# Binding by Synchronisation: a Task-Dependence Hypothesis.

Guido Bugmann

School of Computing, University of Plymouth, Plymouth PL4 8AA, United Kingdom
Email: gbugmann@soc.plym.ac.uk, Fax (+44) 1752 23 25 40

Abstract:

It is argued that binding needs to be task dependent, and cannot usefully be driven by properties of the stimulus alone. However, task dependent binding can only take place after the patterns in a stimulus have been identified. Thus pattern recognition needs to be done prior to binding. Synchronisation may be a consequence of pattern recognition and can be used to localise the pattern and tag its attributes at different levels of information processing.

-------------------------------------------

It is frequently proposed that the function of synchronisation is to bind features belonging to a same stimulus. Let us assume that this is the case. The working hypotheses in early papers on synchronisation were that binding is driven by properties of the stimulus and is mediated by lateral connections within a cortical area (or between areas). However, these hypotheses suffer from several weaknesses:

i) The data most frequently cited as evidence that lateral connections cause synchronisation do not provide such evidence. In strabismic cats, there are no lateral connections between neurons in V1 responding to inputs from different eyes. The reduced synchronisation between these neurons is interpreted as indicating a synchronising role of the lateral connections (König et al, 1993). However, strabismic cats only analyse the visual information of one eye at a time. So, the data can also be interpreted as indicating that (lack of) synchronisation results from the (lack of) processing of the information by higher visual layers. Another set of data shows that severing the callosal link between the visual areas in the two hemispheres eliminates the synchronisation between neurons in these areas (Engel et al., 1991b). This is also taken as evidence for the role of lateral connections for synchronisation. However, callosal fibres also comprise the feedback fibres from higher areas in the contralateral hemisphere. Thus the data could also indicate that synchronisation is mediated by *feedback projections* from the area that uses the information. This is consistent with findings by (Sillito et al, 1994) and (Munk et al., 1995).

ii) The data used to indicate that synchronisation is necessary for perception do not give such evidence. For instance, in cat and humans with strabismic amblyopia, only one eye is normally used. If forced to use the other eye, subjects reveal deficits in connecting contours and discriminating details near contours (see references in Roelfsma et al., 1994a). Only weak synchronisation is observed between neurons responding to this other eye (Roelfsma et al., 1994a). This is interpreted as indicating that synchronisation is needed for correct perception, but it could as well indicate the opposite, namely that correct *perception causes synchronisation*.

iii) The hypothesis of stimulus-driven binding does not explain how neurons know what they should bind. For instance when you are observing your grand-mother's face, you may wish to determine her identity, if she has her eyes open or if her skin is smooth. In the three cases, different elements of the visual input need to be grouped. Thus, if binding by synchronisation takes place, it cannot be stimulus-driven. External inputs are needed to control binding in a *task-dependent* way.

The context field (CF) connections postulated in the paper by Phillips and Singer (1997) open the door to a task dependent binding. However, they are not used for that purpose. Instead, they help in setting up a connectivity that maximises the transfer of relevant information about the input, where the "relevance" is determined by properties of the input. Thus binding remains input-driven rather than task-driven.

Unfortunately, task-driven binding is problematic too. It raises a new question as a counterpart to the point iii) above: How can a high-level process know which parts of the input image to group before it knows what is in the image ?

One way to solve this problem is to abandon the notion that a binding mechanism is necessary for pattern recognition. The feature detectors at various stages of visual information processing may exhibit enough positional dependence to make further binding unnecessary. Further, if a visual scene contains numerous objects or interpretations of objects, why should these not all be recognised by their respective groups of neurons ? The problem is possibly more to *localise* an object or pattern of interest and that is where synchronisation may be useful. In a neural network described in (Bugmann and Taylor, 1993, 1994a, 1994b), when a neuron in a higher layer starts firing, this neuron induces the synchronisation of all neurons in the lower level layers contributing to its response, by virtue of a cascade of feedback projections. The resulting area of synchronised activity in a low layer (e.g. corresponding to V1 or LGN) indicates the location of the pattern. By enabling or disabling high level neurons in a task dependent way,

only the relevant visual features are synchronised. In such a scheme, synchronisation and binding is a consequence of the recognition of a pattern, not its cause.

Task-dependent synchronisation predicts synchronisation between distant neurons which represent an object relevant to the task at hand (Bugmann and Taylor, 1994b). This differs from the familiarity factor suggested by Phillips and Singer (1997).

Another possible role of top-down synchronisation is multi-resolution binding. For instance, in the grand-mother example given above, the texture of the skin may be represented at a relatively low level, the shape of her eyes at a higher level and her identity at an even higher level. Thus different tasks require information to be extracted from different levels. By allowing a high level neuron representing the concept "my grand-mother" to fire, the sub-set of low level texture neurons corresponding to her skin would be synchronised. This sub-set could then be used to produce the response to the question about the smoothness of the skin (fig 1). Thus, a function of synchronisation may be the tagging of attributes of objects at different levels of resolution. Sensory information processing can then be viewed as a *presentation* process in which information is organised for other systems that need it.

References:

Bugmann, G. and Taylor J.G. (1993) "A model for latencies in the visual system" in proc. 3rd Conf. on Artificial Neural Networks (ICANN'93, Amsterdam), Gielen S. and Kappen B. (eds), Springer, Berlin, p.165-168.

Bugmann, G. and Taylor J.G. (1994a) "Role of short-term memory in neural information propagation" Extended Abstract Book of Int. Symp. on Dynamics of Neural Processing, Washington, 132-136

Bugmann, G. and Taylor J.G. (1994b) "A top-down model for neuronal synchronization" Research Report NRG-94-02, School of Computing, University of Plymouth, Plymouth PL4 8AA, UK. (http://www.tech.plym.ac.uk/soc/staff/GuidBugm/Bugmann.html#pub94)

Munk M.H.J., Nowak L.G., Nelson J.I. and Bullier J. (1995) "Structural Basis of Cortical Synchronization. 2. Effects of Cortical-Lesions.", J. of Neurophysiology, 74, pp. 2401-2414.

Engel et al., 1991b (as in the target article)

König et al, 1993 (as in the target article)

Pandya D.N. and Yeterian E.H. (1985) "Architecture and Connections of Cortical Association Areas", in "Cerebral Cortex", Peters A. and Jones E.G. (eds), Plenum Press, New York and London, pp. 3-61.

Roelfsma et al., 1994a (as in the target article)

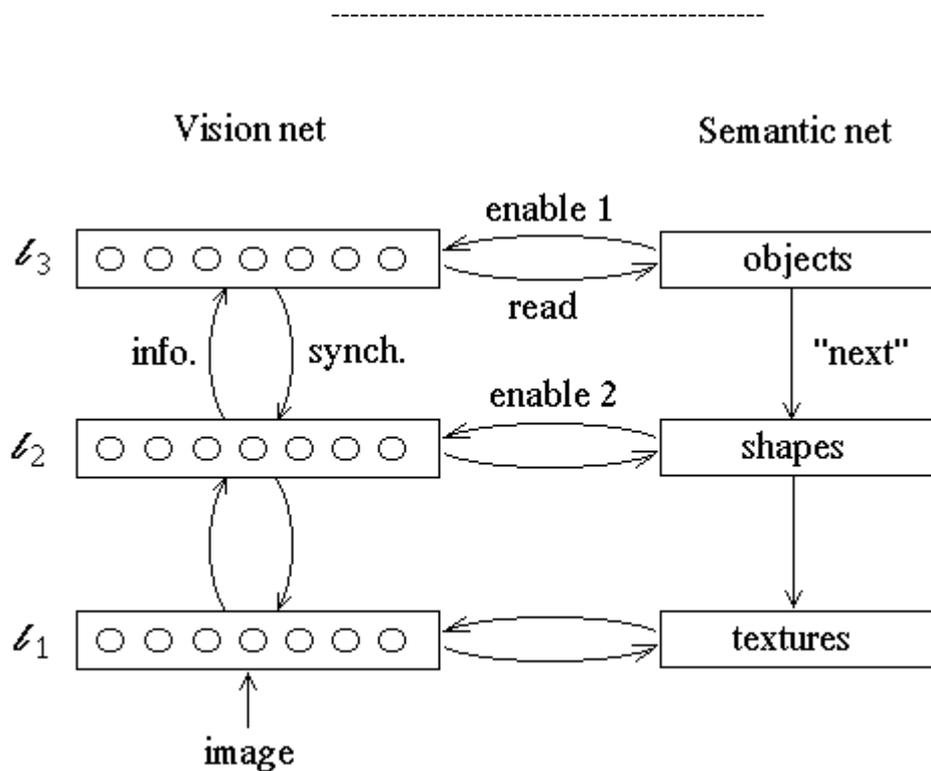Sillito et al, 1994 (as in the target article)

------------------------------------------------



**Figure 1.** Principle of multi-resolution binding. A question such as "how smooth is the skin of my grand-mother's face ?" is answered after following sequence of events: First the grand-mother cell in layer $l_3$ is enabled by the semantic "objects" layer. When the grand-mother is recognised, neurons in layer $l_2$ and $l_1$ representing attributes of the face are synchronised by feedback projections. Finally the semantic "texture" layer can use the group of synchronised neurons in layer $l_1$ to evaluate the smoothness of the skin. If the texture of the eyelids is of interest, one more processing step is needed: After recognition in layer l3 is acknowledged by the semantic "object" layer, signals from the semantic "shape" layer restricts the set of active neurons in layer $l_2$ to those coding for the eyes. This in turn limits the population of synchronised neurons in $l_1$ to those corresponding to the eyes. Anatomical data in (Pandya and Yeterian, 1985) may support such a model.