Storage vs. Processing in Models of Word Inflection.
A Neuro-computational Hebbian Perspective

Vito Pirrelli
Institute for Computational Linguistics, National Research Council, Pisa
vito.pirrelli@ilc.cnr.it

Introduction

The advent of connectionism in the 80's popularised the idea that the human lexical processor consists of a network of parallel processing units selectively firing in response to sensory stimuli (McClelland & Elman 1986, Norris 1994, Luce & Pisoni 1998). Sensory information initiates concurrent activation of specialised processing nodes, which respond to features/units of the input as they unfold with time. Co-activation makes nodes compete with one another for primacy, until an optimal candidate is singled out as the winner. In the light of these assumptions, the most important contribution of connectionism to the theoretical debate on lexical modelling at the time was the utter rejection of the idea that word recognition and production require a dichotomous choice between storage and processing. However, in spite of the prima facie psycho-computational allure of their approach, early connectionist models also embraced a number of unsatisfactory assumptions about word learning and processing: from wired-in conjunctive coding of input symbols in context, to unrealistic output supervision required by the gradient descent algorithm, to a model of word production as a derivational function mapping one lexical base onto fully inflected forms.

More recently, a growing number of approaches to inflection in both Psycholinguistics and Theoretical Linguistics developed the view that surface word relations represent a fundamental domain of morphological competence (Matthews 1991; Bybee 1995; Pirrelli 2000; Burzio 2004; Booij 2010; Ackerman & Malouf 2013; Blevins 2016). Learning the morphology of a language amounts to acquiring relations between fully stored lexical forms, which are concurrently available in the speaker's mental lexicon and jointly facilitate processing of morphologically related forms through patterns of emergent self-organisation. This novel view presupposes an integrative language architecture, where storage and processing, far from being conceived of as insulated and non interacting modules, are the short-term and the long-term dynamics of the same underlying process of adaptive specialisation of synaptic connections. This view, upheld by recent evidence of the neuro-anatomical bases of short-term and long-term memory processes, crucially hinges on Hebbian principles of synaptic plasticity, which are, in turn, in keeping with mathematical models of discriminative learning (Rescorla & Wagner 1972, Ramscar & Yarlett 2007, Baayen et al. 2011). I contend that integrative computer models of Hebbian language learning represent an exciting way forward in current neuro-computational research on word processing, and a persistently fertile legacy of the connectionist revolution.

Recurrent ANNs and Hebbian learning

Over the past 30 years, connectionist architectures have been trying to address some of the above mentioned technical and theoretical limitations of early architectures, while offering ways to conceptualise child word learning and processing that are more in lines with psycholinguistic and linguistic acquisitions in human morphological competence.

In simple recurrent networks (Elman 1990, Jordan 1986), the input to the network at time *t* is represented by the current level of activation of nodes in the input layer augmented with the level of activation of nodes in the hidden layer at the previous time tick (*t-1*). The network keeps track of its past activation states and develops a serial memory of previous inputs. A simple but very powerful task that recurrent networks can carry out most naturally is predicting the upcoming symbol at time *t+1*, given previous exposure to a sequence of symbols up to time *t*. Prediction is very ecological and is known to be heavily involved in human language processing (Altmann & Kamide 2007; DeLong, Urbach, & Kutas 2005; Pickering & Garrod 2007). In addition, it provides a natural way to get instantaneous, observable feedback: if the current input stimulus is different from what is expected and predicted by the network, the network parameters are adjusted for the currently presented input to be the network's future safest guess given the past evidence.

Along similar lines, Temporal Self-organizing maps (TSOMs, Ferro et al. 2011, Marzi et al. 2014, Pirrelli et al. 2015) were recently proposed as models of the dynamic topological organisation of memory nodes selectively firing when symbols are input to the map in specific temporal contexts. A temporal context is loosely defined as a temporal slot (position) in a time series of input symbols, or a window of surrounding symbols. Context-sensitive node specialisation is not wired in the map's connections at the outset (as in traditional connectionist models), but it is something that emerges as a function of input exposure in the course of training. Following equations of Hebbian synchronisation between successively firing neurons (Koutnik 2007), high-frequency input sequences tend to develop deeply entrenched inter-node synaptic connections and highly specialised nodes, functionally corresponding to human expectations for possible continuations (Marzi & Pirrelli 2016). Low-frequency input sequences, on the other hand, tend to fire blended node chains, i.e. sequences of nodes that respond to a class of partially overlapping sequences. This is what distinguishes holistic, dedicated memorisation of full forms from chunk-based storage of low-frequency forms, sharing memory chunks with other overlapping forms.

I will illustrate a few concrete examples of how neuro-computational models of word learning & processing can successfully define an effective research framework whereby Hebbian principles of synaptic plasticity are shown to a) offer a unitary account of short-term and long-term effects for serial memories (D'Esposito 2007; Ma et al. 2014), b) an explanatory account of well-known frequency effects on the processing and acquisition of word families (e.g. Milin et al. 2009a, 2009b), and c) a neurobiologically grounded discriminative view of morphological competence in the mental lexicon.

From this perspective, neurocomputational models have the potential to bridge the persisting gap between Marr's (1982) algorithmic ("how does a cognitive process work?") and implementational (how is the algorithmic level neuroanatomically implemented in the brain?) level. Although Marr originally introduced his hierarchy to emphasize that explanations at different levels are largely independent of each other, we agree with Poggio (2012) that it is now time to clarify the potential for between-level interaction, and investigate the methodological conditions for appropriate integration. In my view, one of these conditions crucially hinges on neurobiologically-inspired principles of Hebbian plasticity and their far-reaching implications on discriminative language learning and the human language architecture in general.

References

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language* 89 (3): 429-464

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4): 502–518.

Baayen, R.H., Milin, P., Durdevic, D.F., Hendrix, P. & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118: 438-481.

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Booij, Geert. 2010. Construction morphology. *Language and Linguistics Compass*, 4(7):543–555.

Burzio, Luigi, 2004. Paradigmatic and syntagmatic relations in Italian verbal inflection, volume 258, 17–44. John Benjamins, Amsterdam-Philadelphia.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5):425–455.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B. Biological Sciences* 362: 761-772.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8): 1117–1121.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2): 179–211.

Elman, J.L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive science*, 33(4), 547-582.

Ferro M., Marzi, C., & Pirrelli, V. (2011). A Self-Organizing Model of Word Storage and Processing: Implications for Morphology Learning. *Lingue e Linguaggio*, X(2): 209-226.

Jordan, M. I. (1986). Serial order: A parallel distributed processing approach. San Diego, CA: University of California.

Koutnik J. (2007). Inductive Modelling of Temporal Sequences by Means of Self-organization. *Proceeding of International Workshop on Inductive Modelling* (IWIM 2007, Prague). 269-277.

Luce, P.A. & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19: 1-36.

Ma, W.J., Husain, M., & Bays, P.M. (2014). Changing concepts of working memory. *Nature neuroscience* 17(3): 347-356.

Marr, D. (1982). Vision. San Francisco: W.H. Freeman.

Marzi, C. & Pirrelli, V. (2015). A neuro-Computational Approach to Understanding the Mental Lexicon. *Journal of Cognitive Science*, 16(4): 491-533.

Marzi, C., Ferro, M., & Pirrelli, V. (2014). Morphological structure through lexical parsability. *Lingue e Linguaggio*, XIII(2): 263-290.

Matthews, P.H. (1991). *Morphology*. Cambridge University Press, Cambridge.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18: 1-86.

Milin, P., Đurđević, D.F., & del Prado Martín, F.M. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. Journal of Memory and Language, 60(1), 50-64.

Milin, P., Kuperman, V., Kostić, A., & Baayen, R. H. (2009b). Words and paradigms bit by bit: An information theoretic approach to the processing of paradigmatic structure in inflection and derivation. In J.P. Blevins & Blevins, J. (Eds.) Analogy in grammar: Form and acquisition, 214-252. Oxford University Press.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52: 189-234.

Pickering, M.J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3): 105–110.

Pirrelli, V. (2000). *Paradigmi in morfologia. Un approccio interdisciplinare alla flessione verbale dell'italiano*. Istituti Editoriali e Poligrafici Internazionali, Pisa.

Pirrelli, V., Ferro, M., & Marzi, C. (2015). Computational complexity of abstractive morphology. In Baerman, M., Brown, D. & Corbett, G. (Eds.), *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press. 141-166.

Poggio, T. (2012). The levels of understanding framework, revised. *Perception*, 41(9): 1017-1023.

Ramscar, M., & Yarlett, D, (2007) Linguistic Self-Correction in the Absence of Feedback: A New Approach to the Logical Problem of Language Acquisition. *Cognitive Science*, 31: 927-960.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64-99.