

# Evaluating Hebbian Self-Organizing Memories for Lexical Representation and Access

Claudia Marzi\*<sup>+</sup>, Marcello Ferro\*, Claudia Caudai\*, Vito Pirrelli\*

\*Institute for Computational Linguistics, “Antonio Zampolli” - National Research Council (CNR)

<sup>+</sup>University of Pavia – Dept. of Theoretical and Applied Linguistics

E-mail: {claudia.marzi, marcello.ferro, claudia.caudai, vito.pirrelli}@ilc.cnr.it

## Abstract

The lexicon is the store of words in long-term memory. Any attempt at modelling lexical competence must take issues of string storage seriously. In the present contribution, we discuss a few desiderata that any biologically-inspired computational model of the mental lexicon has to meet, and detail a multi-task evaluation protocol for their assessment. The proposed protocol is applied to a novel computational architecture for lexical storage and acquisition, the “Topological Temporal Hebbian SOMs” (T2HSOMs), which are grids of topologically organised memory nodes with dedicated sensitivity to time-bound sequences of letters. These maps can provide a rigorous and testable conceptual framework within which to provide a comprehensive, multi-task protocol for testing the performance of Hebbian self-organising memories, and a comprehensive picture of the complex dynamics between lexical processing and the acquisition of morphological structure.

Keywords: Mental Lexicon, Morphology Acquisition, Self-Organizing Maps

## 1. Introduction

The mental lexicon is the store of words in long-term memory, where words are coded as time series of sounds/letters. From this perspective, the question of word coding, storage and maintenance in time is not separable from the issue of how words are accessed and processed. The rich cognitive literature on short-term and long-term memory processes (Miller 1956; Baddeley & Hitch 1974; Baddeley 1986, 2006; Henson 1998; Cowan 2001, among others) has the unquestionable merit of highlighting some fundamental issues of coding, maintenance and manipulation of time-bound constraints over strings of symbols (letters or phonological segments), thus shedding light on well-known aspects of lexical organisation such as primacy and recency effects in lexical access (Aitchison 1987; Gupta et al. 2005), concurrent activation of families of inflectionally and derivationally-related forms (Baayen et al. 1997; Taft 1979; Hay 2001; Ford et al. 2003; Lüdeling & De Jong 2002; Moscoso del Prado Martín et al. 2004), word confusability and alignment in visual recognition (Bowers & Davis 2009; Davis 2010). In spite of their important status, lexical coding issues have nonetheless suffered some neglect by the NLP research community. Both symbolic and connectionist answers to the problem provide ad-hoc built-in devices for string representation (e.g. arrays of characters or conjunctive coding schemata) rather than a principled solution to the problem of lexical storage and its cognitive implications for word access and processing (Sibley et al. 2008). Moreover, to our knowledge, there have been no systematic attempts at evaluating how well lexical encoding schemata meet the requirements of contemporary linguistic and psycholinguistic memory models of the mental lexicon.

## 2. Background

According to the dual-route approach to word processing (Clahsen 1999; Prasada & Pinker 1993; Pinker & Prince 1988; Pinker & Ullman 2002), recognition of a morphologically complex input word involves a preliminary full

form access to the lexicon, and the optional morpheme-based access of sub-word constituents, resulting from application of morphological rules of on-line word processing to the input word. If and only if access to full-form fails to find any matching entry in the lexicon, the second step is taken. The approach endorses a direct functional correspondence between principles of grammar organisation (lexicon vs. rules), processing correlates (storage vs. computation) and localization of the cortical areas functionally involved in word processing (temporo-parietal vs. frontal areas: Ullman 2004).

Over the past decades, the psycholinguistic literature has shed novel light on these issues, and suggested that surface word relations constitute a fundamental domain of morphological competence, with particular emphasis on the interplay between form frequency, family frequency and family size effects within morphologically-based word families such as inflectional paradigms (Baayen et al. 1997; Taft 1979; Hay 2001; Ford et al. 2003; Lüdeling & De Jong 2002; Moscoso del Prado et al. 2004; Stemberger & Middleton 2003; Tabak et al. 2005). However, that more than just lexical storage is involved is suggested by interference effects between false morphological friends (or pseudo-derivations) such as *broth* and *brother*, sharing a conspicuous word onset but unrelated morphologically (Frost et al. 1997; Rastle et al. 2004; Post et al. 2008). The evidence shows that as soon as a given letter sequence is fully decomposable into morphological formatives, word parsing takes place automatically, prior to (or concurrently with) lexical look-up. This suggests that differentiated brain areas devoted to language maximise the opportunity of using both general and specific information simultaneously (Libben 2006; Post et al. 2008), rather than maximize processing efficiency and economy of storage.

Computer models have been successful in tackling different aspects of word structure, but have not been able to provide, to date, a comprehensive picture of the complex dynamics between computation and storage in word processing. The currently emerging view sees word processing as the outcome of simultaneously activating

patterns of cortical connectivity reflecting (possibly redundant) distributional regularities in the input at the phonological, morpho-syntactic and morpho-semantic levels. At the same time, there is evidence to argue for a more complex and differentiated neuro-biological substrate for human language than connectionist one-route models can posit.

T2HSOMs adhere to such a dynamic, non modular view of the interaction between memory and computation, whereby word processing and learning are primarily conceived of as memory-driven processes. They part from both dual-route and one-route approaches in supporting the view that the way words are structured in our long-term memory is key to understanding the mechanisms governing word processing. This perspective focuses on word productivity as the by-product of more basic memory processes that must independently be assumed to account for word learning. Secondly, it opens up new promising avenues of inquiry by tapping the large body of literature on short-term and long-term memories for serial order (see Baddley 2007, for an overview). Furthermore, it gives the opportunity of using sophisticated computational models of language-independent memory processes (Botvinick and Plaut 2006; Brown et al. 2000, among others) to shed light on language-specific aspects of word encoding and storage.

### 3. Topological Temporal SOMs

T2HSOMs are Self-Organizing Maps (SOMs; Kohonen 2001) augmented with first-order re-entrant Hebbian connections that encode probabilistic expectations over letter strings (Koutnik 2007; Pirrelli et al. 2011; Ferro et al. 2010, 2011a). T2HSOMs consist of grids of topologically organised memory nodes which learn stored representations of time-bound input stimuli through training, developing dedicated sensitivity to classes of stimuli. Upon presentation of an input letter, all map nodes are activated synchronously, but only the most highly activated one, the so-called Best Matching Unit (BMU), wins over the others. The activation equation (1) of node  $i$  at time  $t$  is the sum of two functions, corresponding to how close the stored representation in node  $i$  is to the current input ( $y_{S,i}(t)$ ), and how expected it is on the basis of the input shown at the previous time tick ( $y_{T,i}(t)$ ) (Ferro et al. 2011b). Parameters  $\alpha$  and  $\beta$  in (1) weigh up the contribution of the two functions to the overall activation score. By giving more weight to  $\alpha$  we make the map more sensitive to the specific *code* of the current input character (i.e. whether it is an ‘A’, or a ‘B’ or whatever). Comparatively higher values of  $\beta$  make the map more sensitive to the timing of the current input, i.e. its position in the string.

$$(1) \quad y_i(t) = \alpha \cdot y_{S,i}(t) + \beta \cdot y_{T,i}(t).$$

In the learning mode, a BMU adjusts its stored representation to the current letter code and timing, and propagates adjustment to neighbouring nodes. How far propagation goes depends on the propagation radius  $r$ , which decreases as learning progresses. Local propagation causes the development of topologically connected clusters of

nodes that are sensitive to the same letter type.

When a string of letters making up a word form is presented to the map one character at a time, a temporal chain of BMUs is activated. Figure 1 illustrates two such temporal chains, triggered by the German verb forms *gemacht* and *gelacht* (‘made’ and ‘laughed’, past participle) shown to a 20x20 nodes map trained on 30 verb paradigms, sampled from the CELEX German database by decreasing values of cumulative paradigm frequency. In the figure, each node is labelled with the letter the node is most sensitive to. Pointed arrows represent temporal connections linking two consecutively activated nodes, thus depicting the temporal sequence of node activation, starting from the beginning-of-word symbol ‘#’ (anchored in the top left corner of the map) and ending to ‘\$’. Activation chains allow us to inspect the memory structures that a map develops through training (see section 3.1). Moreover, they exhibit a straightforward correlation between morphological segmentation and topological organisation of BMUs on the map (see section 3.2).

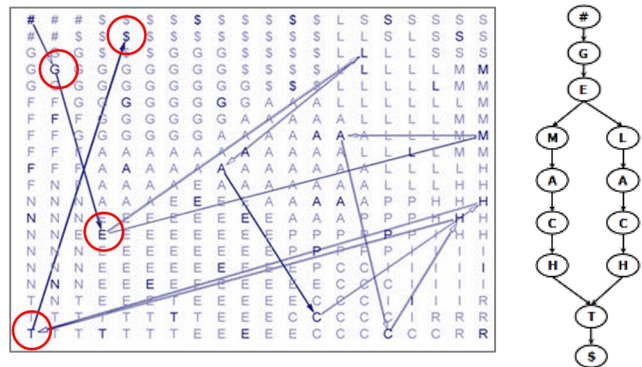


Figure 1: BMU activation chains for gemacht-gelacht (left) and their word-graph representation (right).

#### 3.1 Memory structures and memory orders

Through repeated exposure to word forms encoded as sequences of letters, a T2HSOM tends to dynamically store strings as graph-like hierarchical structures of nodes, such as the one in Figure 1 above, where arrows represent weighted Hebbian connections. Note that the two *a*'s in *gemacht* and *gelacht* activate two different BMUs, as they are temporally preceded by different letters (*m* and *l* respectively). This is an important bias of T2HSOMs, tending to provide dedicated memory resources to identical symbols that happen to be embedded in different contexts. In fact, it can be observed that the map tends to propagate its first-order Markovian expectations. Since the two *a*'s in *gemacht* and *gelacht* trigger different nodes, the latter nodes in turn project different expectations on the upcoming BMUs. As a result, two different nodes are recruited for the ensuing *c*'s and *h*'s in both verb forms.

The length of the chain of dedicated nodes illustrates different memory strategies of the map. It can be shown that this type of organisation maximises the map's expectation of an upcoming symbol in the input string or, equivalently, minimises the entropy over the set of transition probabilities from each BMU to the ensuing one.

This prompts a process of incremental specialisation of memory resources, whereby several nodes are recruited to be sensitive to contextually specific occurrences of the same symbol.

The tendency to store a word form through a uniquely dedicated chain of BMUs, however, is a function of i) the size of the map (i.e. the number of available nodes), and ii) the map's ability to train two adjacent nodes independently (or granularity of the map), defined by the value of the propagation radius. Generally speaking, shortage of memory resources leads to shorter memory orders and a resulting effect of memory compression.

### 3.2 Morphological structure, lexical processing and storage

From a lexical standpoint, the most interesting consequence of memory compression is that T2HSOMs exhibit a straightforward correlation between morphological segmentation and topological organisation of BMUs. Word forms sharing sub-lexical constituents tend to trigger chains of identical or neighbouring nodes. In short, topological distance (proximity) on the map correlates with morphological similarity. In traditional approaches to word segmentation, this is equivalent to aligning morphologically-related word forms by morphological structure. Since chains of activated nodes encode time sequences of symbols, T2HSOMs can be said to enforce alignment through synchrony. For example, German past participles provide a case of discontinuous morphological structure. Turning back to the word-graph in Figure 1, *gemacht* and *gelacht*, after sharing the same sequence of BMUs for *ge-*, part on the roots *mach-* and *lach-* to meet again upon recognition of the suffix *-t*.

This behaviour is modulated by the map parameters: keeping  $\beta$  constant (e.g. 1.0), lower  $\alpha$  values (e.g. 0.087) prompt a greater sensitivity to positional coding of letters, with a tendency to assign different BMUs to the same symbol when shown in different contexts. The left-hand side of Figure 2 shows the memory structures for German *gesprochen* ('spoken'), *gesehen* ('seen'), *gesagt* ('said') and *gefragt* ('asked') on a map with positional coding, with a clear bias towards developing separate activation chains. Higher  $\alpha$  values (e.g. 0.5) considerably affect this bias, as shown by the word graph to the right of Figure 2, where shared past participle endings activate identical BMUs. The interesting implication of this behaviour is that self-organising principles of lexical storage can go a long way in enforcing alignment between morphologically-related chains of activation on a map. Morphologically-motivated alignment is known to be an important, preliminary step in the child's bootstrapping of morphological information proper (Pirrelli & Herreros 2007).

## 4. Evaluating T2HSOMs

As Hebbian SOMs are, by their very nature, memory devices, they can accordingly be probed on several memory-related tasks. Nodes in a SOM can i) store information, ii) fire concurrently upon presentation of an input stimulus, iii) compete for activation primacy, iv) adjust

their memory store for its content to become increasingly more similar to a particular class of stimuli, v) cluster in topologically connected areas of nodes that tend to be sensitive to the same stimuli. Accordingly, we can evaluate a T2HSOM either intrinsically, for its capacity to re-code and recall stored information, or extrinsically, for its capacity to organise stored information into morphologically-sensitive structures.

In this section, intrinsic evaluation will cover two tasks: word activation and word recall. Word activation is defined as a function of the chain of BMUs which are triggered by an input word. The map's activation state is accurate if the information stored by each BMU at time  $t$  matches the input stimulus shown to the map at the same time tick. Given an input word, its activation chain is accurate if every input character activates a node storing that character. Otherwise, the activation chain is not accurate. The task thus probes the map's ability to re-code an input word form accurately, by assigning each incoming symbol its appropriate category. Word recall, on the other hand, defines the map's capacity to reinstate, in the appropriate order, the sequence of BMUs already activated by an input word  $w$ , on the basis of the integrated pattern of node activation triggered by  $w$ . The task requires that word activation is accurate but it also taps the map's expectation of upcoming characters given the characters that were already recalled. Word recall thus probes the notion of lexicality or lexical familiarity, and ultimately the map's ability to discriminate between stored and non-stored forms.

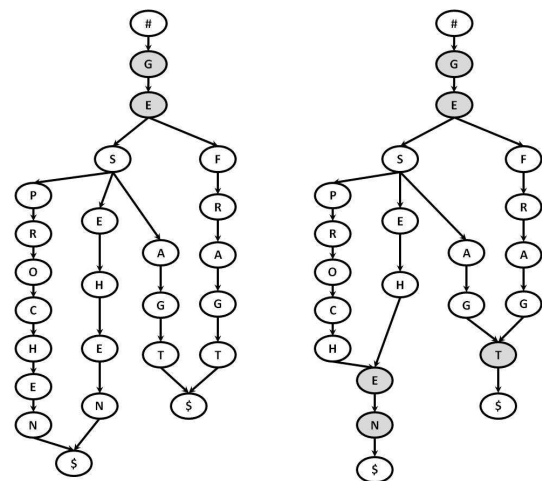


Figure 2: Word-graph representations on a temporal (left) and spatio-temporal map (right) of German past participles.

Extrinsic evaluation, on the other hand, is intended to assess to what extent memory structures mirror principles of morphological organisation, by calculating how well topological proximity of recoded characters correlates with morphological identity. This can be done both intra-paradigmatically, by assessing the consistency of regular and sub-regular paradigms, and inter-paradigmatically, by assessing how uniformly paradigm cells are realised through inflectional markers.

#### 4.1 Experimental design, materials and results

Two sets of 50 German verb paradigms and 50 Italian verb paradigms were sampled by decreasing values of cumulative paradigm frequency from the German Celex database (Baayen et al. 1995) and from the Italian Treebank (Montemagni et al. 2003) respectively. For both languages, 15 inflected forms were selected from each paradigm: the infinitive (*i*), present participle (*pA*, Italian “gerundio”), past participle (*pE*), the six forms of the present indicative (*SIE-PIE*) and the six forms of the past tense (*SKA-PKA*, “präteritum” and “passato remoto”). Each verb form was input as a letter string preceded by ‘#’ and followed by ‘\$’: e.g. ‘#ISTS\$’ for *ist*. All letters common to the German and Italian alphabets were written in upper-case. Umlauted characters were written as lower-case digraphs (e.g. ‘#KoeNNEN\$’ for *können*) and the sharp s ‘ß’ as ‘ss’ (e.g. ‘#HEIssENS\$’ for *heißen*). In both cases, pairs of lower-case letters were processed as one symbol. Letters were represented on the input layer as mutually orthogonal, binary vector codes. Identical letter codes were used for upper-case letters in both German and Italian.

Two configurations of a 40x40 T2HSOM were trained on the selection of 750 German verb forms, using different values for  $\alpha$  and  $\beta$  in equation (1) above. The configurations are hereafter referred to as *Temporal* ( $\alpha=0.087$ ,  $\beta=1$ ) and *Spatio-Temporal* ( $\alpha=0.5$ ,  $\beta=1$ ) respectively, to make it clear that the map is more sensitive to the contextual timing of input characters (temporal) or to their coding (spatio-temporal). Likewise, 750 Italian verb forms were used to train a temporal and a spatio-temporal 40x40 map. Overall, we trained four instances of each map configuration: German temporal, German spatio-temporal, Italian temporal, Italian spatio-temporal. At each training epoch, all forms were shown five times each, to maximise random order effects of input presentation. All map instances were trained for 100 epochs.

##### Experiment 1: word activation

The task is intended to assess to what extent the map can recode an input stimulus accurately. Each input character at time  $t$  is said to be recoded accurately by the map if it matches  $L(\text{BMU}_t)$ , i.e. the label associated with the BMU at time  $t$ . For any map node  $i$  its label  $L_i$  is given by the character  $c$  which  $i$  is most sensitive to. A word is taken to accurately be recoded if all its letters match the labels of the corresponding BMUs. Results are given in Figure 3.

##### Experiment 2: lexical recall

After Baddeley, immediate lexical recall is modelled as the task of reinstating a word form soon after being exposed to it, an experimental protocol highlighting the dynamic interaction between short-term integration/sustainment of memory traces and long-term storage of lexical information. Accordingly, a map is first exposed to an input word  $w$  of length  $n_w$ . Its resulting integrated activation pattern  $\hat{Y}=\{\hat{y}_1, \dots, \hat{y}_N\}$ , with

$$(2) \quad \hat{y}_i = \max_{t=2, \dots, n_w} \{y_i(t)\},$$

is input to the same map ( $n_w-1$ ) times. At each time, the

map’s recall accuracy is calculated according to the activation function. A word is taken to accurately be recalled if all its letters match the labels of the corresponding BMUs. The protocol is thus intended to assess how well the map can output the appropriate sequence of symbols making up  $w$  upon presentation of the whole activation pattern triggered by  $w$ . Results are reported in Figure 3.

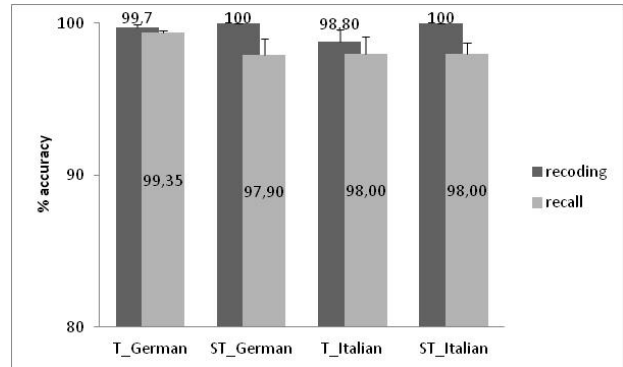


Figure 3: Activation/recall accuracy on German and Italian data, averaged over multiple instances of Temporal (T) and Spatio-Temporal (ST) maps.

##### Experiment 3: word alignment

The task is intended to assess how well the map develops topologically adjacent activation chains upon seeing forms that share morphological structure either intra-paradigmatically (e.g. root-sharing forms of the same verb) or inter-paradigmatically (e.g. affix-sharing forms belonging to different paradigms). This can be stated as a general string alignment problem arising whenever known symbol patterns are presented in novel arrangements, as when human speakers are able to recognise the English word *book* in *handbook*, or the German word *kommen* (‘come’) and *bekommen* (‘become’).

On a trained map, alignment between two words can be expressed analytically in terms of the topological distance between their corresponding activation chains. Figure 4 gives the per-node topological distance of the BMU chains of Italian *vediamo* (‘we see’) and *crediamo* (‘we believe’) on a spatio-temporal 40x40 Italian map.

|    | #    | C    | R    | E    | D    | I    | A    | M    | O    | \$   |
|----|------|------|------|------|------|------|------|------|------|------|
| #  | 0.00 | 0.19 | 0.45 | 0.59 | 0.35 | 0.26 | 0.38 | 0.38 | 0.37 | 0.53 |
| V  | 0.24 | 0.18 | 0.46 | 0.52 | 0.23 | 0.29 | 0.33 | 0.27 | 0.39 | 0.44 |
| E  | 0.21 | 0.19 | 0.37 | 0.39 | 0.34 | 0.18 | 0.31 | 0.36 | 0.29 | 0.48 |
| D  | 0.33 | 0.26 | 0.46 | 0.46 | 0.03 | 0.33 | 0.30 | 0.18 | 0.41 | 0.35 |
| I  | 0.26 | 0.24 | 0.31 | 0.48 | 0.36 | 0.00 | 0.28 | 0.37 | 0.23 | 0.47 |
| A  | 0.38 | 0.33 | 0.29 | 0.33 | 0.33 | 0.28 | 0.00 | 0.30 | 0.28 | 0.32 |
| M  | 0.38 | 0.32 | 0.46 | 0.42 | 0.18 | 0.37 | 0.30 | 0.00 | 0.44 | 0.30 |
| O  | 0.37 | 0.35 | 0.21 | 0.44 | 0.44 | 0.23 | 0.28 | 0.44 | 0.00 | 0.48 |
| \$ | 0.53 | 0.47 | 0.46 | 0.28 | 0.35 | 0.47 | 0.32 | 0.30 | 0.48 | 0.00 |

Figure 4: Distance matrix of BMU chains for the two Italian verb forms *vediamo* (‘we see’) and *crediamo* (‘we believe’) on a spatio-temporal 40x40 T2HSOM.



As the chains unfold, per-node distance progressively narrows down, to converge on the same BMUs for the shared morphological ending *-iamo*, where distance drops to 0. Note furthermore that structure is inherently graded at morpheme boundaries, with an early start corresponding to the shared *-d* in the roots *ved-* and *cred-*.

The next step is to approximate the distance between two words *a* and *b* in terms of the topological alignment of their corresponding activation chains, expressed as a function of the topological distance  $m_{ij}$  of their constituent BMUs in the distance matrix, normalised by the length  $n_a$  and  $n_b$  of the two words as in equation (3):

$$(3) TD(a, b) = \frac{1}{2} \left( \frac{1}{n_a} \sum_{i=1}^{n_a} \min_{j=1, \dots, n_b} \{m_{ij}\} + \frac{1}{n_b} \sum_{j=1}^{n_b} \min_{i=1, \dots, n_a} \{m_{ij}\} \right)$$

Finally, given a specific morphological family, we can quantify its morphological consistency, by assessing how well constituent parts of family members are mutually aligned on the map.

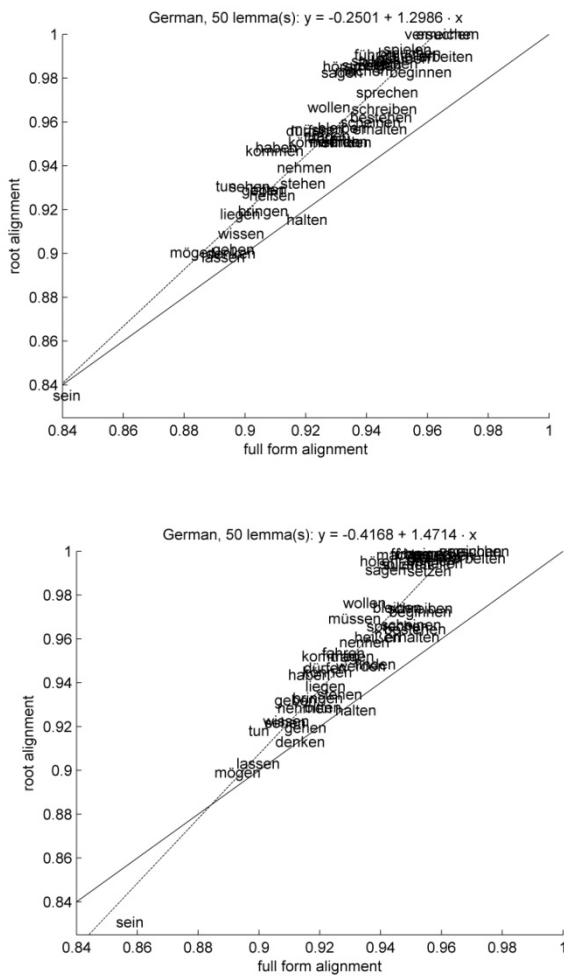


Figure 5: Correlation patterns between word-based vs. root-based alignment scores in German verb paradigms on a temporal (top) and a spatio-temporal (bottom) map.

In Figure 5, we plotted the 50 German subparadigms by assigning each subparadigm i) the average intra-paradigmatic alignment between full forms (x axis), and ii) the average intra-paradigmatic alignment between the corresponding verb roots (y axis).

The two plots are based on a temporal (top) and a spatio-temporal (bottom) German map, both showing a straightforward linear correlation between full form alignment and root alignment. More interestingly, the two maps show differential sensitivity to idiosyncratic, sub-regular and regular paradigms, with temporal maps aligning highly irregular paradigms (bottom left corner in the plots) better than regular paradigms (top right corner), when compared with spatio-temporal maps. Italian data offer a comparable pattern of results on temporal and spatio-temporal maps, providing further evidence that the temporal and spatio-temporal maps organise input word forms according to differently overlapping activation chains (see section 5, General Discussion, for marginal but significant differences between the two verb systems). Mean alignment values averaged across inflectional endings of paradigmatically-homologous verb forms (i.e. forms taking the same paradigm cell in different paradigms) are shown in Figure 6 for both German and Italian maps. The two plots show that spatio-temporal maps organise inflected data in morphologically more coherent families, with alignment values on affixes (black circles) being consistently higher than the corresponding values in temporal maps (white circles). Comparative values for each paradigm cell are reported in Table 1 with their *p-values*.

| cells | German |       |         | Italian |       |         |
|-------|--------|-------|---------|---------|-------|---------|
|       | T      | ST    | p-value | T       | ST    | p-value |
| i     | 0.867  | 0.876 | 0.006   | 0.882   | 0.891 | 0.002   |
| pA    | 0.888  | 0.898 | 0.004   | 0.897   | 0.907 | 0.020   |
| pE    | 0.881  | 0.888 | 0.020   | 0.872   | 0.878 | 0.078   |
| 1SIE  | 0.840  | 0.848 | 0.0100  | 0.851   | 0.853 | 0.753   |
| 2SIE  | 0.861  | 0.884 | 0.000   | 0.854   | 0.860 | 0.092   |
| 3SIE  | 0.838  | 0.845 | 0.023   | 0.841   | 0.839 | 0.347   |
| 1PIE  | 0.866  | 0.874 | 0.009   | 0.903   | 0.913 | 0.011   |
| 2PIE  | 0.847  | 0.853 | 0.035   | 0.887   | 0.889 | 0.438   |
| 3PIE  | 0.865  | 0.873 | 0.006   | 0.880   | 0.890 | 0.085   |
| 1SKA  | 0.836  | 0.843 | 0.025   | 0.857   | 0.863 | 0.007   |
| 2SKA  | 0.874  | 0.890 | 0.001   | 0.897   | 0.904 | 0.173   |
| 3SKA  | 0.837  | 0.845 | 0.019   | 0.842   | 0.850 | 0.171   |
| 1PKA  | 0.870  | 0.880 | 0.001   | 0.897   | 0.905 | 0.009   |
| 2PKA  | 0.865  | 0.872 | 0.061   | 0.896   | 0.908 | 0.007   |
| 3PKA  | 0.870  | 0.879 | 0.001   | 0.886   | 0.891 | 0.074   |

Table 1: Mean alignment values of temporal and spatio-temporal German and Italian maps

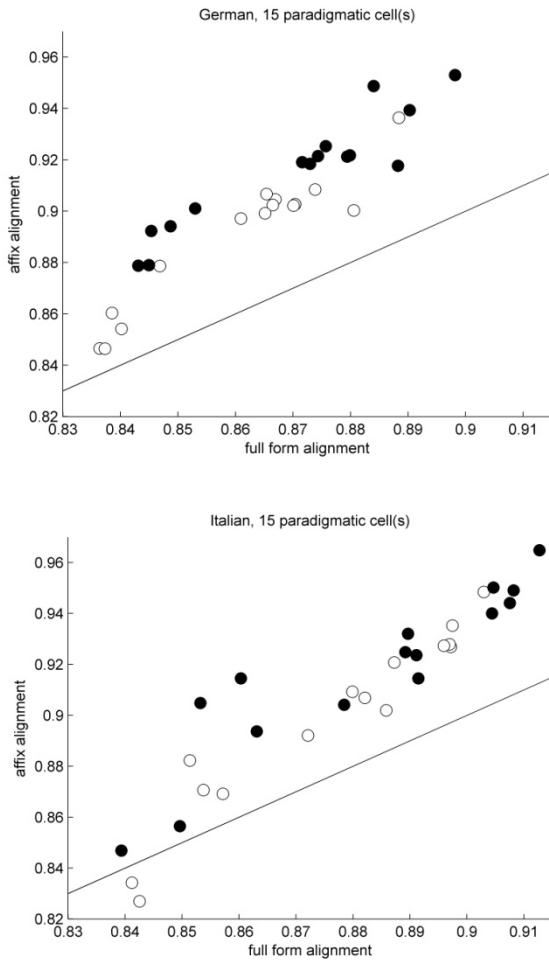


Figure 6: Paradigm-cell alignments averaged over temporal maps (white circles) and spatio-temporal maps (black circles) of German (top) and Italian (bottom).

## 5. General Discussion

There exists a strong mutual implication between principles of language processing – i.e. knowledge of “how” speakers perceive and recode strings in language – on the one hand, and linguistic data structures – or knowledge of “what” speakers know when they learn a language – on the other hand. T2HSOMs allow us to shed a sharp light on this mutual implication. As memory devices, they exhibit a remarkable capacity of recoding an incoming word form correctly, through time-bound activation of contextually appropriate BMUs. This must be a fundamental and very robust behaviour, and a preliminary condition to the development of higher-level memory structures in a long-term store.

Experiment 1 tapped the robustness of this first level of processing, showing that it is based on two local constraints: i) the actual code of the currently processed input letter, and ii) the map’s expectation for an upcoming input letter. Both factors play a role in processing, accounting for the map’s capacity to perceive input signals correctly, and to predict them on the basis of past input signals. This behaviour is the result of a trade-off between sensitivity to space (the signal code) and time (its contextual position in the string). By modulating  $\alpha$  and  $\beta$  in the map’s activation

function (1), we can control the role of the two factors and assess their relative contribution to the map’s behaviour in the face of the same input data. As spatio-temporal maps are more sensitive to the (local) input, they tend to perform consistently better in the activation task.

Results of Experiment 2 show the role of long-term (lexical) storage in recall. This is by far a more complex task than lexical activation, as it requires the capacity to reconstruct an appropriate sequence of BMUs from their integrated pattern of activation. This requires that a map develops accurate time-bound recoding of input signals, to maximise its internal predictive drive and minimise the role of the input signal in conditioning the map’s output. Ideally, words are maximally predicted when they are recoded in the long-term store through dedicated (i.e. non overlapping) activation chains. However, dedicated chains take independently trained nodes and potentially infinite memory resources. Memory self-organisation through training must thus strike a balance between dedicated recoding (full form storage) and finite memory resources.

As reported in Figure 3, temporal maps show a tendency to recall word forms better than spatio-temporal maps, since the former have a stronger bias towards dedicated memory chains. The marginally significant advantage in recalling German verbs over Italian verbs is mainly due to i) Italian verb forms being longer on average than German ones (7.1 vs. 6.4 in number of letters) and ii) Italian presenting a wider range of different word forms, due to the absence of syncretism in the selected paradigms. On the other hand, spatio-temporal maps exhibit a greater sensitivity to morphological structure, and are less prone to develop dedicated memory chains. German verbs present i) a morphological discontinuity in most (circumfixed) past participles and ii) extensive stem alternation in strong forms. These phenomena may account for the marginal advantage of Italian spatio-temporal maps over German spatio-temporal maps. Figure 7 plots the overall root alignment distribution of German and Italian paradigms, showing that Italian verbs tend to concentrate more towards the regular end of the distribution than German verbs do.

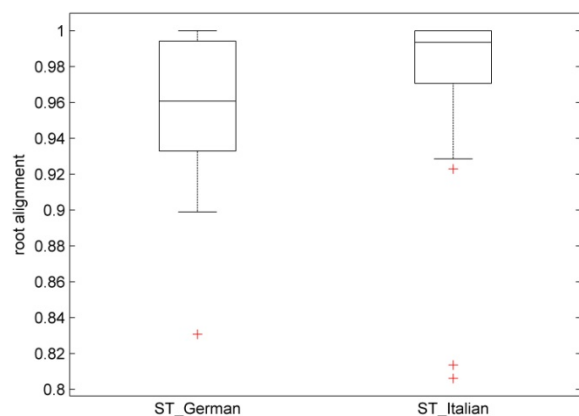


Figure 7: Root alignment distribution of German and Italian paradigms. Edges are 1<sup>st</sup> and 3<sup>rd</sup> quartiles, central mark is the median. Outliers are plotted individually.

Experiment 3 is intended to show that memory self-organisation is conducive to the emergence of morphological structure in the lexical store. This is not to suggest that memory self-organisation is everything we need to account for to explain acquisition of morphological structure, but that the capacity of aligning morphological constituents between morphologically-related word forms is a fundamental precondition to the emergence of word structure. This capacity is something that must in turn be tuned up through incremental storage of morphologically complex strings, as it appears to strongly depend on the morphology of the input language. Moreover, perception of morphological structure can vary depending on the map's recoding strategy. A map with a temporal bias will tend to enforce positional recoding, with letters recruiting different BMUs depending on their position in the string. This makes it more difficult for the map to align the roots in German *machen* and *gemacht*, as the root is shifted by two letters in the past participle form. However, if the map can make letter recoding less position-sensitive (as in a spatio-temporal setting), then alignment can be enforced between time-shifted roots. This is shown in Figure 8, where the two alignment plots of the same paradigm (*finden*) present remarkably different values on the past participle depending on the recoding bias of the two maps.

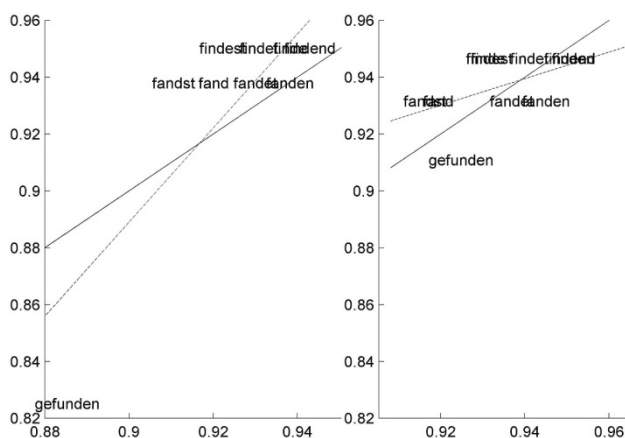


Figure 8: Alignment plot of the *finden* paradigm on a temporal (left) and a spatio-temporal (right) map.

## 6. Concluding remarks

T2HSOMs are biologically inspired computational models of the mental lexicon and useful tools to understand more of the low-level dynamics of lexical recoding and access. In this paper, we showed that they can clarify issues of lexical processing and acquisition of morphological structure, as they allow us to gain insights into their dynamic interdependency, while offering a principled solution to the controversial issue of time-bound recoding in neural networks. In particular, we addressed the problem of how it is possible to provide a rigorous, quantitative evaluation of T2HSOMs by monitoring their behaviour on a battery of memory-related functional tasks such as lexical activation, lexical recall, and word align-

ment. We showed that subtle variations in parameter setting can condition a map's behaviour considerably, and alter its topological organisation from the perspective of its emerging morphological structure.

In this context, a further research question concerns the ability of T2HSOMs to perform traditional classification tasks such as inflectional family membership assignment or possible generalisation of sequence categorisation to sequences that have not been presented in the training phase. The evaluation framework presented here can straightforwardly be extended in this direction, e.g. by putting a T2HSOM to the test of recalling novel word forms. Further experiments will allow comparison of the notion of structure-aware alignment proposed here with other less morphologically-motivated alignment metrics such as the Levenshtein distance or Nosofsky's (1990) analogy. In particular, it will be interesting to assess how well our system can discriminate between the notions of stored word forms, non-stored but morphologically coherent forms and non-stored morphologically incoherent forms, compared with other systems. The current research indicates that morphology learning should play a prominent role in this task.

## 7. References

- Aitchison, J. (1987), *Words in the mind*. London: Blackwell.
- Baayen, H., Dijkstra, T., and Schreuder, R. (1997), Singulars and plurals in Dutch: Evidence for a parallel dual route model, *Journal of Memory and Language*, 37, 94-117.
- Baayen, H., Piepenbrock, R., and Gulikers, L. (1995), *The CELEX Lexical Database (CD-ROM)*. Linguistic Data Consortium, Philadelphia.
- Baddeley, A.D. (1986), *Working memory*, New York: Oxford University Press.
- Baddeley, A.D. (2006), *Working memory: an overview*, in S. Pickering (ed.), *Working Memory and Education*, New York: Academic Press, 1-31.
- Baddeley, A.D. (2007), *Working memory, thought and action*, Oxford: Oxford University Press.
- Baddeley, A.D., and Hitch, G. (1974), *Working memory*, in G.H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory*, New York: Academic Press, 8, 47-89.
- Botvinick, M., and Plaut, D.C. (2006), Short-term memory for serial order: A recurrent neural network model, *Psychological Review*, 113, 201-233.
- Bowers, J. S., & Davis, C. J. (2009), Learning representations of wordforms with recurrent networks: Comment on Sibley, Kello, Plaut, & Elman. *Cognitive Science*, 33, 1183-1186.
- Brown, G.D.A., Preece, T., and Hulme, C. (2000), Oscillator-based memory for serial order, *Psychological Review*, 107, 127-181.
- Clahsen, H. (1999), *Lexical entries and rules of language: A multidisciplinary study of German inflection*, *Behavioral And Brain Sciences*, 22, 991-1060.

- Cowan, N. (2001), The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Behavioral and Brain Sciences*, 24, 87-185.
- Davis, C.J. (2010), The Spatial Coding Model of Visual Word Identification, *Psychological Review*, 117(3), 713-758.
- Ferro, M., Ognibene, D., Pezzulo, G., and Pirrelli, V. (2010), Reading as active sensing: a computational model of gaze planning in word recognition, *Frontiers in Neurobotics*, 4(6), 1-16.
- Ferro, M., Marzi, C., Pirrelli, V. (2011a), T2HSOM: Understanding the Lexicon by Simulating Memory Processes for Serial Order, in *Proceedings of the First International Workshop on Lexical Resources WoLeR 2011*, Ljubljana, Slovenia, 32-41.
- Ferro, M., Marzi, C., Pirrelli, V. (2011b), A self-organizing model of word storage and processing: implications for morphology learning, in *Lingue e Linguaggio*, X, 2, Bologna: Il Mulino, 209-226.
- Ford, M., Marslen-Wilson, W., and Davis, M. (2003), Morphology and frequency: contrasting methodologies, in H. Baayen and R. Schreuder (eds.), *Morphological Structure in Language Processing*, Berlin-New York: Mouton de Gruyter.
- Frost, R., Forster, K.I., and Deutsch, A. (1997), What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 829-856.
- Gupta, P., Lipinski, J., Abbs, B., Lin, P.H., (2005), Serial position effects in nonword repetition, *Journal of Memory and Language*, 53, 141-162.
- Hay, J. (2001), Lexical frequency in morphology: is everything relative?, *Linguistics*, 39, 1041-1111.
- Henson, R.N. (1998), Short-term memory for serial order: The start-end model, *Cognitive Psychology*, 36, 73-137.
- Kohonen, T. (2001), *Self-Organizing Maps*, Heidelberg: Springer.
- Koutnik, J. (2007), Inductive Modelling of Temporal Sequences by Means of Self-organization, in *Proceeding of International Workshop on Inductive Modelling (IWIM 2007)*, Prague, 269-277.
- Libben, G. (2006), Why studying compound processing? An overview on the issues. In G. Libben and G. Jarema (Eds.) *The representation and processing of compound words*, Oxford: Oxford University Press, 1-22.
- Lüdeling, A., and Jong, N. de (2002), German particle verbs and word formation, in N. Dehé, R. Jackendoff, A. McIntyre and S. Urban, (eds.), *Explorations in Verb-Particle Constructions*, Berlin: Mouton der Gruyter.
- Miller, G.A. (1956), The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychological Review*, 63 (2), 81-97.
- Montemagni, S. et al. (2003). Building the Italian syntactic-semantic treebank. In Abeillé, A. (ed.) *Building and Using Parsed Corpora*, 189-210. Dordrecht: Kluwer.
- Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., and Baayen, H. (2004), Morphological Family Size in a Morphologically Rich Language: The Case of Finnish Compared With Dutch and Hebrew, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(6), 1271-1278.
- Nosofsky, R. M. (1990), Relations between exemplar similarity and likelihood models of classification, *Journal of Mathematical Psychology*, 34, 393-418.
- Pinker, S., and Prince, A. (1988), On language and connectionism: Analysis of a parallel distributed processing model of language acquisition, *Cognition*, 29, 195-247.
- Pinker, S., and Ullman, M.T. (2002), The past and future of the past tense, *Trends in Cognitive Science*, 6, 456-463.
- Pirrelli, V., and Herreros, I. (2007) Learning Morphology by Itself, in G. Booij, L. Ducceschi, B. Fradin, E. Guevara, A. Ralli and S. Scalise (eds.) *Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*, University of Bologna, 269-290.
- Pirrelli, V., Ferro, M., and Calderone, B. (2011), Learning paradigms in time and space. Computational evidence from Romance languages, in M. Goldbach, M.O. Hinzelin, M. Maiden and J.C. Smith (eds.) *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, Oxford: Oxford University Press, 135-157.
- Post, B., Marslen-Wilson, W., Randall, B., and Tyler, L.K. (2008), The processing of English regular inflections: Phonological cues to morphological structure, *Cognition*, 109, 1-17.
- Prasada, S., and Pinker, S. (1993), Generalization of regular and irregular morphological patterns, *Language and Cognitive Processes*, 8, 1-56.
- Rastle, K., Davis and M.H. (2004), The broth in my brothers brothel: Morpho-orthographic segmentation in visual word recognition, *Psychonomic Bulletin and Review*, 11(6), 1090-1098.
- Sibley, D.E., Kello, C.T., Plaut, D., and Elman, J.L. (2008), Large-scale modeling of wordform learning and representation, *Cognitive Science*, 32, 741-754.
- Tabak, W., Schreuder, R., and Baayen, R.H. (2005), Lexical statistics and lexical processing: semantic density, information complexity, sex and irregularity in Dutch, in M. Reis and S. Kepser (eds.), *Linguistic Evidence*, Berlin: Mouton de Gruyter, 529-555.
- Taft, M. (1979), Recognition of affixed words and the word frequency effect, *Memory and Cognition*, 7, 263-272.
- Ullman, M.T. (2004), Contributions of memory circuits to language: The declarative/procedural model, *Cognition*, 92, 231-270.6