Computational Natural Language Learning:
±20years ±Data ±Features ±Multimodal ±Bioplausible

David M. W. Powers

Artificial Intelligence and Cognitive Science Group
School of Computer Science, Engineering and Mathematics
Flinders University, Adelaide, South Australia
David.Powers@flinders.edu.au

Abstract

This speech celebrates the 20th anniversary of the CoNLL conference and looks back 20 years before CoNLL and 20 years into the future in an attempt to paint a longterm roadmap of Computational Natural Language Learning. The founders of CoNLL agonized hard and long over what to call our nascent field, and how to ensure that we kept all the interdisciplinary diversity that we had in those early days, including preserving the richness of views in a field that encompassed many controversies. We will explore this diversity with a focus on new directions that are developing; we will reflect on the changing nature of our technology including the deceleration of Moore's Law and the emergence of Big Data; and we will consider the impact of and on ubiquitous technologies ranging from wearables to multimedia, from intelligent phones to driverless cars.

1 Introduction

Machine Learning has moved out of the lab and into the field, and the explosion of language-related learning-research is a massive part of this. Companies like Google, Facebook and Amazon, as well as IBM, Apple and Microsoft, have emerged as huge players in our playground, and important sponsors!

One of the most important changes we've seen over the last two decades is that we've fallen off the Moore's Law curve as far as single core processors are concerned[1] - for CoNLL's first decade SPECInt performance was a around 50%, for our second decade it was more like 20% and according to NVidia it looks like being more like 5% for the next decade. The emphasis must now be on working smart and processing in parallel, but at the same time we are introducing much higher software and data overheads with managed code and extensive markup.

20 years ago few households had a personal computer or a mobile phone, while today few western families wouldn't possess a range of equipment from phones to pads to MP-players to cars and the list goes on with microwaves, washing machines and airconditioners, averaging over 20 processing units per household. All of these can be expected to be part of the Internet of Things in the very near future. Recent developments are not focussing on doing more with our computers, but using our computers more flexibly and universally, with "the cloud" and "hands free" operation being major drivers of the technology race, with language technology "in the cloud" helping to deal with spoken or typed operations that would until recently have been regarded as office functions, with call centres being outsourced to computers, with IBM's Watson challenging the quizmasters at their own game, in their own language, and moving onto a wide range of "Cognitive Computing" applications. I will focus initially on this changing context.

The Association for Computational Linguistics reached 50 recently, starting with a strong focus on Machine Translation that was originally part of the name, and SIGNLL affiliated itself with ACL in 1993 to reflect our driving interest in language[2]: "Computational Linguistics is the scientific study of language from a computational perspective...providing computational models of various kinds of linguistic phenomena."

---

[1] http://preshing.com/20120208/a-look-back-at-single-threaded-cpu-performance

[2] http://www.aclweb.org/portal/what-is-cl

Artificial Intelligence is at least 60 years old, and SIGNLL adopted the "Natural Language" teminology of AI to make the connection to a rich history of AI research in NL that seldom, however, made use of Machine Learning, and rarely considered the Linguistic and Psycholinguistic problems of how human language works and how children learn language. But we didn't want NLL to be seen as just another application of Machine Learning, particularly as we (and our sister SIG-DAT) developed increasingly large and expensive resources, including tagged corpora. Thus we moved away from the name "Machine Learning of Natural Language" which was the title of the book based on my own PhD thesis (Powers and Turk, 1989) while the AAAI event I organized added "and Ontology" (Powers, 1991a).

It is worth citing the aims of SIGNLL as presented to ACL when we formalized our affiliation as a SIG, and noting that all areas of language and ontology, linguistics and psycholinguistics, are intended to be in scope for SIGNLL and CoNLL:

SIGNLL aims to promote research in:

- automated acquisition of syntax, morphology and phonology

- automated acquisition of semantic and ontological structure

- automated acquisition of inter-linguistic correspondences

- learning to recognize or produce spoken and written forms

- modelling human language acquisition theory and processes[3]

I reviewed over the last couple of days CoNLL's two full decades of proceedings, but its hard to single out some for special attention when others are equally worthy. So I will instead review from the perspective of a set of "±features", of *computation* and *language* and *learning*, that are I feel becoming increasingly important to CoNLL, and will set the stage for the next two decades. In relation to the SIGNLL aims, my feature-driven wandering will visit each of them broadly, but I see huge opportunity for crossover between these five categories and their internal subareas. So rather than separating them arbitrarily, I will draw them

together, emphasizing aspects that I think are important to the future of the field as a whole. I particularly want to encourage interdisciplinary collaboration and a Computational Cognitive Science (CoCoSci) that not only seeks to engineer better technologies, but also seeks to exploit and model and inform our understanding of human language and cognition.

From the beginning, SIGNLL and CoNLL promoted and welcomed interdisciplinary researchers and collaborations, but today most of us have a primary background in computing, and we attract mainly computing and engineering students. While the founders of CoNLL all had very interdisciplinary background, it is a daunting prospect to try to keep up with related fields when our own has grown so massive. In conclusion, I will make some suggestions as to how we can address this.

2 ±Applications

One of the new buzzwords the last two decades have brought is "Applications" or "Apps". When we set up SIGNLL we boringly referred to "software and data", and APIs were libraries, and tools and applications weren't distinguished; personal computers didn't have real user level multiprocessing capabilities, and it was only after founding SIGNLL in 1993 that the path through Windows 95/98/XP opened up the PC world, and Apple made its resurgence with the iMac and OSX, while SUN with Java started to blur the distinction between computer and network.

Software as a commodity is probably the biggest single change to the scene since CoNLL was started, and this is reflected in CoNLL through our sponsorships and the major corporate labs in this area. But the new challenge for us is mobility in a borderless world-wide cloud... Already the World Wide Web, and Google, with the web as a data resource, and information retrieval as a major application and focus for research, are putting a new complexion on our field. And now the mobile age has arrived and Apps on phones offer unprecedented ubiquitous interactivity, with new demands on language technology ranging from speech recognition to automated help desk, to instant speakable machine translation, to educational games and location-aware monitoring and advice.

The older focus on Machine Translation and Speech Recognition, which brought AI into dis-

---

[3]http://ifarm.nl/signll/about/

2

Figure 1: ECAs for teaching children in 2008 & 2010 and assisting the elderly in 2010 & 2015 (AnnaCares.com)

repute in the 1960s and 1970s with overconfident claims and predictions, has now resurged with successes due largely to the Big Data resources available (particularly to search engine companies) as well as the huge parallel computational resources available (particularly to search engine companies) and the ability to drive our statistical machine learning or artificial network tools harder and deeper (as exploited by the same companies). Now the challenge is to get these technologies into a mobile format that is interactive and dynamic, location and activity aware, and not so dependent on instant cheap access to the cloud.

We also have new and rich opportunity to collect data from these ubiquitous, multisensor devices and their increasingly intelligent Apps. We can do eyetracking, we are already developing applications that improve speech recognition and machine translation by offering choices, and a major research direction of mine is Unconscious Computer Interface where these choices and corrections are made below the level of consciousness, like our everyday articulation choices in speech and writing. This tracking and choice data is an immensely rich resource, but is also associated with ethical and privacy considerations, so it is best used for dynamic online training on device.

Traditionally, CoNLL's focus has been on *understanding* language, although Machine Translation has transferred to same-language paraphrase and summarization, and multilingual representation can be useful for monolingual *generation*. Additionally, CoNLL's focus has been on *learning* language, but from the beginning we have also interacted with the Computer Assisted Language Learning community, and there are interesting synergies as the resources and techniques developed for language learning are turned into *teaching* Apps. We've developed systems for teaching English (Powers et al., 2008; Anderson et al., 2008; Chiu et al., 2012; Anderson et al., 2012), teaching social skills to children with autism (Milne et al., 2010), as well as applications in aged care and health space[4] (Powers et al., 2010). These recent systems incorporate bots and games and simulated environments into talking/thinking/teaching heads and Embodied Conversational Agents (ECAs), while allowing us to understand the effect of different features of the system on human acceptability, understanding and

---

[4] http://annacares.com

3

learning (Stevens et al., 2016). An interesting 'uncanny valley' aspect of this has been our regression from human-indistinguishable disembodied heads to cartoon-like embodied agents, while at the same time controlling age, sex and ethnicity components as determined by focus groups for each application (Fig. 1).

Could we put more Mobile Apps into CoNLL?

3 ±Parallelism

Parallel computers, distributed computers, the internet and the cloud, are complemented by naturally parallel paradigms, include the Functional/Logic Programming paradigms, or the MapReduce paradigm that seems to have taken on a life of its own, and of course the Parallel Distributed Processing of the Artificial Neural Network. The brain is of course massively parallel, but at the same time spoken language and conscious thought are both intrinsically sequential.

When we founded SIGNLL and CoNLL, there were several competing massive custom-designed supercomputers out there, but the time and dollar cost of the custom design meant that they were seldom really that much ahead of the mainstream servers. It was actually animation, games and graphical hardware that drove parallelism to the mainstream, with the incorporation of GPUs in most modern PCs, and the development of the GPGPU and Intel's Phi, bringing enormous power to our fingertips - as well as NVidia's supercomputer (and deep learning network) in a box.[5]

Many of our *low level* operations can be performed in parallel - as elementary keyword lookups are performed in search, with AND and OR operations turning into streaming INTERSECT and UNION operations. Semantic networks and activation models are naturally parallel, as are Artificial Neural Networks. Yet the regular systolic nature of current ANNs is highly suited to GPU architectures, and there is significant challenge associated with exploiting them for more *ad hoc* network structures.

On the other hand *high level* modularity and multimodality naturally give rise to components that effectively run in parallel but need to coordinate efficiently. For example our HeadX (Luerssen et al., 2010) employs both shared memory and socket streams to coordinate speech and face synthesis while managing keyboard, mouse and speech interactivity. Although our focus on language might seems straightforward when we consider text, the natural form of language is speech, and the natural grounding of semantics in our physical, social and cultural world. Processing video for person, face, lip and eye tracking is an increasing load as NLL moves out of the lab and into a multimodal world, sensed through a phone with limited power in terms of both processing and battery capacity.

Furthermore, much of our limited power, along with the subtle features embedded in our data, is lost in repeated compression and decompression, as well as dealing with massive amounts of multiple kinds markup that must be selectively processed or skipped when embedded in a sequential stream. Conversely, access to and addition of additional streams of annotation is more efficiently achieved in a distributed parallel way,[6] and we are not only producing synchronized parallel streams for individual microphones and cameras, but for subregions and macroblocks, and image frames, and unidirectional and bidirectional prediction frames that provide information about motion and allow utilizing information about attention.

Parallelism will be increasingly key to CoNLL.

4 ±Data

The research of SIGNLL, like its sister SIGDAT, is driven by data, often Big Data. My own interest is more on unsupervised learning (Powers, 1984; Powers, 1991b; Leibbrandt and Powers, 2012), but even unsupervised systems need to be evaluated, and formal evaluation was a missing element in our early research. For supervised learning, annotated data is essential, and the Penn Treebank (Marcus et al., 1993) was a great resource in those early days and is still influential today.

In his presentation for the 10th anniversary of CoNLL, Walter Daelemans (2006) notes that there are huge costs in developing such corpora, that there are issues with annotator agreement, and that our trained systems might give high overall accuracies, or low error rates, but for key ambiguity problems error rates of 20-30% are common - and my own exploration of "problems" with my unsupervised learning using BNC2 for evaluation actually showed that BNC tags were wrong as much as

---

[5]http://www.nvidia.com/object/
deep-learning-system.html

[6]http://alveo.edu.au/

60% of the time for certain specialized cases (e.g. for the PoS labeling of "work" in "going to work" the *corpus* tags have an accuracy of only 39.1%).

We need more of a focus on understanding our data, ensuring it is clean and accurate, and that the numbers that we use to characterize its accuracy are actually reflective of the hard cases rather than just the easy cases (Entwisle and Powers, 1998). Zipf's law tells us that the top 150 words of English suffice to account for half the tokens of running English text, and these and other functional words, as well as words with unambiguous or highly predictive affixes, quickly leads us to what in normal accuracy terms would be regarded as a creditable performance. This is a particular case of the 80:20 rule - the first 80% of accuracy is achieved very easily. Another sign of the problem with our tagging and parsing is limiting consideration to sentences of 40 words or less.

A further problem with corpora is that they wear out quickly! That is datasets get overused, and we treat them as an ML resource where we try to tweak every last percentage point of accuracy, using every trick in the book, but in the end with 0.05 significance testing one in twenty researchers is likely to show an improvement over their baseline system just by being different - adding noise can do the trick! As Cohen (1994) suggests, Statistical Hypothesis Inference Testing is worthy of its acronym.

We need further work on developing good corpora, including multimodal corpora where this is a richer basis for unsupervised learning, and for automated validation and correction, that is not just text corpora with syntactic or semantic annotations, but corpora involving audio-visual speech and longitudinal contextual data (Roy, 2009). One driving force of this is having the same data available for CoNLL as babies have when they are learning data, but another is the Memories4Life GrandChallenge of capturing all the important moments of our lives and exploiting the ubiquitous computing and audiovisual resources of today's mobile devices, with practical applications already being developed for alleviating dementia.

Data will continue to be pivotal to CoNLL.

5 ±Features

In the previous section, I mentioned the issue with tags for supervised training and/or evaluation of systems. Now I want to focus more on the development and evaluation of the utility of unsupervised features, and to connect back to our Apps as I propose once again an approach that allows this through the use of application-oriented evaluation Powers(1991a; 2005) - don't try to evaluate tags or features or structures directly, but comparatively in real world (or in early stages toy world) applications. Often when we talk about features we think specifically of features in visual, auditory or other signals. This is indeed what I am talking about – linguistic features derive from these modalities.

One special case is unsupervised discovery of tags, including both syntactic (PoS) and semantic tags. Much of the "unsupervised" learning we see at CoNLL already assumes a "linguistically" motivated set of tags, and often a pre-tagged corpus, for which there is no psychological, neurological or other empirical evidence other than professors/linguists have managed to formalize some kind of system and students/annotators have managed to learn the system.

If we think of the tagging and parsing with the tags as their "applications", we then have a special case of the proposal in the previous section to compare systems based on evaluation of the application, but in the end I'd still like to take it to a real world application eventually - after all nobody actually knows (yet) how we process language in our heads, and understanding this is our fifth goal. The same learning techniques applied to different sets of tags gives us a comparison - but we must be sure to use an evaluation technique that doesn't implicitly bias when we have different numbers of classes (Powers, 2008). Furthermore when we do compare techniques or parameterizations or tweak biases and thresholds, we should be wondering whether the difference are real and bioplausible, and whether they are universal or an artefact of specific data.

Tags are in fact a special kind of feature in that they are descrete and there are few enough of them to give them names (outside of Categorial Grammar). Letters of the alphabet (graphemes) also belong to this special subclass by definition, and similarly by assumption the same applies to other emic units (e.g. phonemes and morphemes, including affixes and functional words). The situation is more difficult at the level of words and sentences - I actually have my doubts as to whether those are real units psychologically in pre-literate language, as they are defined somewhat arbitrarily by the placement of punctuation: why 'out of' and

not 'into' - and what is the current status of 'upto'? Is it 'today' or 'to-day' or 'to day' or 'the day'; is it 'one of' or 'one off', 'would have' or 'would of'? Do these units even have well-defined boundaries in speech? Can we tell which word a particular speech code vector actually belongs to? And all this is without descending to the etic level... There is an intrinsically fuzzy aspect to language, and it is hugely context dependent in a way that transcends the traditional phonetic, morphemic, syntactic and semantic levels implicit in CoNLL's first two aims. But it is in working towards the last two aims, in a ubiquitous, mobile, multimodal context, that we will effectively address the dynamics of language, and finally resolve all the fuzziness and ambiguity.

CoNLL will discover new vistas of features.

6   ±Multimodal

CoNLL has been very successful in relation to SIGNLL's first three "automated acquisition" aims! But what of the fourth "[hand]written and spoken" aim? And have we really got "semantics" and "ontology" under control as per the second aim? Since I first used it in the early 1980s, and incorporated it in my 1991 symposium title and the SIGNLL aims, the word "ontology" has come to mean something more like "taxonomy" rather than its traditional and etymological idea of "our understanding of the world". In the 1980s our focus was in trying to find an alternative to the word "semantics", which in practice was becoming "look it up in the dictionary" or "follow a link in a semantic net".

Feldman et al. (1990) introduced the idea of L0 as the basic bootstrap language task based on a simple toy world model, while Harnad (1987; 1990; 1991) used the term "symbol grounding" and argued strongly that even simulated worlds weren't enough, and earlier still (Hayes, 1979) had used the phrase "naive physics" to describe what he thought was needed.

Between 1984 and 1991 my students had developed and were using the Magrathea robot world extension to Prolog (Powers and Turk, 1989) but between 1995 and 1997 we built a physical humanoid baby that could crawl, feel touch to arms and legs, had omnidirectional auditory perception and stereovision, and could orient towards a touch or sound, and "feed" (charge and download) via a USB bottle/umbilical, but it was a brittle heavy system that was not suitable for a child to "mother" as originally envisaged (Powers, 2001), so that I eventually transitioned to the model of an Intelligent Room with half a dozen microphones and cameras.

Luc Steels (1995; 1997; 2003; 2015) adopted a simpler approach, transitioning from simple graphical animations, to turtle-like robots to a pair of cameras that viewed a "real world scene" that was constrained to be very simple (manipulation of cutout shapes on a board), and produced some very interesting interactions and learning.

There's another advantage to multimodal data, that is that you can use supervised techniques in a directed but unsupervised way. The 1980s idea of (holographic) autocoding of the input to self-organize features, can become a more efficient and bioplausible system where intramodal feature discovery and intermodal feature discovery are distinguished - for example visemes can be self-organized as facial/lip patterns that correlate to certain groups of phonemes. The use of multimodal autosupervision allows more bioplausible features to be self-organized, as well as facilitating a cognitive approach to learning phonology, morphology, syntax and semantics (Powers, 1997). An additional advantage is that techniques like eye-tracking and gaze-tracking can augment our user interfaces and help identify context or distinguish alternatives, boosting the accuracy of our NLL systems.

Today, Google Glass and Microsoft's Hololens (with Kinect-like 3D), are examples of the integration of multiple cameras and microphones, and a heads-up augmented-reality type display, into an efficient platform that can keep track of the 3D world in a way that will naturally complement speech capabilities as well as augment the capabilities of language learning with its richer data.

Multimodal will open a door to a new CoNLL.

7   ±Bioplausibility

So we are now up to the final "modeling human language acquisition" aim. In the end, language is a product of human biology and ecology, but Linguistics and Computational Linguistics have largely been developing without any input from Biology, Psychology or Neuroscience although there are interesting crossovers, and CoNLL has always strongly encouraged the modelling of human language acquisition theory and processes,

and we do get a trickle of papers with a Psycholinguistic flavour. But I would encourage the CoNLL community to look beyond Computational Linguistics and Artificial Intelligence to the evidence being amassed in Computational Neuroscience and Cognitive Psychology, and to seek to connect to people studying language and learning from these different perspectives.

There is no compelling reason we have to make our systems bioplausible, and just because we use a neural network doesn't necessarily make it bioplausible model. But there are advantages in taking on board this aim. Indeed the introduction of neural and computational plausibility revolutionized the behavioural and cognitive sciences, with terms like neurons and agents replacing vague concepts from earlier theories of Psychology and Philosophy as they talked about demons (Selfridge, 1959) and zombies (Dennett, 1995), with the zombie argument turning up in AI in the well known guise of Mary's Room (Jackson, 1986) and the Chinese Room (Searle, 1980). Philosophers tend to shift the focus from Turing's (1950) behavioural test of indistinguishability of language performance (as a surrogate for behaviour and cognition in general), to mind, consciousness, awareness and feelings, while Computational Neuropsychology seeks to model what we find in the brain and show how that can explain and reproduce human-like language and behaviour.

Part of SIGNLL and CoNLL's charter includes the understanding and modeling of the behaviour of another, the theory of mind, as a component that is absolutely necessary for conversation, for effective communication and learning, for understanding the affective, emotional and physical states of the person we are talking to, and for understanding the human factors of the interfaces we are building.

So beyond just looking at the latest work across CoCoSci, we could be looking at our language learning systems as scientific models in their own right, or looking at interdisciplinary theories as a basis for our systems, and presenting them in a way that makes them into behaviourally, biologically and computationally plausible hypotheses and theories that are testable by their predictions about human behaviour (Popper, 1934; Popper, 1963; Lakatos, 1970).

We thus encourage collaborations with other parts of Cognitive Science who can help us improve and extend our language learning models, as well as help others keep their theories computationally realistic. CoNLL system may well give insights to other disciplines, and certainly our methodologies can and should be utilized in cognitive programs, more over our tools can help with their data collection and behavioural analysis (Stevens et al., 2016).

CoNLL is key to unlocking the human psyche.

8 ±20 Years

SIGNLL and CoNLL were born *into* an environment where Cognitive Science had brought people out of their silos and interdisciplinary research and computational modeling were recognized as essential to a proper understanding of language and cognition. Indeed Cognitive Science was born *out of* the controversies over language and learning: what was innate, what was learned, and what was biologically and computationally feasible.

This is a story for another time, and indeed one of SIGNLL's sponsored workshops, Cognitive Aspects of Natural Language Learning (CACLL) is continuing the debate as another stream during CoNLL/ACL, and there are other relevant workshops on Morphology and Phonology, and Representations and Evaluations, etc. Workshops have driven advances in the application of unsupervised and semisupervised learning, as well as covering some of the more forward thinking application areas like Companionable Dialog Systems. We don't regard them as in competition with CoNLL, but we actively sponsor workshops as additional streams that allow a focus that is not possible in the main thread of CoNLL - our sponsored workshops are just as much a part of CoNLL as of ACL, and often they capture elements that are emerging in CoNLL and will be important in the future!

While our successes with taggers and parsers and semantic models, and their applications in information retrieval and machine translation and intelligent assistants, will remain the bread and butter for CoNLL, we actively encourage people to be proactive and propose workshops and tutorials that will broaden the background of CoNLL researchers as Speech and Language become pivotal in a mobile, wearable age of ubiquitious computing and communication. It is also impressive to see how important the shared tasks have been in setting directions, and becoming a focus for future years - and we aim to improve the way in which these are made available for future use, compari-

son and evaluation. Your suggestions about shared tasks will also be most welcome.

In the next 20 years, I expect ubiquitous longitudinal multimodal multiangle multidirectional data, wearable processing, and cloud connectivity, will multiply both the opportunities and the challenges for CoNLL, and of course lead to further successes and new technologies to wow the world one year, and be taken for granted the next.

I hope we will make similar progress in understanding the human that wears the tech, in exploring the similarities and differences between Artificial Intelligences and Human Intelligence.

CoNLL has an exciting future ahead.

Acknowledgments

References

Tom AF Anderson, Wu-Yuin Hwang, and Ching-Hua Hsieh. 2008. A study of a mobile collaborative learning system for chinese language learning. In *Proceedings of International Conference on Computers in Education*, pages 217–222.

Tom AF Anderson, Zhi-Hong Chen, Yean-Fu Wen, Marissa Milne, Adham Atyabi, Kenneth Treharne, Takeshi Matsumoto, Xi-Bin Jia, Martin Luerssen, Trent Lewis, Richard Leibbrandt, and David M. W. Powers. 2012. Thinking head mulsemedia. In *Multiple Sensorial Media Advances and Applications: New Developments in MulSeMedia*. IGI Global.

Yi-Hui Chiu, Tom A. F. Anderson, and David M. W. Powers. 2012. Dubbing of virtual embodied conversation agents for improving pronunciation. In *Fifteenth International CALL Conference*, pages 188–193. Taylor and Francis.

Jacob Cohen. 1994. The earth is round (p ¡ .05). *American Psychologist*, 12:997–1003.

Walter Daelemans. 2006. A mission for computational natural language learning. In *Proceedings of the Tenth Conference on Computational Natural Language Learning (CoNLL-X)*, pages 1–5, New York City, June. Association for Computational Linguistics.

Daniel Dennett. 1995. The unimagined preposterousness of zombies. *Journal of Consciousness Studies*, 2(4):322–326.

Jim Entwisle and David M. W. Powers. 1998. The present use of statistics in the evaluation of nlp parsers. In *Joint Conferences on New Methods in Language Processing and Computational Natural Language Learning (NeMLaP/CoNLL*, pages 215–224.

Jerome A. Feldman, George Lakoff, Andreas Stolcke, and Susan H. Weber. 1990. Miniature language acquisition: A touchstone for cognitive science. *Ann. Conf. of the Cog. Sci. Soc.*, pages 686–693.

Stevan Harnad, S. J. Hanson, and J. Lubin. 1991. Categorical perception and the evolution of supervised learning in neural nets. In David M. W. Powers, editor, *AAAI Spring Symposium on Machine Learning of Natural Language and Ontology*.

Stevan Harnad. 1987. *Categorical perception: The groundwork of Cognition*. Cambridge Univ. Press, New York NY.

Stevan Harnad. 1990. The symbol grounding problem. *Physica D*, 42:335–346.

Pat J. Hayes. 1979. The naive physics manifesto. *in Expert Systems in the Micro-electronics Age*, 1979:242–270.

Frank Jackson. 1986. What mary didn't know. *Journal of Philosophy*, 83:291–295.

Musgrave Lakatos, editor. 1970. *Criticism and the Growth of Knowledge*. Cambridge Univ. Press., Cambridge. ISBN 0-521-07826-1.

Richard E Leibbrandt and David M. W. Powers. 2012. Robust induction of parts-of-speech in child-directed language by co-clustering of words and contexts. In *Proceedings of the Joint Workshop on Unsupervised and Semi-Supervised Learning in NLP*, pages 44–54. Association for Computational Linguistics.

---

[7]`//computationalcognitivescience.com`

Martin Luerssen, Trent Lewis, and David Powers. 2010. Head x: Customizable audiovisual synthesis for a multi-purpose virtual head. In *Australasian Joint Conference on Artificial Intelligence*, pages 486–495. Springer.

Mitchell P Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini. 1993. Building a large annotated corpus of english: The penn treebank. *Computational linguistics*, 19(2):313–330.

Marissa Milne, Martin H Luerssen, Trent W Lewis, Richard E Leibbrandt, and David MW Powers. 2010. Development of a virtual agent based social tutor for children with autism spectrum disorders. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9. IEEE.

Karl Popper. 1934. *The Logic of Scientific Discovery*. Routledge. (1959 trans: Logik der Forschung).

Karl Popper. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper and Row.

David M. W. Powers and Chris Turk. 1989. *Machine Learning of Natural Language*. Springer.

David Powers, Richard Leibbrandt, Martin Luerssen, Trent Lewis, and Mike Lawson. 2008. Peta: a pedagogical embodied teaching agent. In *Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments*, page 60. ACM.

David MW Powers, Martin H Luerssen, Trent W Lewis, Richard E Leibbrandt, Marissa Milne, John Pashalis, and Kenneth Treharne. 2010. Mana for the ageing. In *Proceedings of the 2010 Workshop on Companionable Dialogue Systems*, pages 7–12. Association for Computational Linguistics.

David M. W. Powers. 1984. Natural language the natural way. *Computer Compacts*, 2(3):100–109.

David M. W. Powers. 1991a. Goals, issues and directions in machine learning of natural language and ontology. In David M. W. Powers, editor, *AAAI Spring Symposium on Machine Learning of Natural Language and Ontology*, pages 1–22.

David M. W. Powers. 1991b. How far can self-organization go? results in unsupervised language learning. In David M. W. Powers, editor, *AAAI Spring Symposium on Machine Learning of Natural Language and Ontology*, pages 131–136.

David M. W. Powers. 1997. *Perceptual Foundations for Cognitive Linguistics*. International Cognitive Linguistics Conference.

David M. W. Powers. 2001. The robot baby meets the intelligent room. In *AAAI Spring Symposium on Learning Grounded Representations*.

David M. W. Powers. 2005. Biologically-motivated machine learning of natural language and ontology a computational cognitive model. In *Workshop on Multi-Modal User Interface*. HCSNet/NICTA.

David M. W. Powers. 2008. Evaluation evaluation: A monte carlo study. In *European Conference on Artificial Intelligence*, pages 843–844.

David MW Powers. 2015. A critical time in computational cognitive science. *Computational Cognitive Science*, 1(1):1.

Deb Roy. 2009. New horizons in the study of child language acquisition. In *Proceedings of Interspeech*.

John Searle. 1980. Minds, brains and programs. *Behavioral and Brain Sciences*, 3(3):415–457.

Oliver G. Selfridge. 1959. Pandemonium: A paradigm for learning. In D. V. Blake and A. M. Uttley, editors, *Proceedings of the Symposium on Mechanisation of Thought Processes*, pages 511–529, London.

Luc Steels. 1995. A self-organizing spatial vocabulary. *Artificial life*, 2(3):319–332.

Luc Steels. 1997. The synthetic modeling of language origins. *Evolution of communication*, 1(1):1–34.

Luc Steels. 2003. Evolving grounded communication for robots. *Trends in cognitive sciences*, 7(7):308–312.

Luc Steels. 2015. *The talking heads experiment*. Language Science Press.

Catherine J. Stevens, Bronwyn Pinchbeck, Trent Lewis, Martin Luerssen, Darius Pfitzner, David M. W. Powers, Arman Abrahamyan, Yvonne Leung, and Guillaume Gibert. 2016. Mimicry and expressiveness of an eca in human-agent interaction: familiarity breeds content! *Computational Cognitive Science*, 2(1):1–14.

Alan Turing. 1950. Computing machinery and intelligence. *Mind*, LIX(236):433–460.