

Reservoir computing and the Sooner-is-Better bottleneck

Commentary on: Christiansen & Chater, "The now-or-never bottleneck: a fundamental constraint on language", to appear in *Behavioral and Brain Sciences*. © Cambridge University Press.

Stefan L. Frank^a and Hartmut Fitz^b

^a*Centre for Language Studies, Radboud University Nijmegen, P.O. Box 9103, 6500 HD Nijmegen, The Netherlands;* ^b*Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.*

Abstract: Prior language input is not lost but integrated with the current input. This principle is demonstrated by "reservoir computing": Untrained recurrent neural networks project input sequences onto a random point in high-dimensional state space. Earlier inputs can be retrieved from this projection, albeit less reliably so as more input is received. The bottleneck is therefore not "Now-or-Never" but "Sooner-is-Better."

Christiansen and Chater argue that the "Now-or-Never" bottleneck arises because input that is not immediately processed is forever lost when it is overwritten by new input entering the same neural substrate. However, the brain, like any recurrent network, is a state-dependent processor whose current state is a function of both the previous state and the latest input (Buonomano & Maass 2009). The incoming signal therefore does not wipe out previous input. Rather, the two are integrated into a new state that, in turn, will be integrated with the next input. In this way, an input stream "lives on" in processing memory. As prior input is implicitly present in the system's current state, it can be faithfully recovered from it, even after some time. Hence, there is no need to immediately "chunk" the latest input to protect it from interference. This does not mean that no part of the input is ever lost. As the integrated input stream grows in length, it becomes increasingly difficult to reliably make use of the earliest input. Therefore, the sooner the input can be used for further processing, the more successful this will be: There is a "Sooner-is-Better" rather than a "Now-or-Never" bottleneck.

So-called "reservoir computing" models (Lukoševičius & Jaeger 2009; Maass et al. 2002) exemplify this perspective on language processing. Reservoir computing applies untrained recurrent networks to project a temporal input stream into a random point in a very high-dimensional state space. A "read-out" network is then calibrated, either online through gradient descent or offline by linear regression, to transform this random mapping into a desired output, such as a prediction of the incoming input, a reconstruction of (part of) the previous input stream, or a semantic representation of the processed language. Crucially, the recurrent network itself is not trained so the ability to retrieve earlier input from the random projection cannot be due to learned chunking or other processes that have been acquired from language exposure. Indeed, Christiansen

and Chater (1999) found that even before training, the random, initial representations in a simple recurrent network's hidden layer allow for better-than-chance classification of earlier inputs. Reservoir computing has been applied to simulations of human language learning and comprehension, and such models accounted for experimental findings from both behavioural (Fitz 2011; Frank & Bod 2011) and neurophysiological studies (Dominey et al. 2003; Hinaut & Dominey 2013). Moreover, it has been argued that reservoir computing shares important processing characteristics with cortical networks (Rabinovich et al. 2008; Rigotti et al. 2013; Singer 2013), making this framework particularly suitable to the computational study of cognitive functions.

To demonstrate the ability of reservoir models to memorize linguistic input over time, we exposed an echo-state network (Jaeger & Haas 2004) to a word sequence consisting of the first 1,000 words (roughly the length of this commentary) of the Scholarpedia entry on echo-state networks. Ten networks were randomly generated with 1,000 units and static, recurrent, sparse connectivity (20% inhibition). The read-outs were adapted such that the network had to recall the input sequence 10 and 100 words back. The 358 different words in the corpus were represented orthogonally, and the word corresponding to the most active output unit was taken as the recalled word. For a 10-word delay, the correct word was recalled with an average accuracy of 96% (SD = 0.6%). After 100 words, accuracy remained at 96%, suggesting that the network had memorized the entire input sequence. This indicates that there was sufficient information in the system's state-space trajectory to reliably recover previous perceptual input even after very long delays. Sparseness and inhibition, two pervasive features of the neocortex and hippocampus, were critical: Without inhibition, average recall after a 10-word delay dropped to 51%, while fully connected networks correctly recalled only 9%, which equals the frequency of the most common word in the model's input. In short, the more brain-like the network, the better its capacity to memorize past input.

The modelling results should not be mistaken for a claim that people are able to perfectly remember words after 100 items of intervening input. In order to steer the language system towards an interpretation, earlier input need not be available to explicit recall and verbalization. Thus, it is also irrelevant to our echo-state network simulation whether or not such specialized read-outs exist in the human language system. The simulation merely serves to illustrate the concept of state-dependent processing where past perceptual input is implicitly represented in the current state of the network. A more realistic demonstration would take phonetic, or perhaps even auditory, features as input, rather than pre-segmented words. Since the dynamics in cortical networks is vastly more diverse than in our model, there is no principled reason why such networks should not be able to cope with richer information sources. Downstream networks can then access this information when interpreting incoming utterances, without explicitly

recalling previous words. Prior input encoded in the current state can be used for any context-sensitive operation the language system might be carrying out – for example, to predict the next phoneme or word in the unfolding utterance, to assign a thematic role to the current word, or to semantically integrate the current word with a partial interpretation that has already been constructed.

Because language is structured at different levels of granularity (ranging from phonetic features to discourse relations), the language system requires neuronal and synaptic mechanisms that operate at different timescales (from milliseconds to minutes) in order to retain relevant information in the system's state. Precisely how these memory mechanisms are implemented in biological networks of spiking neurons is currently not well-understood; proposals include a role for diverse, fast-changing neuronal dynamics (Gerstner et al. 2014) coupled with short-term synaptic plasticity (Mongillo et al. 2008) and more long-term adaptation through spike-timing dependent plasticity (Bi & Poo 2001). The nature of processing memory will be crucial in any neurobiologically viable theory of language processing (Petersson & Hagoort 2012), and we should therefore not lock ourselves into architectural commitments based on stipulated bottlenecks.

Acknowledgements

H.F. would like to thank Karl-Magnus Petersson for helpful discussions on these issues. S.L.F. is funded by the European Union Seventh Framework Programme under grant no. 334028.

References

- Bi, G. & Poo, M. (2001). Synaptic modification of correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience* 24 :139–66.
- Buonomano, D. V. & Maass, W. (2009) State-dependent computations: Spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience* 10:113–25.
- Christiansen, M. H. & Chater, N. (1999) Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23:157–205.
- Dominey, P. F., Hoen, M., Blanc, J. -M. & Lelekov-Boissard, T. (2003) Neurological basis of language and sequential cognition: Evidence from simulation, aphasia and ERP studies. *Brain and Language* 86:207–25.
- Fitz, H. (2011) A liquid-state model of variability effects in learning nonadjacent dependencies. In: *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, ed. L. Carlson, C. Hölscher & T. Shipley, pp. 897–902. Cognitive Science Society.
- Frank, S. L. & Bod, R. (2011) Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science* 22:829–34.
- Gerstner, W., Kistler, W.M., Naud, R. & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press.
- Hinaut, X. & Dominey, P. F. (2013) Real-time parallel processing of grammatical structure in the fronto-striatal system: A recurrent network simulation study using reservoir computing. *PLoS ONE* 8(2): e52946.
- Lukoševičius, M. & Jaeger, H. (2009) Reservoir computing approaches to recurrent neural network training. *Computer Science Review* 3:127–49.
- Jaeger, H. & Haas, H. (2004) Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science* 304:78–80.
- Maass, W., Natschläger, T. & Markram, H. (2002) Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation* 14:2531–60.
- Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, 319, :1543–46.
- Petersson, K. M. & Hagoort, P. (2012) The neurobiology of syntax: beyond string sets. *Philosophical Transactions of the Royal Society B* 367:1971–83.
- Rabinovich, M., Huerta, R. & Laurent, G. (2008) Transient dynamics for neural processing. *Science* 321:48–50.
- Rigotti, M., Barak, O., Warden, M. R., Wang, X. -J., Daw, N. D., Miller, E. K. & Fusi, S. (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–90.
- Singer, W. (2013) Cortical dynamics revisited. *Trends in Cognitive Sciences* 17:616–26.