The Implications of Connectionism and Neural Networks for Simultaneous Interpreting

Baaziz Termina

Mohamed-V-University-Souissi, Rabat, Morocco

The aim of this paper is to approach simultaneous interpreting adopting some connectionist models and conceptions and, thus, nurture interpreting modeling. Using neural networks stimulations, connectionism looks at the brain as a network of inter-connected nodes/units that substrate any intelligent ability or cognitive process through the propagation of activation. Neural networks, in connectionist models, acquire and learn diverse cognitive patterns through the strengthening and weakening of nodes associations. Drawing on this seminal framework, this paper is an attempt to account for interpreting and novice interpreters as neural networks. Throughout the matrix of argumentation, we suggest that the accumulation of a number of input units in novice interpreters' neural nodes, as micro agents for processing informational units, and the building of the necessary knowledge associations among such nodes, through the strengthening and weakening principles, are prerequisites for being an interpreter. The establishment of a huge network of knowledge nodes enables the network to interpret in a given velocity. More to the point, such endeavor entails an extensive training on the propagation of a distributed neural activation in distributed circuits. Such neural circuits give emerge to the required neural associations echoing the declarative and procedural abilities crucial for interpreting as a quasi-automatic activity.

Keywords: connectionism, networks, inter-connection, association, activation, interpreting, learning and training

Introduction

Connectionism is a new paradigm emerged in 1980's to replace the very view that the mind processes information as symbols and combinatorial rules manipulated by computing operations. Using the concept "cognitive architecture" of neural connections, it proposes the implementation of neural networking models to account for diverse aspects of cognitive activities including language processing. In interpreting, the interpreter engages in a number of overlapped processes, and such processes lie in the heart of connectionism. These include the reception of the input, comprehension, and the activation of representation patterns necessary to produce a grammatically, syntactically, and semantically congruent output. Thus, linguistic behavior and the related cognitive processes constitute an overarching central point of both interpreting and connectionism, something which substantiates the adoption of what we view to be the implications and applications of connectionist networks for the domain of interpreting and translation. For accounting for the above-stated hypothesis, the paper is structurally designed as follows: The first short section draws a line on an overview about connectionist modeling of cognition. The second section is

Baaziz Termina, Ph.D. student, Institute of Studies and Research for Arabization, Faculty of Educational Sciences, Mohamed-V-University-Souissi.

intended to approach interpreting from connectionist lens; whereas, the third section discerns connectionist insights relevant for training future interpreters.

Connectionist Modeling of Cognition

Researchers in the paradigm of connectionism, also known the PDP (parallel distributed processing) (Rumelhart, Hinton, & McClelland, 1986) use the architecture of neural networks to model the neural interconnectedness and the way its networking results in emergent cognitive aspects. They analogized the human mind to a huge network of interconnected units. Hinton (1989) stated that "Each unit has a 'state' or 'activity level' that is determined by the input received from other units in the network" (p. 186). PDP revolves around the centerpiece idea in neurology that the brain consists of billions of highly interconnected neurons (as synonymous to agents, units, and nodes) engaging in neural processes to generate cognitive products including language comprehension and production. An individual neuron is a micro- processor that calculates the sum of its received inputs. Exceeding a certain threshold, it generates an output, which is an input to other neurons (i.e., units) to which it is connected through synapses as neuro-connectors (Cilliers, 1988).

Each synapse excites or inhibits the target neuron through the signals that come from a sensory organ, for example, and each synapse determines the strength of influence. Information is processed in parallel neural operations, and it is distributed to other parts of the brain (Cilliers, 1988, p. 16) in different cortical regions (Anderson, 2009). The synaptic connection between any two interconnected nodes has a value which can be either positive (excitatory) or negative (inhibitory) depending on the weight of their connections (Bechtel & Abrahamsen, 1991) during activation. Excitatory connections increase the activity of neural units; whereas, the inhibitory ones decrease it (Bechtel & Abrahamsen, 1991).

It should be noted that inspirational insights of connectionism came from Minsky (1986) who deemed the cognitive system as a society composed of micro-agents capable of locally providing solutions to diverse problems. However, connectionists persistently refer to what is widely known as Hebb's rule, which states that when two *neurons fire together wire together*. The neurons that get activated through neural signals in a simultaneous manner strengthen their synaptic connections.

Connectionist modeling of cognition accounts for two types of representations: distributed and localist. Distributed representations are assumed to be vectors in a representational state space. Parallel distributed representations are processed in simultaneity by a number of associated nodes in a connectionist network of nodes (McClelland, Rumelhart, & The PDP Research Group, 1986). Whereas, in localist models, localist representations are envisioned to be encoded in individual neural nodes for representing the entire concept such as cat. However, this paper is intended to work out the common ground between interpreting and connectionism in general. It is beyond the paper's scope of inquiry to provide details on the two above-mentioned types of representations on which there is an ongoing controversial debate.

Approaching Interpreting From Connectionist Perspectives

The Necessary Skills for Interpreting

As a process of neural transferring of lexico-semantic units between two spoken languages, simultaneous interpreting occurs between two neural talking networks in the real time. Interpreters must provide a topnotch output to the target audience without lagging behind the speaker. Thus, simultaneous interpreting is a complex task that is very demanding in terms of language skills and memory capacity. Accordingly, the ultimate

objective of interpreting courses is to develop in trainees a bunch of the required perceptive and cognitive processing skills such as problem solving and analytic abilities crucial for being an interpreter. This requires constructing a huge neural network of associations reflecting a high level in conceptual, procedural, and declarative knowledge or, let us put it another way, a high level in neuro-linguistic plasticity to shift between two languages in a lexico-semantically symmetrical manner. By the end of training, novice interpreters are required to have been armed with diverse inter-complementary skills in order to be able to perform the three macro-processes involved in discourse mediation (Padilla, 1999): (1) *analysis and understanding* the source text or discourse; (2) *translation or reformulation* of the source text or discourse; and (3) *production* of the text or discourse in the target language.

Among the recognized skills involved in simultaneous interpreting, which worth mentioning, are: (1) *selective listening* that requires concentration and attention to cope with velocity and mitigating perturbing elements; (2) *understanding* that necessarily requires predicting skills and speed of representation; (3) *memorizing the information in SL (Source Language)*; (4) *mental translation* which includes editing; and (5) *verbalization* in the same rate of speech speed (see Kornakov, 2000, pp. 246-248).

The Interpreter as a Networking Unit

Language processing, as well as other cognitive forms, demands from agents to internalize an enormous number of units or exemplars (i.e., items, instances) for establishing a large network of knowledge. It is prerequisite to train the network on establishing different associative strengths between those units/nodes in tune with "the frequency and regularity effects of the input", adjusting the activation values among the inter-connected units on the basis of stimuli frequency (Hulstijn, 2003, p. 418). Knowledge network in connectionism is taken to be neural, i.e., a result of neural connections inside the neural society existing in our mental system. On this basis, we suggest that an interpreter must develop his/her system's networking through knowledge units the system receives from external sources in terms of internal prior conceptual, declarative, and procedural connections of knowledge network among his/her neurons. Without such connections maintained by neural agents communicating in the mental society, probably, the interpreter's neural system cannot manage to segment the distribution and juxtaposition of sound units in chunks received from a talking unit.

Adopting the notions of recurrent networks, the author of this paper suggests that an interpreter is hypothesized to be a unit among a network of interconnected units (i.e., the interpreter, speakers, listeners, or audience). Such unit is supposed to be well-trained phonologically, lexico-grammatically, morpho-syntactically, semantically, and pragmatically to process the message input (the talking unit's output) and map it into a symmetrical output to the other nodes supposed to logically and emotionally be plugged into the net. The line of communication makes the activated neurons of all the plugged units interconnect and interact on the basis of their previous beliefs, knowledge network and the organizational map of the world matrix stored in neurons.

Human Mirror Neurons System Hypothesis supports such conception in that specific neurons fire in inferior frontal cortex and superior partial lobe when someone performs an action. The same neurons fire again when humans see someone performing that action (Iacoboni et al., 1999). Suffice to say, one's mirror neuron mirror system is connected to that of the sender to reflect his/her pattern of activity. This theorem supports the very idea that, in a communication encounter, the motor concept related to words is characterized by the transmission from one brain network to another through neuron mirrors system (Schippers et al., 2010). It is mentioning here that mirror neurons are also hypothesized to be responsible for self and others awareness

(Oberman & Ramachandran, 2009) and understanding their goals and intentions (Leonardo et al., 2005), etc. Since interpreting involves such awareness and understanding, mirror neurons theorems, we conjecture, might have brought about workable results for buttressing theorizing in the field of simultaneous interpreting where mirror neurons interact and stimulate each other.

In a similar vein to that of mirror neurons hypothesis, multi-agent connectionist model of inter-personal communications suggests that many recurrent networks communicate in straight lines with each other. The individual recurrent network of individual inter-connected units "stimulates the process of information uptake, integration and memorization within individual agents". Each agent's recurrent network contains an auto-associative recurrent network. Communication is controlled by trust weights (Overwalle & Heylighen, 2006, pp. 606, 608). Trust in multi-agent talking nets is "determined by the consistency with the receiving agents existing beliefs, and results" (Overwalle & Heylighen, 2006, pp. 606, 608). The concept "trust weight", in this context, is synonymous with changes occurring during communication process in terms of *connection between individual networks*. These weights trigger selective propagation and thus lead to the filtering out of irrelevant information. In such context, communicators implement, as Overwalle and Heylighen noted, Grice's (1975) maxims of quality and quantity.

Recall that inter-connected agents are pragmatically conscious and aware of the requirements of communication set forth by Grice (i.e., *quality*, *quantity*, *relevance*, and *manner*). Connectionist vehicle theory of consciousness makes clear that representation in the brain networks is a conscious experience of the very states (such as the activation of representations patterns that constitute the brain concurrent response to prevailing conditions); the communicating brains are intentional about something (Opie, 1998, p. 128). For an interpreter to work out the coherence of the received message, he/she shall be cooperative and participative within the network so as to reach the intentionality and aboutness activated in the mental unit of the speaker as a talking net. Such cooperation is tantamount to neural trust weight. However, as the next section will show, such level demands from novice interpreters/students a lot of training on the network of knowledge in two languages and exercise in a given frequency on the basis of *trials and errors*. As such, they come to familiarize their neural cells with the morpho-syntactic and conceptual units used in conferences and other interpreting situations and settings.

Interpreters Training and Connectionism

Interpreter's Neural Networking and Environment

Interpreting students necessarily come to translation classes with linguistic lexico-grammatical elements/nodes mirroring the neural interactions and associations in their neural dynamic systems. Presumptively, they have a given command in the structure (i.e., lexical system impending semantic system) and meaning (semantic system of lexical units) of two languages or more. Their neural associations represent their bilingual fluency and their conceptual network, which they would refine in the new environment for becoming future interpreters. Cilliers (1988) drew a line on such adaptability of the mental system of networking nodes stating: "A complex dynamic system has to develop its internal structure and has the ability to adapt that structure in order to cope with the changes taking place in the environment" (p. 12), which is the classroom in our case. He further explained such dialectical relationship between the environment matrix and the brain networks permeating such matrix in that the brain as a complex system can develop a complex structure. This latter develops its structure from fairly unstructured beginnings through changing the

distribution of elements composing the system. This occurs under influence of both the external environment and the history of the system, because the system is plastic (Cilliers, 1988, p. 16).

The dialectical relationship between the internal structure of the student's networking system and the environment within which he/she develops that structure might best be understood in the light of the didactic triangle of the three axes: *student*, *knowledge*, and *professor* (Lee-Jahnke, 2005). The triad elements result in three processes that, the author views, are necessary for constructing neural networks crucial for quasi-automatic interpreting task. These are: (1) learning process—occurs between student and knowledge; (2) training process—takes place between professor and student; and (3) teaching process—occurs between professor and knowledge (Lee-Jahnke, 2005, p. 361).

Recurrent Networks in Interpreting Classroom

The author contends that the concept "recurrent network" in connectionism is workable for describing the interaction of those three above-mentioned elements. Cognition in recurrent models is not limited to specific neurons nor to the individual mind as a sum of its associations, but it involves interactions between interconnected brains in a network of abstract ties. While levy (1997) used the concept "collective intelligence" to designate the multi-agent networking, Hutchins (1995) referred to it as a "distributed cognition".

In the class environment, the processes of *learning*, *training*, and *teaching* are to be understood in terms of the recurrent multi-agent nets for the fact that it involves many agents/units communicate with each other once knowledge is activated: That is, it involves auto-associative nets that are students neural networks in addition to that of the teacher. Apparently, the teacher uses supervising procedures in propagating knowledge that might be activated by any unit, be it a teacher or a student. Then, each unit receives the activation from the external source outside the neural net. The activation, which reflects the term *short term memory* in connectionist theorizing*, spreads automatically among the units of the net of recurrent networks, for they are interconnected in an auto-associative recurrent manner (Overwalle & Heylighen, 2006, pp. 606-608). Thus, agents learn from each other since their neuronal nodes are interconnected. We conjecture that all the three processes: *Learning process*, *training process*, and *teaching process* are interactively involved in such classroom recurrent networking.

In activating knowledge in classroom, we shall be mindful that the target is the neural nodes network with an explicit aim of establishing associations between them in congruence with those maintained among chunks or stretches of knowledge under activation. Classroom didactic activities, from this connectionist lens, are constructive in their essence for they enable the interconnected units (i.e., students) to turn the quasi-randomly connected neurons into well constructed connections. The author pinpoints that the activation is distributed among the network of students through activating mirror neurons of the network. Mirror neurons make students observe their actions and performance and stimulate each other in a multi-agent learning process while they are armed with meta-cognitive skills in their meta-memory. Each student adopts his/her learned strategies to have self- regulation and self- monitoring (Flavell, 1976).

Knowledge and Neural Networks

Knowledge network organizing the world patterns is represented via the linguistic cohesive associations maintained among the syntactic units that implement the semantic system (Cilliers, 1988); our "conceptual categories give rise to linguistic categories, which enable us not only to communicate, but also to guide our understanding of the world" (Lee-Jahnke, 2005, p. 361). An interpreter's mental system exists within the

network of knowledge; it neurally shapes its knowledge out of that network's components through changing the values of weight. When we decode an input or speak an output, we reflect the level of our neural connectionism. Hinton (1989) stated this conception as follows:

> All the long-term knowledge in a connectionist network is encoded by where the connections are or by their weights, so learning consists of changing the weights or adding or removing connections. The short-term knowledge of the network is normally encoded by the states of the units, but some models also have fast-changing temporary weights or thresholds that can be used to encode temporary contexts or bindings. (p. 186)

Pursuing such line of reasoning, the ultimate purpose behind activating knowledge network in classroom shall be the building of a high level of long-term neural plasticity necessary for performing the interpreting task. Knowledge network in its socio-economic and socio-cultural frame work consists of a number of component nodes/units including actors (i.e., individuals, groups, organizations, and relations between those actors within social networks, that can be characterized by form, content, and intensity). They also include resources (used by actors in relationships) in addition to institutional properties, i.e., structural and cultural dimensions such as control mechanisms, rules and norms, standard operation procedures and communication patterns (see Seufert, Krogh, & Bach, 1999).

The interpreter's neural network, following networks theory, has to echo the network of knowledge represented as exemplars (i.e., items, instances). These are institutional and socio-cultural nodes as large coherent input. Such input enables interpreters, as embodied agents, to interpret within the network of interconnected organizational and associative nodes constituting our micro and macro-cosmopolitan world (Cronin, 2006). They are required to construct a neuro-cognitive representation to those nodes and conceptual relationships among them, including enablement, part-whole conception, analogy, inheritance, as well as other knowledge patterns.

Training, Constructivism, and Connectionism

In his input hypothesis, Krashen (1985) postulated that learners are in need to a comprehensible input data to processes information, which is an equivalent in connectionism to exemplar nodes without which weight of trust cannot be achieved. Learners cannot process an input that contains linguistic features than their current knowledge (Krashen, 1985). Interpreting students have to accumulate a number of knowledge units to use them as prior knowledge input in their schemata that would enable them to process what Krashen (2003) terms "comprehensible input" into a congruent "comprehensible output". The accumulation of a large network of interconnected neural units must be achieved in a piecemeal processor, using Shastri (1999) terms, "long-term potentiation". In similar vein, recent studies (Tse et al., 2007, 2011), as McClelland (2013) pointed out, showed that neocortical circuits can rapidly acquire new neural associations that are consistent with prior knowledge.

The network associations are neutrally established through what Lakoff (1999) postulated to be the aspects of conceptual structures as embodied neural properties: *image schemas*, *frame semantics*, *conceptual metaphor*, *radial categories*, *basic-level categories*, and *mental spaces* within which we organize knowledge reflecting the world. He stated that these conceptual systems emerge within the mind as embodied results of neural representations. On this basis, the learning networks training for being interpreters engage in subjective projects using selective propagation caused by trust weights principle. The subjective neural associations conception supports the constructivist approach of learning that:

> … The learner is building an internal representation of knowledge, a personal interpretation of experience. This representation is constantly open to change, its structure and linkages forming the foundation to which other knowledge structures are appended. Learning is an active process in which meaning is developed on the basis of experience… Conceptual growth comes from the sharing of multiple perspectives and simultaneous changing of our internal representations in response to those perspectives as well as through cumulative experience. (Bednar, Cunningham, Duffy, & Perry, 1991, p. 21)

There are three inter-complementary meta-cognitive elements that a student is supposed to have constructed in his/her mental network: (1) *Metacognitive knowledge* which has to do with an agent's knowledge about his and others cognitive process; (2) *Metacognitive regulation* that is tantamount to controlling the process of learning; and (3) *Metacognitive experiences* which is related to the current endeavor (Flavell, 1979). Resting on connectionism, such metacognitive parameters in metacognitive memory, one might assume, are emergent states.

Interpreting as an Emergent State Proportional to Training

Taking for granted the premise that any cognitive behavior is the emergence of global states in a network of interconnected elements (McClelland, 2010) going under transformations due to the internal history of the system and its environment, interpreting behavior would logically be deemed an emergent global phenomenon. To make it clear, in proto-emergenism model, mental agents aggregate in mental conspiracies in memory representations of items to produce emergent category representations corresponding to individual items or experiences (i.e., exemplars in exemplar models by Medin and Shaffer, 1978) as McClelland noticed (2010). In such conception, the world is inseparable from the neural structures embodying the neural global states within which neural assemblies compete with each other to design conscious phenomena such as those related to interpreting. In distributed connections models, cognitive nodes contain units that stand for individual things (McClelland, 2010), including conceptual properties if we follow Shastri (1999). Building on Hopefield, Rumelhart, Hinton, and McClelland (1986) used the concept schemata to refer to those emergent states of neural networks. The aggregate co-occurrence statistics of properties of many schemata rooms in the neural system are represented by the connection weights (see McClelland, 2010, pp. 756-758).

Training should place major focus on neuro-linguistic factors for preparing efficient interpreters "taking into account the fact that a major part of the work depends on the self-training of the students" (Kornakov, 2000, p. 248) in constant connection with other networks (i.e., other students) to stimulate interpreting situations. Neuro-linguistic interactive training enables students to strengthen the ties between their cognitive nodes in a sophisticated neural architecture reflecting the conceptual categories and common patterns needed for interpreting as a neutral-chemically emergent state. Ellis (2002) pointed out that the acquisition of common patterns and categories is driven by the principle of frequency. Interpreters acquire the common patterns through practice and experience that would expose them to large input. Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976, pp. 382-493) noted that experience refines one's capacity to categorize the world into structured categories.

Practice and Neuro-plasticity

Practice is of vital importance for establishing the interpreter's network capable of mediating messages in the spot. It teaches "neurons to fire and hire together" and makes Hebb's assemblies in local and distributed circuits. Everything we do modifies our brains physically thanks to neuro-plasticity. Anderson (2009) argued that creativity has to do with neuro-plasticity, with mobilizing different motor and cognitive representations of memory. This includes the capacity to switch in a flexible manner from an experience to another when

analyzing information. Neuro-plasticity enables interpreters to make associations easily simply because in their mental systems, a large number of neurons have been recruited through learning and practice to perform the interpreting task. In the context of learning, Anderson (2009, p. 21) gleaned to the conclusion that students with higher networks perform better than those with weak connection.

Similarly, representing the recruitment learning framework, Shastri (1999, p. 1) referred to the process of learning to be based on the 'quasi-randomly' connected networks of neural nodes. In the mental architecture, there are two kinds of neural nodes: recruited and non-recruited cells. The recruited ones "have acquired a distinct meaning of functionality by virtue of their strong interconnections to other recruited nodes or sonsori-motor nodes" (Shastri, 1999, p. 1). Whereas, the free nodes have weak association with other free or recruited nodes or sensory-motor nodes, waiting to be recruited for participating in neural activity by representing new items or concepts. Shastri (1999) provided an illustration how those are recruited function, stating that "when a novel concept Y is a conjunct of existing concept x1 and x2 in long term memory" by means of recruiting the free nodes that receive connection weight "from both x1 and x2" (p. 1); a number of nodes get recruitment through the process of strengthening the synaptic weight. The node that "receives a large number of links can potentially participate in a large number of functional circuits" (Shastri, 1999, p. 7). The recruitment process makes transformation in the quasi-random network of nodes into structures of nodes to perform some specific functionalities. This framework might have certain applicability for training simultaneous interpreters in that it offers a roadmap to envision an interpreter as a network of recruited units. Hence, the more recruited cells in bilingualism and conceptual, declarative and procedural systems the mental society of an interpreter contains, the more effective he/she is.

Controlled and Automatic Processing in Interpreting

We assume that an automatic processing in interpreting necessarily requires recruiting a large number of neural interconnections we have discussed throughout our argumentation. In drawing a line on the controlled and automatic processing, Nikolic (1999, p. 7) demonstrated that prolonged training on doing something leads to the quasi-automatic processing. It enables one to overcome the capacity limits of WM (Working Memory) in terms of the items that it can hold, the number of stimuli that one can trigger in addition to that of categories in absolute judgments when processing becomes automatic (Miller & Campbell, 1953, for the size of WM 7 +/- 2 for controlled processing), automatic processing requires shorter time reaction to processed material (Nikolic, 1998, p. 8). P. Padilla, Bajo, Cañas, and F. Padilla (1995) demonstrated that professional interpreters showed higher WM capacity span contrary to students. That high capacity permits interpreters to meet the demands of interpreting simultaneously, avoiding articulatory suppressing effect (Padilla, 1995, p. 208). To illustrate the dialectical overlap between controlled and non-controlled processing, Nikolic (1999) enunciated that: "Controlled processing can control the automatic processes. Controlled processes can trigger automatic ones, shut them off or retrieve them into consciousness when the situation requires it (i.e., novel situations)" (p. 9). Hulstijn (2003) explained the dual usage of capacity saying that fluent reading and listening are automatic at what he saw to be "lower levels of word cognition" (p. 419) and sentence parsing. The capacity is left free to focus attention on the higher information levels relating to semantics and content (Harrington, 2001; Rost, 2002). Interpreters, as listeners are exposed to concatenated speech and they are required to distinguish between the borderlines, that is, between interconnected local units in local networks. Through being exposed to extensive training, the mental network can distinguish the segments of speech easily and rapidly.

Conclusions

Connectionist modeling of cognitive processing is a promising field that is replete with implications and applications for many fields. Interpreting and translation might benefit from such field in a scientific manner on the basis of experiments. Throughout the lines of argumentation, we attempted to provide a connectionist framework for modeling interpreting that demands, above all, the building of a huge network of nodes through learning, training and experience. We tried to bridge the distance between connectionism and neural networks, as a paradigm, and interpreting in particular and translation in general. Thus, the paper provides a new scope of research in connectionist models, which might propel inquiry in interpreting that has always been a challenging subject matter in both theory and practice.

References

Anderson, O. R. (2009). Neurocognitive theory and constructivism in science education: A review of neurobiological, cognitive and cultural perspectives. *Brunei International Journal of Mathematics and Science Education, 1,* 1-32.

Bechtel, W., & Abrahamsen, A. A. (1991). *Connectionism and the mind: An introduction to parallel processing in networks.* Oxford: Basil Blackwell.

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1992). Theory into practice: How do we link?. In T. M. Duffy, & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction* (pp. 17-34). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cilliers, P. (1988). *Complexity & postmodernism: Understanding complex systems.* London: Routledge.

Cronin, M. (2006). *Translation and identity.* London and New York: Routledge, Taylor & Francis Group.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Learning, 24,* 143-188.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Hillsdale, N.J.: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-development inquiry. *American Psychologist, 34,* 906-911.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28,* 3-71.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics, 36,* 193-202.

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). New York: Academic Press.

Harrington, M. (2001). Sentence processing. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 91-124). Cambridge: Cambridge University Press.

Hebb, D. O. (1949). *The organization of behavior: A neurophysiological theory.* New York: Wiley.

Hinton, G. E. (1981). Implementing semantic networks in parallel hardware. In G. E. Hinton, & J. A. Anderson (Eds.), *Parallel models of associative memory.* Hillsdale, N.J.: Erlbaum.

Hinton, G. E. (1989). Connectionist learning procedures (Technical report CMU-CS-87-115, Carnegie-Mellon University, Pittsburgh, P.A.).

Hulstijn, J. H. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *Computer Assisted Language Learning, 16*(5), 413-425.

Hutchins, E. (1995). *Cognition in the wild* (p. 381). Cambridge, M.A.: The MIT Press.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526-2528.

Kornakov, P. (2000). Five principles and five skills for training interpreters. *META, 45*(2), 241-248.

Krashen, S. (1985). *The input hypothesis.* Beverly Hills: Laredo Publishing Company.

Krashen, S. (2003). *Explorations in language acquisition and use.* Portsmouth: Heinemann.

Lakoff, G. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought.* New York: Basic Books.

Lee-Jahnke, H. (2005). New cognitive approaches in process-oriented translation training. *Translators' Journal, 50*(2), 359-377.

Leonardo, F., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science, 308*(5722), 662-667.

Levy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. New York: Michigan University.

McClelland, J. L. (2010). Emergence in cognitive science. *Topics in Cognitive Science, 2,* 751-770.

McClelland, J. L. (2013, August). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General, 142*(4), 1190-1210.

McClelland, J. L., Rumelhart, D. E., & The PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 2: Psychological and biological models*. Cambridge, M.A.: Bradford.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207-238.

Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *Journal of Abnormal and Social Psychology, 59,* 1-9.

Minsky, M. L. (1986). *Society of mind*. New York: Simon and Schuster.

Nikolic, D. (1999). Neurophysiological mechanisms underlying the distinction between automatic and controlled processes (Ph.D. thesis, University of Oklahoma, USA).

Oberman, L., & Ramachandran, V. S. (2009). Reflections on the mirror neuron system: Their evolutionary functions beyond motor representation. In J. A. Pineda (Ed.), *Mirror neuron systems: The role of mirroring processes in social cognition* (pp. 39-62). New York: Humana Press.

Opie, J. (1998). Consciousness: A connectionist perspective (Doctoral thesis, University of Adelaide).

Overwalle, F. V., & Heylighen, F. (2006). Talking nets: A multivalent connectionist approach to communication and trust between individuals. *Psychological Review, 113,* 606-627.

Padilla, P. (1999). Proposal for a cognitive theory of translation and interpreting: A methodology for future empirical research. *The Interpreters' Newsletter, 9,* 61-78.

Padilla, P., Bajo, M. T., Cañas, J. J., & Padilla, F. (1995). Cognitive processes of memory in simultaneous interpretation. In J. Tommola (Ed.), *Topics in interpreting research* (pp. 61-71). Turku: University of Turku Press.

Perkins, D. N. (1981). *The mind's best work*. Cambridge: Harvard University Press.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Rost, M. (2002). *Teaching and researching listening*. Harlow, England: Pearson Education.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge M.A.: MIT Press/Bradford Books.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundations* (pp. 45-76). Cambridge, M.A.: MIT Press.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundations*. Cambridge, M.A.: Bradford.

Schippers, M., Marleen, B., Roebroeck, A., Renken, R., Nanetti, L., & Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences of United States of America-PNAS, 107*(20), 9388-9393.

Seufert, A., Krogh, G. V., & Bach, A. (1999). Towards knowledge networking. *Journal of Knowledge Management, 3*(3), 180-190.

Shastri, L. (1999). Biological grounding of recruitment learning and vicinal algorithms in long-term potentiation. *Emergent Neural Computational Architectures Based on Neuroscience Lecture Notes in Computer Science*. Berkeley, California, USA.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., & Morris, R. G. M. (2007). Schemas and memory consolidation. *Science, 316,* 76-82. doi:10.1126/science.1135935.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., & Morris, R. G. M. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science, 333,* 891-895. doi:10.1126/science.1205274.