Neural Networking, Connectionism and Parallel Distributed Processing

(cf. thesis: ch. 4.2)

Serial and parallel processing

Before looking at connectionists models *per se*, there is another issue which has concerned psychologists, that of serial or parallel processing. To what extent should the processing of language be characterised as a serial process (first identify sounds, combine them into words and then into sentences) or to what extent should it be seen as number of different processes acting at the same time and at different levels?

The information processing model (cf. Randall, 2007, pp. 14-15), although not stating it explicitly, suggest a serial processing approach; information is taken in via the senses and then various features extracted through a series of memory stores. The symbolicist approach (following from linguistic descriptions which are hierarchical in nature) also suggests serial processing of language input. From this perspective, syntactical processing precedes semantic processing. The brain first of all decodes the input from a rule-governed syntactical viewpoint which then accesses a semantic representation (for language comprehension). The brain uses a similar reverse path for language production; the semantics generate the syntax which then produces output (O'Halloran, 2003; Randall, 2007, pp. 101-124).

However, serial processing models have been challenged by parallel processing models. Based on what we know about the structure and function of the brain, a simple serial model would seem to be inadequate *(and unlikely (JH))* (although there will be times that a serial approach will provide good explanations of how language can be interpreted). The brain contains a vast number of neurons connected into neural networks which carry out myriad simultaneous and complex operations. This neural architecture has led to a general theory

of language processing and storage known as "connectionism" (also known as "parallel distributed processing", "interactive activation" or "spreading activation"). These theories postulate that the brain is able to carry out multiple levels of activity simultaneously and thus several processes can take place at the same time and not in serial order, spreading activation through many parts of the brain through a highly complex system of neural networks.

(…) Parallel distributed processing and connectionism rest on the strength of the connections between different language features (such as words) and present a very different picture of language processing than is assumed by symbolist approaches (cf. Hulstijn, 2002).

Connectionist models of language

One of the earliest connectionist models was McClelland and Rumelhart's "interactive activation model" (McClelland & Rumelhart, 1981), designed to explain how individual letter features can be seen to produce word recognition (cf. Randall, 2007, pp. 53-86). Connectionist approaches set out to explain how language can be decoded through the operation of simple processes whereby large numbers of neurons co-operate to process information. They involve the simultaneous activation of internal "nodes" within the brain. These nodes also interconnect to either inhibit or suppress the activation of all the other nodes involved until a threshold has been reached.

Nodes are assumed to be numeral processors and the information passed between them is numerical rather than symbolic. The output from a node is assumed to be the numerical sum of its inputs. The strength of connections between the nodes can be given a numerical value representing the probability that one node will co-exist with another. Through this process the interconnectivity in the brain reflects the probabilistic relations between features in the

language. The input from the language data will "train" the network. Thus the brain "learns" a language from the input. Language rules "emerge" from the input as a series of probabilities of the co-occurrence of certain features not as symbolic representation such as grammar rules.

Thus the simple sentence "It is running" would be seen as a correct piece of language and *"They is run" as incorrect due to the fact that

1. the verb form "is" is highly frequent after "it" and not after "they"
2. a verb in the "-ing" form frequently follows "is", but not a verb in its base form.

The input "it" would activate a node for "is" which would then activate nodes for "ing". A symbolic representation, however, would describe the situation in terms of grammatical features, such as verb-subject agreement, the auxiliary verb "to be" and a present participle, the present continuous tense form. The connectionist model relies on basic associative learning principles as did behaviourist principles, but with the associative learning connected into associative networks.

Evidence for the validity of such systems involves the degree to which artificial systems can be set up to mirror actual human learning. For typical examples of the process see Rumelhart and McClelland (1986) who produced a computer model to replicate the learning of the irregular past tense in English that we have described above or Van Heuven's evaluation of how second language scripts are processed (2005).

Although the evidence from such investigations is highly computational, reflecting, as it does, the close connection of PDP with artificial intelligence, the theory is a powerful metaphor for understanding possible mechanisms for language learning and language processing. *The keyword here is "possible". So far we have learned how the computer can process language, provided the software leads to this*

process. It says nothing about the way the human brain works (JH). It has instigated a number of interesting lines of research. For example, in the study cited above, MacWhiney (2001) demonstrated that the salience of different language features across languages is reflected in different patterns of noticing by the native speakers of those languages. MacWhiney suggests that such evidence demonstrates the psychological validity in language processing of the frequency of surface features within a language.

Connectionism offers an explanation of the micro-processes through which language structure (grammar and semantics) is implemented based on a highly plausible model of neural activity. Connectionism is design as a complete *(? (JH))* explanation of language processing, but only a few areas of language activity have, as yet, been investigated, such as letter recognition and past tense acquisition (cited above). However, it can be seen as a promising alternative explanation of language processing, combining language performance with neurological mechanisms. It is also possible to suggest that language processing "can be described at two levels: at the psychological level, in terms of symbol processing; and at the implementation level, in neuro-scientific terms (to which connectionism approximates)" (Chater & Christiansen, 1999, p. 236). At times the symbolic level may appear to be most applicable (for example in designing a language programme) at others the connectionist model may be more applicable (for example in designing the types of activity).

(Randall, 2007, pp. 18-21)

The contributions the connectionist model can make to the understanding of language processes in the brain are limited and theoretical. Connectionism is a theory and all theories are prone to refutation. In my opinion, the main merit of connectionism lies in its contributions to research into artificial intelligence and maybe the producing of (commercial) language programs such as translation. Nothing has occurred so far in the neuro-sciences which has confirmed without doubt that the brain works in the way of the computer programmes of the

connectionists. In the past, memory was seen as treasure room, warehouse, dovecote, mineshaft, living magnet, photographic film, hologram etc. Now it is the turn of the computer. Maybe it has to do with the language of the researchers and writers. They mostly write (myself included) of the storage of information. This is too narrow. The brain stores meaning.

It has long been assumed that the neuronal processes in the brain work in linear way. This is apparently not the case.

The up to date wisdom is that cognitive functions are being carried out by large cell assemblies which not necessarily have to consist of tight neighbouring neurons, but can act in dispersed networks. Individual neurons can also act in different networks at different times….

During the last 30 years the neurosciences have been increasingly concerned with the phenomenon of the electric nerve activity of the brain: the correlation of increased higher cognitive functions (such as language) with the increase or decrease of synchronous activity of neurons in relevant frequency bands (frequency dependant oscillation).

(Müller, 2013, pp. 165-166)

Oscillation

  In the ERP analysis of the EEG/MEG signal the neural activities are exclusively examined by looking at chronological processes of combined amplitudes. From experiments with animals it is known that cognitive processes are not based on some ten thousand neighbouring neurons which show a combined activity at a fixed location of the cortex. Cognitive processes are based on short-term combinations of many thousands of widely dispersed neurons which form a temporary functional network. Such widely connected activities are

frequently synchronised in different frequencies and appear therefore as so-called oscillatory activities of neurons.

(Müller, 2013, p. 135)

Also significant is the speed of the information transfer between different neuron networks is of importance for the functioning of the brain. Schack et al (2003) demonstrated that the speed of diffusion of oscillatory processes is different for the processing of concrete or abstract nouns. Concrete nouns showed especially between right-hemispheric measuring points a slower diffusion rate of 10 meters per second. Abstract nouns, on the other hand, showed a diffusion rate of 14 meters per second.

(Weiss, 2009) explains this with the higher number of mental motor and sensory simulations which go hand in hand with the processing of concrete nouns……Because of the many multimodal simulations for concrete nouns, the activation needs initially more time, but leads in the end to a more robust, more efficient and faster aces to the respective lexicon entry (Rickheit, Weiss, & Eikmeiyer, 2010). Pulvermüller (1999) developed a model in which at first differently acting neurons build increasingly stronger links by constant co-activation (e.g. the processing of a word). Here are also linked mental concepts with motor and sensory processes to constitute the meaning of words (Pulvermüller, 2005; Pulvermüller & Fadiga, 2010).

The above is an extract from Müller (Müller, 2013, p. 171), including the citations.

Since our brain and memory became the subject of enquiry, it was known that we can remember concrete words better than abstract ones. Early practitioners of the art of memory e. g. (Auctor ad Herennium, 1st Century BC) used this knowledge to learn better abstract words by substituting them for concrete words (peace – dove). It is only now that the neurosciences are showing us why this is the case. (see also: my thesis p.64)

Auctor ad Herennium (Ed.). (1st Century BC). *Retorica ad Herennium* (Vol. 15). Hildesheim: Georg Olms Verlag.

Chater, N., & Christiansen, M: H. (1999). Connectionism and natural language processing. In S. Garrod & M. Pickery (Eds.), *Language Processing* (pp. 233-279). Hove: Psychology Press.

Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*(18(3)), 193-223.

MacWhiney, B. (2001). The competition model: The input, the context and the brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: CUP.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1. An account of the basic findings. *Psychological Review*(88), 375-407.

Müller, H. M. (2013). *Psycholinguistk-Neurolinguistik*. München: Wilhelm Fink.

O'Halloran, K. (2003). *Critical Discours Analysis and Language Cognition*. Edingburgh: Edingburgh University Press.

Pulvermüller, F. (1999). Words in the brain's language. *Behavior and Brain Sciences, 22*, 253-336.

Pulvermüller, F. (2005). Brain mechanisms linking language and action *Nature Reviews Neuroscience, 6*, 576-582.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor cicuits as a cortcal basis for language. . *Nature Review Neuroscience, 11*, 351-360.

Randall, M. (2007). *Memory, Psychology and Second Language Learning* (Vol. 19). Amsterdam: John Benjamins Publishing Company.

Rickheit, G., Weiss, D. J., & Eikmeiyer, H.- J. (2010). *Kognitive Linguistik: Theorien, Modelle, Methoden* Tübingen: UTB.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. . In J. C. McClelland & D: E. Rumelhart (Eds.), *Parallel Distributed Processing*. Cambridge MA: The MIT Press.

Schack, B., Weiss, S., & Rappelsberger, P. (2003). Cerebral information transfer during word processing: Where and when does it occur and how fast is it? . *Human Brain Mapping, 19*, 18-36.

Van Heuven, W. J. B. (2005). Bilingual interactive activation models of of word recognition in a second language. In V. Cook & B. Bassetti (Eds.), *Second Language Writing Systems*. Clevedon: Multilingual Matters.

Weiss, D. J. (2009). *Gehirnoscillationen und neural Kommunikation während der Verarbeitung von Sprache. .* (postdoctoral lecture qualification), University of Bielefeld Bielefeld.