1}^{n} \mathcal{V}_R, \tag{12}$$

known from quantum field theory (Haag 1992).

In order to apply the tensor product representation to our example CFG G (1) with phrase structure trees T, we identify the categories $f \in \mathbf{T} \cup \mathbf{N}$ of G with their associated filler vectors $\psi(f)$. On the other hand, we represent the roles with the "one-particle" Fock space basis

$$\psi(r_1) = |1\rangle, \quad \psi(r_2) = |2\rangle, \quad \psi(r_3) = |3\rangle \tag{13}$$

This notation has the advantage, that both, the tensor products in item 4, $\psi(f_i) \otimes \psi(r_j)$, and the direct sums in item 5, $\bigoplus_{i \in I} \psi(\beta(s_i))$, can be omitted, simply writing $\psi(f_i)|r_j\rangle$, and $\sum_{i \in I} \psi(\beta(s_i))$, respectively.

The tree s in (8) is than mapped onto the Fock space vector

$$\begin{aligned} \psi(\beta(s)) &= \mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle + (\mathtt{VP}|1\rangle + \mathtt{ate}|2\rangle + \mathtt{grass}|3\rangle)|3\rangle \\ &= \mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle + \mathtt{VP}|13\rangle + \mathtt{ate}|23\rangle + \mathtt{grass}|33\rangle \\ &\in \mathcal{F}, \end{aligned} \tag{14}$$

where we wrote the tensor products $|i\rangle \otimes |j\rangle = |i\rangle|j\rangle$ as "two-particle" Fock space vectors $|i\, j\rangle$.

Now, we are able to map a parse, namely a trajectory of trees $U \in T^*$ generated by a sequential dynamic top-down parser π (5) onto a trajectory of Fock space vectors

$$\mathcal{U} = (\psi(\beta(\pi^l(\mathtt{S}))))_{0 \le l \le L}. \tag{15}$$

Correspondingly, the parser π is represented by a nonlinear *operator*

$$P_\pi : \mathcal{F} \to \mathcal{F}, \quad \mathbf{x} \mapsto \mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in \mathcal{F} \tag{16}$$

defined through

$$(P_\pi \circ \psi \circ \beta)(x) = (\psi \circ \beta \circ \pi)(x) \tag{17}$$

for all $x \in T$ belonging to a parse U.

For our example above, the Fock space representation $\mathcal{U}$ of the parse U shown in (3) and in Fig. 2, is obtained as

$$\begin{aligned} \mathcal{U} = (&\mathtt{S}|1\rangle, \mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle + \mathtt{VP}|3\rangle, \mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle \\ &+ \mathtt{VP}|13\rangle + \mathtt{ate}|23\rangle + \mathtt{grass}|33\rangle) \end{aligned} \tag{18}$$

In Section "Dynamic fields", we are going to describe the Fock space parser P_π by a dynamically evolving field. A suitable function space for these fields will be constructed in Section "Spherical wave functions". To this aim, we need an embedding of the time-discrete dynamics (15) into continuous time. We achieve this construction by an *order parameter expansion* (Haken 1983) of the form

$$\mathbf{u}(t) = \sum_{l=0}^{L} \lambda_l(t) \psi(\beta(\pi^l(\mathtt{S}))), \quad t \in \mathbb{R}_0^+ \tag{19}$$

where the time-dependent coefficient $\lambda_l(t)$ is the order parameter for the lth subtree $\pi^l(\mathtt{S})$ of the parse U. Each order parameter $\lambda_l(t)$ assumes a unique maximum at time $T_l \in \mathbb{R}_0^+ (T_{l+1} > T_l)$ when the lth tree has been established.

Accordingly, T_L denotes the duration of the whole parse in continuous time. The functions $\lambda_l(t)$ form a Lagrange basis with $\lambda_l(T_k) = A_k\delta_{lk}$ and A_k the amplitude of the k-th order parameter (Kress 1998, Chap. 8)

The order parameter dynamics is usually governed by *order parameter equations*

$$\tau_l \frac{\mathrm{d}\lambda_l(t)}{\mathrm{d}t} + \lambda_l(t) = g_l(\lambda_0(t), \lambda_1(t)\ldots, \lambda_L(t)) \tag{20}$$

with appropriate functions g_l (Haken 1983).

Order parameter dynamics

We will discuss two different approaches to determine the time evolution of the coefficients $\lambda_l(t)$, $l = 0,\ldots, L$ and $t \in \mathbb{R}_0^+$ The first approach will be based on a simple recursion formula leading to a partition of unity for the coefficients. The second approach is build on order parameter dynamics as suggested by Haken (1983). Here, we will incorporate the neural background of our mapping and use a system of coupled nonlinear differential equations based on a leaky integrator neuron model.

Dynamics based on a recursion formula

For our recursive dynamics we start with some function δ_0 shaping the decay dynamics of the our coefficients. Here, the constant Δ denotes the maximal time interval on which a coefficient is active. For this first approach this means that $\lambda_l(t) > 0$. Later we will also work with coefficients which are not compactly supported, then activation is more complex and might be understood as the period in time in which the coefficient superseeds some threshold.

We assume that δ_0 is monotonic and continuous on $[0, \Delta/2]$, $\delta_0(0) = 1$ and $\delta_0(\Delta/2) = 0$. Also, we define $\delta_1 = 1 - \delta_0$ on $[0, \Delta/2]$, which yields $\delta_1(0) = 0$ and $\delta_1(\Delta/2) = 1$. Now, we set

$$\lambda_0(t) = \begin{cases} \delta_1(t), & t \in [0, \Delta/2] \\ \delta_0\left(t - \frac{\Delta}{2}\right), & t \in [\Delta/2, \Delta] \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Then, we define the coefficients λ_l for $l \geq 1$ by

$$\lambda_l(t) = \lambda_{l-1}\left(t - \frac{\Delta}{2}\right), \quad l = 1, \ldots, L-1. \tag{22}$$

The functions λ_l build a *partition of unity*, i.e. we have the property

$$\sum_{l=0}^{L-1} \lambda_l(t) = 1, \quad t \in \mathbb{R}. \tag{23}$$

Further, the support of the coefficient λ_l is a subset of $l \cdot \frac{\Delta}{2} + [0, \Delta]$.

Neural order parameter dynamics

Basically, Eq. 20 is a *leaky integrator equation* that is often used in neural modeling (beim Graben and Kurths 2008; beim Graben 2008). It can also be seen as a discretized version of the Amari equation for neural/dynamical fields (Wilson and Cowan 1973; Amari 1977; Jirsa and Haken 1996; Coombes et al. 2003; Jirsa 2004; Wright et al. 2004; Erlhagen and Schöner 2002; Schöner and Thelen 2006; Thelen et al. 2001). Thus it is promising to relate (20) with brain dynamics.

Since each well-established parse state s_l at time T_l triggers its successor s_{l+1}, we choose a similar delay ansatz for the coupling functions g_l in (20) as in Section “Dynamics based on a recursion formula”:

$$g_0(t) = w \cdot f_{\eta,\sigma}\left(\frac{1.5\,\Delta - t}{\Delta}\right), \tag{24}$$

$$g_l(\lambda_1, \ldots, \lambda_{l-1})(t) = w \cdot f_{\eta,\sigma}(\lambda_{l-1}(t - \Delta)), \quad l \geq 1, \tag{25}$$

for $t \geq 0$.

Here, the sigmoidal *logistic function* f with cut constant η and spread parameter σ is defined by

$$f_{\eta,\sigma}(\rho) = \frac{1}{1 + e^{-\sigma\cdot(\rho-\eta)}} \tag{26}$$

and w, η, σ and $\tau_l = \tau$ are real positive constants. For the case $\sigma = 0$ we use the jump function

$$f_{\eta,0}(\rho) = \begin{cases} 0, & \rho < \eta \\ 1, & \text{otherwise,} \end{cases} \tag{27}$$

which corresponds to the limit of $f_{\eta,\sigma}$ for $\sigma \to \infty$. The properties of the solutions to (20)–(27) depend strongly on σ. For $\sigma = 0$ it has singular points, for $\sigma > 0$ it is a smooth function.

Spherical wave functions

Let $\mathcal{V}_F = \mathrm{Span}(\mathbf{f}_k)_{1\leq k \leq n}$ be the n-dimensional space spanned by the (linearly independent) filler vectors $\mathbf{f}_k$, and $\mathcal{V}_R = \mathrm{Span}(|1\rangle, |2\rangle, |3\rangle)$ be the 3-dimensional space spanned by the “one-particle” roles (13).

Our approach relies upon a separation ansatz where the fillers are described by functions of time, while the roles are given by spherical harmonics at the unit sphere S. First, we identify the n fillers $\mathbf{f}_k$ with functions $f_k(t)$.

Next, we regard the tree in Fig. 3 as a “deformed” term schema for a spin-one triplet (Fig. 4).

Figure 4 indicates that the three role positions $|1\rangle$, $|2\rangle$, $|3\rangle$ in a labeled binary tree have been identified with the three z-projections of a spin-one particle:

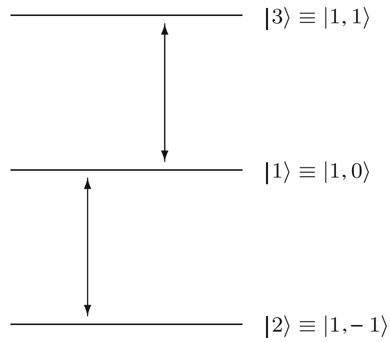

Fig. 4 Tree roles in a spin-one term schema

$$\begin{aligned}|2\rangle &\equiv |1,-1\rangle\\ |1\rangle &\equiv |1,0\rangle\\ |3\rangle &\equiv |1,1\rangle,\end{aligned} \tag{28}$$

which have an $L^2(S)$ representation by spherical harmonics

$$|j,m\rangle \cong Y_{jm}(\varphi,\vartheta), \quad \varphi \in [0,2\pi[\ ,\vartheta \in [0,\pi]. \tag{29}$$

In order to deal with complex phrase structure trees, we have to describe the tensor products of role vectors $|i\rangle \otimes |j\rangle = |ij\rangle$. Inserting the spin eigenvectors from (28), yields expressions like

$$|j_1,m_1\rangle \otimes |j_2,m_2\rangle \equiv |j_1,m_1,j_2,m_2\rangle, \tag{30}$$

well-known from the spin coupling in quantum mechanics (Edmonds 1957).

In quantum mechanics, the product states (30) generally belong to different multiplets, which are given by the irreducible representations of the spin algebra $sl(2)$. These are obtained by the Clebsch–Gordan coefficients in the expansions

$$|j,m,j_1,j_2\rangle = \sum_{m_1,m_2=m-m_1} \langle j_1,m_1,j_2,m_2|j,m,j_1,j_2\rangle |j_1,m_1,j_2,m_2\rangle, \tag{31}$$

where the total angular momentum j obeys the triangle relation

$$|j_1-j_2| \le j \le j_1+j_2. \tag{32}$$

However, our aim appears to be a bit different. Instead of computing the state of the coupled system by means of (31), we have to express the particular state vector (30) through higher harmonic wave functions. Therefore, we have to invert (31), leading to

$$|j_1,m_1,j_2,m_2\rangle = \sum_{j=|j_1-j_2|}^{j_1+j_2} \langle j,m,j_1,j_2|j_1,m_1,j_2,m_2\rangle |j,m,j_1,j_2\rangle, \tag{33}$$

with the constraint $m = m_1 + m_2$.

Equation (33) has to be applied recursively in order to obtain the role positions of more and more complex phrase structure trees. Finally, a single tree $s \in T$ is represented by its filler/role bindings in the basis of spherical harmonics

$$s(\varphi,\vartheta,t) = \sum_{jkm} a_{jkm} f_k(t)\, Y_{jm}(\varphi,\vartheta), \tag{34}$$

where the coefficients $a_{jkm} = 0$ if filler k is not bound to pattern Y_{jm}. Otherwise, the a_{jkm} encode the Clebsch–Gordan coefficients in Eq. (33).

Combining (34) with the order parameter ansatz (19), yields the spatio-temporal parsing dynamics

$$u(\varphi,\vartheta,t) = \sum_{l=0}^{L} \lambda_l(t) s_l(\varphi,\vartheta,t), \tag{35}$$

indicating a separation of time scales (Haken 1983): the fast functions $f_k(t)$ in $s_l(\varphi, \vartheta, t)$ encode instantaneous trees, while the transient evolution of the order parameters $\lambda_l(t)$ reflects the time course of the parsing process.

Let us illustrate this construction in the light of our example CFG G (1). The five fillers $f_1 = \mathtt{S}$, $f_2 = \mathtt{VP}$, $f_3 = \mathtt{Susan}$, $f_4 = \mathtt{ate}$, $f_5 = \mathtt{grass}$ are encoded by functions $f_k(t)$ with $k = 1, 2,\ldots, 5$.

The first state, $\mathtt{S}|1\rangle$, of the parse $\mathcal{U}$ (18) in Fock space is simply given by

$$s_0 = f_1(t)\, Y_{1,0}$$

since $|1\rangle \cong Y_{1,0}$. For the second state, $\mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle + \mathtt{VP}|3\rangle$, we straightforwardly obtain the representation

$$s_1 = f_1(t)\, Y_{1,0} + f_3(t)\, Y_{1,-1} + f_2(t)\, Y_{1,1}.$$

Only computing the third and final state, $\mathtt{S}|1\rangle + \mathtt{Susan}|2\rangle + \mathtt{VP}|13\rangle + \mathtt{ate}|23\rangle + \mathtt{grass}|33\rangle$, turns out to be somewhat cumbersome. In a first step, we get

$$\begin{aligned}s_3 = {}& f_1(t)\, Y_{1,0} + f_3(t)\, Y_{1,-1} + f_2(t)|1,0\rangle|1,1\rangle\\ &+ f_4(t)|1,-1\rangle|1,1\rangle + f_5(t)|1,1\rangle|1,1\rangle.\end{aligned}$$

Expressing the tensor products by (33), yields firstly

$$\begin{aligned}|1,0\rangle|1,1\rangle = |1,0,1,1\rangle &= \sum_{j=0}^{2} \langle j,1,1,1|1,0,1,1\rangle |j,1,1,1\rangle\\ &= \langle 0,1,1,1|1,0,1,1\rangle |0,1,1,1\rangle\\ &\quad + \langle 1,1,1,1|1,0,1,1\rangle |1,1,1,1\rangle\\ &\quad + \langle 2,1,1,1|1,0,1,1\rangle |2,1,1,1\rangle.\end{aligned}$$

The first Clebsch–Gordan coefficient $\langle 0, 1, 1, 1 \mid 1, 0, 1, 1\rangle = 0$ because a spin $j = 0$ particle cannot have an $m = 1$ projection. On the other hand, the Clebsch–Gordan coefficients are $\langle 1,1,1,1|1,0,1,1\rangle = -1/\sqrt{2}$ and $\langle 2,1,1,1|1,0,1,1\rangle = 1/\sqrt{2}$.

Correspondingly, we obtain for

$$\begin{aligned}|1,-1\rangle|1,1\rangle=|1,-1,1,1\rangle&=\sum_{j=0}^{2}\langle j,0,1,1|1,-1,1,1\rangle|j,0,1,1\rangle\\&=\langle 0,0,1,1|1,-1,1,1\rangle|0,0,1,1\rangle\\&\quad+\langle 1,0,1,1|1,-1,1,1\rangle|1,0,1,1\rangle\\&\quad+\langle 2,0,1,1|1,-1,1,1\rangle|2,0,1,1\rangle.\end{aligned}$$

Here, $m = 0$ is consistent with $j = 0, 1, 2$ such that all three terms have to be taken into account through $\langle 0,0,1,1|1,-1,1,1\rangle=1/\sqrt{3}, \langle 1,0,1,1|1,-1,1,1\rangle=-1/\sqrt{2}$, and $\langle 2,0,1,1|1,-1,1,1\rangle=1/\sqrt{6}$.

Finally, we consider

$$\begin{aligned}|1,1\rangle|1,1\rangle=|1,1,1,1\rangle&=\sum_{j=0}^{2}\langle j,2,1,1|1,1,1,1\rangle|j,2,1,1\rangle\\&=\langle 0,2,1,1|1,1,1,1\rangle|0,2,1,1\rangle\\&\quad+\langle 1,2,1,1|1,1,1,1\rangle|1,2,1,1\rangle\\&\quad+\langle 2,2,1,1|1,1,1,1\rangle|2,2,1,1\rangle.\end{aligned}$$

Obviously, only the last term contributes to the sum with $\langle 2, 2, 1, 1 | 1, 1, 1, 1\rangle = 1$ for the same reason as above.

Summarizing, the parse U in Fig. 2, possessing the abstract Fock space representation (18), translates into the time-discrete dynamics

$$\begin{aligned}&s_0 = f_1(t)\,Y_{1,0}\\ \to\ &s_1 = f_1(t)\,Y_{1,0}+f_3(t)\,Y_{1,-1}+f_2(t)\,Y_{1,1}\\ \to\ &s_2 = f_1(t)\,Y_{1,0}+f_3(t)\,Y_{1,-1}+\frac{f_2(t)}{\sqrt{2}}\left(Y_{2,1}-Y_{1,1}\right)\\&\quad+f_4(t)\left(\frac{1}{\sqrt{3}}Y_{0,0}-\frac{1}{\sqrt{2}}Y_{1,0}+\frac{1}{\sqrt{6}}Y_{2,0}\right)+f_5(t)\,Y_{2,2}.\end{aligned} \tag{36}$$

Dynamic fields

Dynamic field theories (DFT) are a phenomenological account for continuum models in the cognitive sciences (Erlhagen and Schöner 2002; Schöner and Thelen 2006; Thelen et al. 2001). They are mathematically equivalent to neural field theories (Wilson and Cowan 1973; Amari 1977; Jirsa and Haken 1996; Coombes et al. 2003; Jirsa 2004; Wright et al. 2004; beim Graben 2008), yet not referring to a particular neurophysiological description but rather to dynamics in abstract *feature space*.

In order to obtain such dynamic fields, we bring (36) together with (35) to generate the time-continuous parsing dynamics:

$$\begin{aligned}u(\varphi,\vartheta,t)=&\lambda_0(t)f_1(t)Y_{1,0}\\&+\lambda_1(t)\left[f_1(t)Y_{1,0}+f_3(t)Y_{1,-1}+f_2(t)Y_{1,1}\right]\\&+\lambda_2(t)\left[f_1(t)Y_{1,0}+f_3(t)Y_{1,-1}+\frac{f_2(t)}{\sqrt{2}}\left(Y_{2,1}-Y_{1,1}\right)\right.\\&\left.+f_4(t)\left(\frac{1}{\sqrt{3}}Y_{0,0}-\frac{1}{\sqrt{2}}Y_{1,0}+\frac{1}{\sqrt{6}}Y_{2,0}\right)+f_5(t)Y_{2,2}\right].\end{aligned} \tag{37}$$

In order to complete our description, we have to determine the filler functions $f_k(t)$ in (37). One possible choice is to assign eigenfrequencies

$$\omega_k = \frac{2\pi k}{n} \tag{38}$$

to the five fillers $f_1 = \mathtt{S}$, $f_2 = \mathtt{VP}$, $f_3 = \mathtt{susan}$, $f_4 = \mathtt{ate}$, $f_5 = \mathtt{grass}$, such that the fillers become represented as harmonic oscillations

$$f_k(t) = \mathrm{e}^{\mathrm{i}\omega_k t} \tag{39}$$

in time.

Then, the parse U in Fig. 2, possessing the Fock space representation (37), translates into

$$\begin{aligned}u(\varphi,\vartheta,t)=&\lambda_0(t)\mathrm{e}^{\mathrm{i}\omega_1 t}Y_{1,0}\\&+\lambda_1(t)\left[\mathrm{e}^{\mathrm{i}\omega_1 t}Y_{1,0}+\mathrm{e}^{\mathrm{i}\omega_3 t}Y_{1,-1}+\mathrm{e}^{\mathrm{i}\omega_2 t}Y_{1,1}\right]\\&+\lambda_2(t)\left[\mathrm{e}^{\mathrm{i}\omega_1 t}Y_{1,0}+\mathrm{e}^{\mathrm{i}\omega_3 t}Y_{1,-1}+\frac{\mathrm{e}^{\mathrm{i}\omega_2 t}}{\sqrt{2}}\left(Y_{2,1}-Y_{1,1}\right)\right.\\&\left.+\mathrm{e}^{\mathrm{i}\omega_4 t}\left(\frac{1}{\sqrt{3}}Y_{0,0}-\frac{1}{\sqrt{2}}Y_{1,0}+\frac{1}{\sqrt{6}}Y_{2,0}\right)+\mathrm{e}^{\mathrm{i}\omega_5 t}Y_{2,2}\right].\end{aligned} \tag{40}$$

Simulations

The stationary waves representing the three parse steps s_0, s_1 and s_2 of $\mathcal{U}$ in (36) are shown in Fig. 5(a–c) as a sequence of snapshots, respectively.[1]

Note that the initial state s_0 is given by the constant "p_z orbital" (Fig. 5a).

In the representation of s_1 we have a superposition of the constant function from s_0 with two higher harmonics, indicating the fillers $f_2(t)$ and $f_3(t)$ assigned to the tree position roles r_3 and r_2, respectively (Fig. 5b). The final state s_2 is given by an even more involved oscillation (Fig. 5c).

Additionally, we present in Fig. 6. the dynamics of $|s_1|$ (Fig. 5b) with higher temporal resolution.

Now, the first column of Fig. 6. corresponds to the first six images in the row of Fig. 5b; the toroidal dynamics accounted for by the fillers $f_2(t)$ and $f_3(t)$ is clearly visible.

[1] Animations for our simulations are available as supplementary material online.

Fig. 5 Snapshot sequences of the moduli of the stationary waves $|s_1|$ (36). (**a**) for the initial state s_0, (**b**) for s_1, (**c**) for s_2

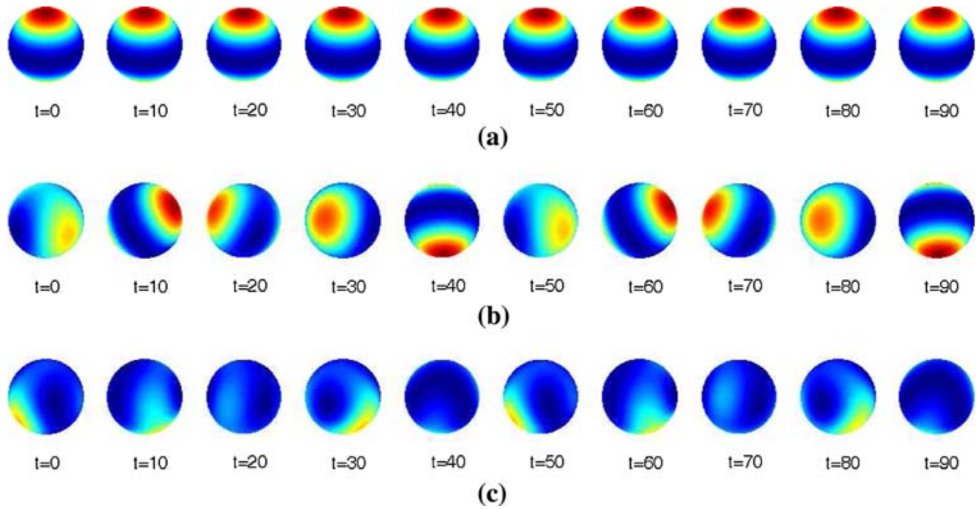

Fig. 6 Snapshot sequence of the state $|s_1|$ with higher temporal resolution

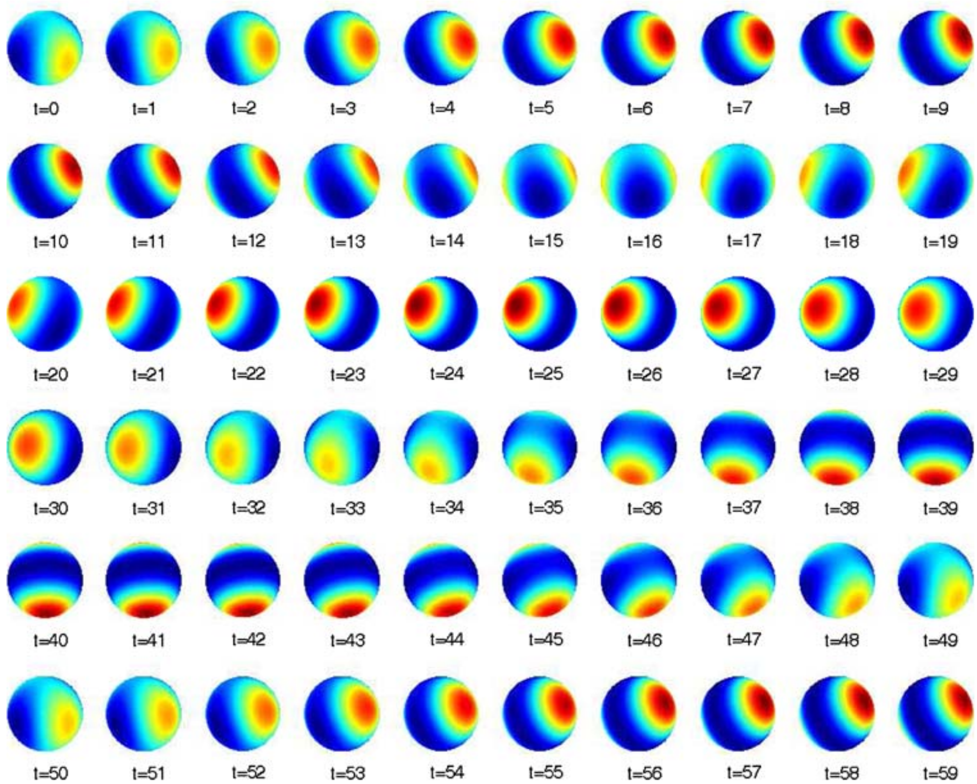

Figure 7 displays the temporal evolution of the three order parameters $\lambda_l(t)$, denoting the amplitudes of the corresponding parse states s_l, according to the order parameter equation (20) with the coupling functions g_l (24). For the numerical simulation of the neural order parameter equation we used Euler's method (Kress 1998, Chap. 10).

Figure 7 reveals that the parse states s_0, s_1, and s_2 are fully established at times $T_0 \approx 50$, $T_1 \approx 150$, and $T_2 \approx 250$, where the order parameters assume their respective local maxima.

Finally, Fig. 8 presents the transient evolution of the dynamic parse field $|u(\varphi, \vartheta, t)|$ according to (40) as a sequence of snapshots.

Comparing Fig. 8 with Fig. 5 shows that the parse states s_0, s_1, and s_2 become established at times $T_0 \approx 50$, $T_1 \approx 150$, and $T_1 \approx 250$ in accordance with Fig. 7.

Discussion

In Section "Introduction" we have raised the question: Can the basic assumptions and mechanisms underpinning a language processing model be expressed in terms that are compatible with viable models of the brain? The basic assumptions and mechanisms of linguistics are that

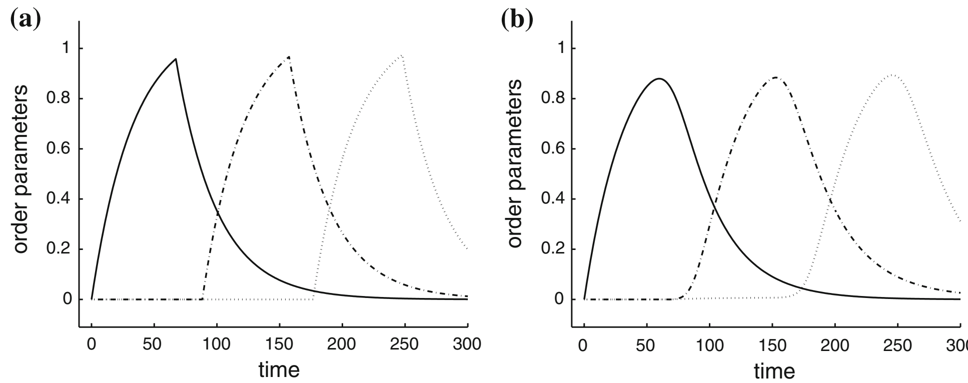

Fig. 7 Dynamics of the three order parameters $\lambda_0(t)$ (*solid*), $\lambda_1(t)$ (*dashed-dotted*), and $\lambda_2(t)$ (*dotted*), generated via (20)–(27). (**a**) for $\sigma = 0$, (**b**) for $\sigma = 10$. The calculations have been carried out with the parameter choices $\eta = 0.5$ and delay time $\Delta = 70$

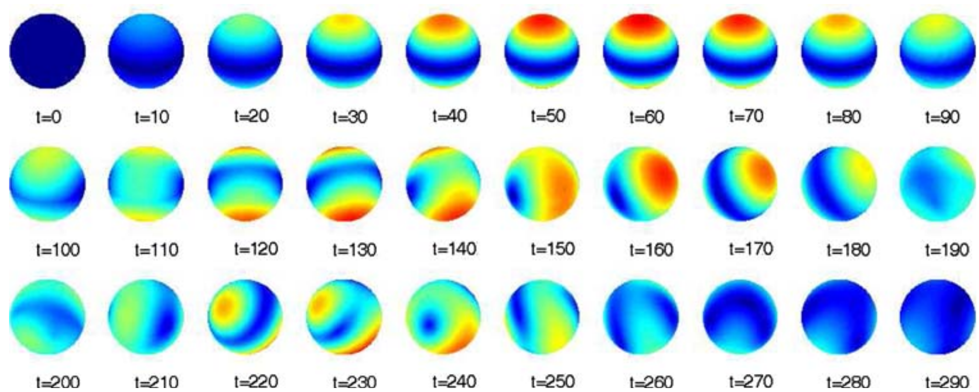

Fig. 8 Snapshot sequence of the dynamic field $|u|$ representing the transient evolution of the parse in Fock space

symbolic expressions are complex and recursive hierarchical data structures which are symbolically processed by appropriate cognitive architectures (Chomsky 1957; Hopcroft and Ullman 1979; Lewis and Vasishth 2006). On the other hand, macroscopic brain function is often modeled in terms of neural field dynamics (Wilson and Cowan 1973; Amari 1977; Jirsa and Haken 1996; Coombes et al. 2003; Jirsa 2004; Wright et al. 2004; Erlhagen and Schöner 2002; Schöner and Thelen 2006; Thelen et al. 2001).

In order to respond to that question, we have constructed a faithful mapping from linguistic phrase structure trees to the infinite-dimensional function space of spherical wave functions, using Smolensky's filler/role bindings and tensor product representations (Smolensky 1990, 2006; Smolensky and Legendre 2006). The abstract feature space of this representation is the unit sphere S. Language processing is than described by the transient evolution of spherical patterns, governed by order parameter equations for their respective amplitudes (Haken 1983). For the order parameter equations we chose a neural leaky integrator model with delayed coupling between the parse states.

Since spherical harmonics are also often employed in analyzing electroencephalographic brain waves (Nunez and Srinivasan 2006), it is tempting to simply identify the feature space S of our model with a spherical head model, thereby interpreting the dynamic field $u(\varphi, \vartheta, t)$ of the parser with the actual voltage distribution across the human scalp. However, such a straightforward interpretation is not tenable as it would require the whole brain to be in only a few representative states necessary to maintain one particular cognitive task. This is obviously not the case. Therefore we would need another mapping from the abstract feature space representation of our model to a neural representation in the brain in order to answer the question in the end. We shall leave this issue for future research on the cognitive neurodynamics of brain waves.

References

Amari SI (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern 27:77–87

Anderson JR (1995) Cognitive psychology and its implications, 4th edn. W. H. Freeman and Company, New York

beim Graben P (2008) Foundations of neurophysics. In: beim Graben P, Zhou C, Thiel M, Kurths J (eds) Lectures in supercomputational neuroscience: dynamics in complex brain networks. Springer Complexity Series. Springer, Berlin, pp 3–48

beim Graben P, Kurths J (2008) Simulating global properties of electroencephalograms with minimal random neural networks. Neurocomputing 71(4):999–1007

beim Graben P, Gerth S, Saddy D, Potthast R (2007) Fock space representations in neural field theories. In: Biggs N, Bonnet-

Bendhia AS, Chamberlain P, Chandler-Wilde S, Cohen G, Haddar H, Joly P, Langdon S, Lunéville E, Pelloni B, Potherat D, Potthast R (eds) Proceedings of Waves 2007. The 8th international conference on mathematical and numerical aspects of waves. Department of Mathematics, University of Reading, Reading, pp 120–122
beim Graben P, Gerth S, Vasishth S (2008) Towards dynamical system models of language-related brain potentials. Cogn Neurodynam. doi:10.1007/s11571-008-9041-5
Bornkessel I, Zysset S, Friederici AD, von Cramon DY, Schlesewsky M (2005) Who did what to whom? The neural basis of argument hierarchies during language comprehension. NeuroImage 26:221–233
Chomsky N (1957) Syntactic structures. Mouton, Den Haag
Christiansen MH (1992) The (non)necessity of recursion in natural language processing. In: Proceedings of the 14th annual conference on Cognitive Science Society, Lawrence Erlbaum Associates, Hillsdale, pp 665–670
Christiansen MH, Chater N (1999) Toward a connectionist model of recursion in human linguistic performance. Cogn Sci 23(4):157–205
Coombes S, Lord G, Owen M (2003) Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. Physica D 178:219–241
Drenhaus H, beim Graben P, Saddy D, Frisch S (2006) Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. Brain Lang 96(3):255–268
Edmonds AR (1957) Angular momentum in quantum mechanics. Princeton University Press, Princeton
Elman JL (1995) Language as a dynamical system. In: Port RF, van Gelder T (eds) Mind as motion: explorations in the dynamics of cognition. MIT Press, Cambridge, pp 195–223
Erlhagen W, Schöner G (2002) Dynamic field theory of movement preparation. Psychol Rev 109(3):545–572
Everett DL (2005) Cultural constraints on grammar and cognition in Pirahã – another look at the design features of human language. Curr Anthropol 46(4):621–646
Friederici AD (1995) The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. Brain Lang 50:259–281
Garagnani M, Wennekers T, Pulvermuller F (2007) A neuronal model of the language cortex. Neurocomputing 70:1914–1919
Haag R (1992) Local quantum physics: fields, particles, algebras. Springer, Berlin
Hagoort P (2005) On Broca, brain, and binding: a new framework. Trends Cogn Sci 9(9):416–423
Haken H (1983) Synergetics. An introduction, Springer series in synergetics, vol 1, 3rd edn. Springer, Berlin. 1st edition 1977
Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: What is it, who has it, and how did it evolve? Science 298:1569–1579
Hopcroft JE, Ullman JD (1979) Introduction to automata theory, languages, and computation. Addison–Wesley, Menlo Park
Jirsa VK (2004) Information processing in brain and behavior displayed in large-scale scalp topographies such as EEG and MEG. Int J Bifurcat Chaos 14(2):679–692
Jirsa VK, Haken H (1996) Field theory of electromagnetic brain activity. Phys Rev Lett 77(5):960–963
Kempson R, Meyer-Viol W, Gabbay D (2001) Dynamic syntax. The flow of language understanding. Blackwell
Kress R (1998) Numerical analysis, graduate texts in mathematics, vol 181. Springer-Verlag, New York
Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207:203–205
Kutas M, Hillyard SA (1984) Brain potentials during reading reflect word expectancy and semantic association. Nature 307:161–163
Lewis RL, Vasishth S (2006) An activation-based model of sentence processing as skilled memory retrieval. Cogn Sci 29:375–419
Maye A, Werning M (2004) Temporal binding of non-uniform objects. Neurocomputing 58:941–948
Mizraji E (1989) Context-dependent associations in linear distributed memories. Bull Math Biol 51(2):195–205
Mizraji E (1992) Vector logics: the matrix-vector representation of logical calculus. Fuzzy Sets Syst 50:179–185
Moore C, Crutchfield JP (2000) Quantum automata and quantum grammars. Theor Comput Sci 237:275–306
Noveck IA, Posada A (2003) Characterizing the time course of an implicature: an evoked potentials study. Brain Lang 85:203–210
Nunez PL, Srinivasan R (2006) Electric fields of the brain: the neurophysics of EEG. Oxford University Press, New York
Osterhout L, Holcomb PJ (1992) Event-related brain potentials elicited by syntactic anomaly. J Mem Lang 31:785–806
Schöner G, Thelen E (2006) Using dynamic field theory to rethink infant habituation. Psychol Rev 113(2):273–299
Smolensky P (1990) Tensor product variable binding and the representation of symbolic structures in connectionist systems. Artif Intell 46:159–216
Smolensky P (2006) Harmony in linguistic cognition. Cogn Sci 30:779–801
Smolensky P, Legendre G (2006) The harmonic mind. From neural computation to optimality-theoretic grammar, vol 1: cognitive architecture. MIT Press, Cambridge
Stabler EP (1997) Derivational minimalism. In: Retoré C (ed) Logical aspects of comutational linguistics, Springer Lecture Notes in Computer Science, vol 1328. Springer, New York, pp 68–95
Tabor W, Juliano C, Tanenhaus MK (1997) Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. Lang Cogn Process 12(2/3):211–271
Thelen E, Schöner G, Scheier C, Smith LB (2001) The dynamics of embodiment: A field theory of infant perseverative reaching. Behav Brain Sci 24:1–86
van der Velde F, de Kamps M (2006) Neural blackboard architectures of combinatorial structures in cognition. Behav Brain Sci 29:37–108
Vosse T, Kempen G (2000) Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. Cognition 75:105–143
Werning M, Maye A (2007) The cortical implementation of complex attribute and substance concepts: synchony, frames, and hierarchical binding. Chaos Complex Lett 2(2/3):435–452
Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13:55–80
Wright JJ, Rennie CJ, Lees GJ, Robinson PA, Bourke PD, Chapman CL, Gordon E, Rowe DL (2004) Simulated electrocortical activity at microscopic, mesoscopic and global scales. Int J Bifurcat Chaos 14(2):853–872