# Introduction: Cognitive Issues in Natural Language Processing

Thierry Poibeau, Shravan Vasishth

▶ **To cite this version:**

Thierry Poibeau, Shravan Vasishth. Introduction: Cognitive Issues in Natural Language Processing. Revue TAL : traitement automatique des langues, 2014, Traitement Automatique des Langues et Sciences Cognitives, 55 (3), pp.7-19. 10.1201/b21583-2 . hal-01386307

**HAL Id: hal-01386307**
**https://hal.science/hal-01386307**

Submitted on 23 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Introduction: Cognitive Issues in Natural Language Processing

**Thierry Poibeau*** — **Shravan Vasishth****

* *Laboratoire LATTICE — CNRS, École normale supérieure and Université Sorbonne Nouvelle, France*
** *University of Potsdam, Germany*

## 1. Introduction

From the very beginning, pioneers in computer sciences have tried to establish a link between practical issues and more fundamental ones, like language processing in the brain, knowledge representation or the nature of communication. Modeling language was supposed to open a window on the brain, or at least give an idea of how things can be processed by the brain. The famous Turing Test (Turing, 1950; Peters and Epstein, 2004) itself was not directly about language modeling, but mimicking a conversation through a computer that was considered as a proof of intelligence, i.e. the fact that computers could interpret utterances, infer new information and produce relevant responses.

These concerns never disappeared but the field quickly moved towards more applied research interests. It was especially the case for machine translation, as soon as in the 1950s, since it was important then to show that usable results could be obtained rapidly (Hutchins, 1986; Hutchins, 2000). However, there has been a continuous interest since the Second World War for what could be called "cognitive issues". Even the deception about the poor performances of machine translation systems in the 1960s led to reflections about language complexity and whether it was possible to model this complexity or not. In the 1960s and 1970s, Robert Dreyfus' works have for example been largely influenced by the results of programmes dealing with computers and languages (Dreyfus, 1965; Dreyfus, 1972).

TAL. Volume 55 – n° 3/2014, pages 7 à 19

Dreyfus advocates, among other things, that a large part of language production is unconscious or, at least, cannot be modeled via explicit rule-based systems. Neural networks have then been seen as a possible solution to model unconscious reasoning: the architecture of natural language processing was supposed to offer a more sensible representation of the problem (Rumelhart and McClelland, 1986). It is now generally assumed that neural networks do not directly simulate the human brain but they allow to implement processes that are required to model language and other complex tasks.

It seems that this interest for cognitive issues decreased in the late 1980s and more dramatically during the 1990s. The need for practical and real world applications pushed forward more applied and engineering approaches, with a decreasing interest for cognitive issues, even if of course some researchers have always been active in this field. One sign of this situation is for example that even if recent conferences like the Conference on Natural Language Learning (CoNLL) or Empirical Methods in Natural Language Processing (EMNLP) always had a track on cognitive and psycholinguistic issues, very few submissions were generally received and even less accepted [1]. The two domains have then largely evolved separately, cognitive science researchers being maybe overwhelmed by the high technicality of NLP nowadays, while natural language processing researchers do not always see the added value of taking cognitive aspects into account for their problems.

However, things seem to be changing. Several workshops are now regularly held on these topics (see for example the "Cognitive Aspects of Computational Language Learning" workshop regularly held since 2007 (Villavicencio *et al.*, 2013) or the "Cognitive Modeling and Computational Linguistics" workshop during NAACL 2015). Natural language processing would benefit from a better understanding of human processes: as it has been said by several researchers recently, traditional machine learning approaches have brought rapid and important improvements in different natural language processing tasks but these successes may be "low hanging fruits" (Church, 2011), which means the field may have to face from now on more difficult problems (e.g. discourse planning, argumentative analysis, etc.) that would benefit from a better understanding of the processes involved in the brain.

Additionally, researchers are again interested in evaluating the relevance of their models according to a cognitive dimension. Many of the existing computational models attempt to study language tasks under cognitively plausible criteria (such as memory and processing limitations) that humans face (Blache, 2013). New machine learning techniques, especially deep learning, bring back to the front scene a new version of neural networks that seems both more powerful and more sound, from a technical as well as a cognitive point of view (Socher *et al.*, 2013). Last but not least, cognitive science also benefit and sometimes takes inspiration from computational models.

1. One should however note that more established conferences like ACL or COLING never had a cognitive track *per se*, even in the 1960s or 1970s. However, each year several papers were dealing with issues related to cognitive science, like the structure of the mental lexicon or the psychological plausibility of parsing algorithms.

## 2. Current research trends

In this section, we examine some recent research trends concerning language comprehension, language acquisition, language pathologies and language evolution. There are of course other domains where computational linguistics meets cognitive science but we think that these four areas have seen important improvements recently.

### 2.1. *Language comprehension*

Cognitive science has a strong tradition of developing models of cognitive processes using tools that originated in the natural language processing world. One of the early examples in sentence comprehension research is the work by Joshi and colleagues (Joshi, 1990; Rambow and Joshi, 1994), who developed an explanation of the differences in processing between crossed and nested dependencies in Dutch and German (Bach *et al.*, 1986) by explicitly linking a type of pushdown automaton, the Bottom-up Embedded Pushdown Automaton (BEPDA), to Tree Adjoining Grammar (TAG) (Joshi *et al.*, 2000). A unique aspect of their work was that they did not directly use TAG to develop a processing model, but defined an equivalence relationship between the automaton and TAG grammars.

Over the last fifteen years, an information-theoretic approach to language comprehension processes, led by (Hale, 2001; Hale, 2003; Hale, 2006) and Roger Levy (Levy, 2008), has generated a lot of interest, especially because it was able to explain certain empirical findings (Konieczny, 2000; Vasishth and Lewis, 2006) that needed additional assumptions to account for using classical working memory-based accounts (Just and Carpenter, 1980; Gibson, 2000). The Hale and Levy approach relied on large-scale probabilistic context free grammars derived from treebanks, and consists of computing a complexity metric, such as surprisal or entropy reduction, to characterize processing difficulty. The distinguishing feature of this body of work is that it formalizes the effect that the frequency of grammatical continuations has on expectations about upcoming words and phrases. The surprisal idea had already been investigated decades earlier, and has a rich tradition in EEG research (Kutas and Hillyard, 1984), but the Hale and Levy approach provided a formal basis for computing surprisal on a moment-by-moment basis. The Levy paper directly led to attempts to experimentally evaluate the predictions of the surprisal and entropy reduction ideas using large-scale grammars and eyetracking corpora or other reading data (Boston *et al.*, 2008; Demberg and Keller, 2008; Boston *et al.*, 2011; Frank, 2013). In addition, planned experiments also tested specific predictions of the surprisal account, with mixed results. Several studies have found evidence largely consistent with surprisal and related metrics (Staub, 2010; Levy and Keller, 2013; Frank *et al.*, 2015; Jäger *et al.*, 2015), but other research has shown mixed evidence consistent with both the surprisal and the classical working memory-based accounts (Levy *et al.*, 2013; Husain *et al.*, 2014; Husain *et al.*, 2015; Safavi *et al.*, 2015). As Levy pointed out in his 2008 paper, it is likely that both classes of explanation play a role in determining comprehension difficulty.

Nevertheless, a major contribution of the probabilistic grammar-based approach has been to give a formal foundation to the idea of expectation driving parsing processes. Future work would profit from building probabilistic parsing frameworks that include both kinds of constraints; this has been attempted in the past (Boston *et al.*, 2011), but the field would greatly benefit from more principled model development and model evaluation over cross-linguistic data. A cross-linguistic investigation is necessary as a corrective to the tendency to focus on English as the target language. A further gap in the field is the apparent lack of connection between the Hale and Levy formalization of expectation, and the rich body of work on surprisal in the EEG literature. Testing the detailed predictions of surprisal and related ideas using EEG is an important empirical test that is yet to be conducted. Another area where proposals such as surprisal and entropy reduction can make important contributions is individual-level variability in expectations, arising from differences in working memory capacity and the varying grammatical knowledge of the comprehender (Nicenboim *et al.*, 2015).

Ten years ago, it seemed that two areas where probabilistic models needed most work was to (i) demonstrate that moment-by-moment processing can be explained in terms of probabilistic expectation, and to (ii) attempt to model a broader range of phenomena. In the last few years, excellent progress has been made regarding the first point, but the empirical coverage of the probabilistic parsing account remains rather limited. For example, there is a large literature—and a great deal of empirical evidence—on interference effects arising in agreement attraction configurations (Wagers *et al.*, 2009), and antecedent-reflexive constructions where principles of the binding theory play a role (Dillon *et al.*, 2011). Probabilistic models obviously have little to say about phenomena that involve linguistic as opposed to probabilistic constraints. This highlights the need for developing more comprehensive models of cognitive processes such as sentence comprehension that take a broader view of the constraints that act on cognition.

### 2.2. *Language acquisition*

Language acquisition is another domain where computation models and cognitive sciences have a fruitful dialogue. To a certain extent, neuro-imaging has renewed the study of language by making it possible to directly observe processes in the brain (Dehaene, 2007) but for the rest, language is only known through direct production, i.e. language utterances. Therefore, the study of language acquisition by children is crucial, since it gives an overview on what vocabulary and structures are mastered first, what untypical constructions (when compared to adult speech) are used by children during learning, etc.

The study of language acquisition has seen important developments recently thanks to the development of automatic approaches. The main reason for these new approaches is the availability of large amont of data in the Childes database (MacWhinney, 1995), especially for English, but also for other languages, French being especially well represented here (Morgenstern and Parisse, 2012). Most record-

ings concern discussions between children and adults in real life situations (during dinner, bath, play, etc.). Videos are generally available so as to allow a multimodal approach to the study of language acquisition. Most corpora are annotated with morphosyntactic information (Sagae *et al.*, 2007), and parts of the corpus have received other kinds of annotation, including syntactic tags (Sagae *et al.*, 2010) and more rarely semantic or multimodal annotations (concerning gestures for example) (Morgenstern and Parisse, 2012). The Talkbank project has accumulated multimodal information in a more systematic way (MacWhinney, 2004). One should also mention Roy, who has systematically recorded the first years of his son, producing thousands of hours of videos (Roy, 2009). However, this mass of data is not structured and not tagged, making it hard to be used by researchers, on top of the ethic problem that this kind of recordings may pose. One specificity of most of these corpora is to provide longitudinal data. They generally make it possible to follow language development of a single child over several months.

Language acquisition studies concern mainly the development of the language ability in young children, the complexity of their speech in accordance with their age, and the comparison of children productions with adult ones (Tomasello, 2003). Corpora make it possible to observe statistical properties in children speech: size of the vocabulary used, average length of the utterances, constructions effectively used, etc. (Poibeau *et al.*, 2013). Some other variables can be observed like the influence of adult speech on children production as well as social and gender influence. Corpora make observation more direct but do not provide a direct answer to more theoretical questions like the famous "poverty of stimulus" (Chomsky, 1986; Pullum and Scholz, 2002; Legate and Yang, 2002). Lastly, researchers also tried to develop computational models that reproduce specific parts of the language acquisition process itself, like the acquisition of subcatgorization frames or of semantic categories (MacWhinney, 2004; Buttery and Korhonen, 2007; Alishahi and Stevenson, 2008).

Most data included in Childes concern on average one hour of recording per month per child or more rarely per week, which means only a very small fractions of the child production is available. Not all the situations are represented so that mixing different corpora together does not always solve the problem: data may still be biased, quite unbalanced and unrepresentative. The project from Roy mentioned above (Roy, 2009) was precisely supposed to overcome some of these limitations by providing a nearly exhaustive recording of the child input, but we have seen that this massive set of data is unstructured and therefore hard to process.

### 2.3. *Language pathologies*

The study of the production of people with language pathologies has also attracted a high interest in the last decades, see for one example among many others (Haarmann *et al.*, 1997). This field can be compared, to a certain extent, to the research done in language acquisition: the idea is to get an accurate description of the production of people with languages pathologies so as to find what is deficient in their speech

and then propose relevant treatments or relevant measures to help them overcome their difficulties (Grasemann *et al.*, 2013; Kiran *et al.*, 2013). Additionally cognitive science has of course a long tradition of mapping language deficiencies with specific areas in the brain.

One of the main challenges for studies in language pathologies is to access large and representative corpora that are still lacking in the field (Ferguson *et al.*, 2008). Most studies are based on small size data that may not be representative enough. Gathering large corpora is possible but pause multiple problems, from finding enough people with a similar pathology to ethical issues that are especially high here.

### 2.4. *Language evolution*

At first sight, language evolution can be seen more as a social process than as a cognitive one. However, language evolution has to take into account how a group of individuals master a language and transmit this knowledge to their infants. This is the core of social cognition, that aims at studying how individual knowledge interacts so as to give birth to social processes. Language evolution can thus be seen as one of the central topics of social cognition (Seyfarth and Cheney, 2014; Fitch *et al.*, 2010). It is our position here.

Research on language evolution has been booming in the late 1990s and early 2000s giving birth to a series of important conference called Evolang (held every other year since 1996). Computational models play a major role in this field since they make it possible to observe the influence of various parameters in the evolution of languages. Christiansen and Kirby (Christiansen and Kirby, 2003) propose to categorize research in this field in three broad categories:

1) *evaluation*: computational models require to make explicit all the variables and parameters used and can thus be seen as a "rigorous check" of model hypotheses and help identify hidden variables;

2) *exploration*: simulations can be used to see how a population of agents evolve from a particular initial situation, materialized in the computational model through a particular settings of its parameters;

3) *exemplification*: computational models can be used as a tool to demonstrate "how an explanation works", especially for pedagogical purposes.

In fact, these different categories are not exclusive, and nearly all simulations serve at the same time as a proof, as an illustration, and provide new ideas for further experiments. Most simulations are based on multi-agent systems, where large populations of agents give birth to new generations of agents at regular intervals. They transmit their language to new generations according to various parameters as said above. The evolution of different linguistic features (addressing phonology, morphology, syntax or even semantics) can thus be studied in this framework, see for example (Briscoe, 2009; Oudeyer, 2013; Kirby *et al.*, 2014) among many other references.

It should finally be noted that most simulations with multi-agent systems offer a quite abstract representation of the evolution of real languages (these simulations can even be completely abstract and related to populations of robots for example (Steels, 2005)). This is probably one of the main reasons of the apparent recent decrease of interest towards this field of research: computational models proposed so far were quite abstract and are hard to prove or justify against real world data. One of the challenge is thus now to better connect these models with real world data, which is a long term and highly challenging task (see (Van Trijp, 2012) for a recent example).

## 3. Content of this special issue

This special issue is dedicated to get a better picture of the relationships between computational linguistics and cognitive science. It specifically raises two questions: "what is the potential contribution of computational language modeling to cognitive science?" and conversely: "what is the influence of cognitive science in contemporary computational linguistics?"

The call for papers specifically targeted contributions on actual applications of methods from computational linguistics aiming at modeling cognitively motivated phenomena, as well as applications of cognitive theories to the computational modeling of language.

The following topics were proposed:

– Computational models of natural language acquisition, word clustering and word segmentation

– Psycholinguistically motivated phonetic, phonological, morphological syntactic, semantic, pragmatic studies of language

– Statistical and probabilistic modeling of factors encouraging one production or interpretation over its competitors

– Models of language emergence, change and evolution

– Models of language processing and surprisal

– Experimental or corpus driven modeling and analysis of language

We have received seven submissions, and four of them have been selected for publication (two of them written in English and two in French). We think they give a good overview of some recent research trends: the four articles cover very different areas of the field, although they of course do not represent the whole field.

The first paper, by Maxime Amblard, Karën Fort, Caroline Demily, Nicolas Franck and Michel Musiol is entitled "Analyse lexicale outillée de la parole transcrite de patients schizophrènes" ("Machine-assisted lexical analysis of speech transcripts from schizophrenic patients"). The paper reports some recent results obtained from the analysis of transcriptions of audio recordings of people suffering from schizophrenia. The corpus at stake contains 375.000 tokens, which is considerably larger than previ-

ous similar corpora and makes it necessary to use natural language processing tools for the analysis. The paper mainly investigates disfluencies and the use of lexical forms. The main conclusion is that although people with schizophrenia produce more disfluencies, their discourse does not seem specific concerning the use of lexical forms, contrary to a commonly held hypothesis. This paper illustrates how natural language processing now plays a major role in understanding language pathologies.

The second paper, by Quentin Feltgen, Benjamin Fagard and Jean-Pierre Nadal is entitled "Représentation et évolution du langage dans les modèles numériques : la grammaticalisation comme perspective" ("Language representation and models of language evolution: A grammaticalization perspective"). The authors investigate the notion of "grammaticalization", a well known linguistic phenomenon where some words or expressions become gradually frozen so as to play a grammatical role (the most well known example in French is the noun "*pas*" that has completely lost its original meaning when the word is used as a negation in contemporary French). Even if a large number of recent projects in computational linguistics have explored language evolution, phenomena such as grammaticalization seem to have been put aside in the computational community, despite the high success of this notion among historical linguists. Within this context, the authors propose a new model illustrating, among other things, the loss of semantic content, known as "semantic bleaching" (Givón, 1979) (often translated as "javellisation sémantique" in French).

The third paper, by Ákos Kádár, Afra Alishahi and Grzegorz Chrupała is entitled "Learning word meanings from images of natural scenes". The paper investigates how to associate word meanings with objects (through images) in noisy environments. To address this question, the authors use a large collection of images with natural language descriptions. The idea is to detect regularities and gradually create associations between words and image features. The authors show that their model correlates with human similarity judgments of word pairs when taking into account ambiguity and referential uncertainty. A parallel can be drawn with learning words in real life environments where children need to identify which aspects of the scenes are related to which parts of the perceived utterances.

The last paper, from Bruno Gaume, Karine Duvignau, Emmanuel Navarro, Yann Desalle, Hintat Cheung, Shu-Kai Hsieh, Pierre Magistry and Laurent Prévot is entitled "Skillex: a graph-based lexical score for measuring the semantic efficiency of used verbs by human subjects describing actions". The authors are interested in the conceptual organization of lexical networks. They describe a technique to derive a semantic network from a dictionary and show that, despite surface disagreements, networks derived from different dictionaries all share the same topological structure. They assume that this structure is meaningful from a cognitive point of view and they show how this assumption can be used to evaluate action labelling tasks. They present a technique for the evaluation of what they call "naming efficiency" and detail a comparison involving children and adults.

**Acknowledgements**

We thank the journal editors Éric de la Clergerie, Yves Lepage, Jean-Luc Minel and Pascale Sébillot, as well as the following reviewers, for their efforts in making this special issue possible:

– F. Alario (LPC, Univ. Aix-Marseille, France),
– A. Alishahi (Tilburg University, The Netherlands),
– M. Amblard (LORIA, Univ. de Lorraine, France),
– P. Blache (LPL, Aix-en-Provence, France),
– A. Christophe (LSCP, Paris, France),
– S. Colonna (SFL, Univ. Paris-VIII, France),
– I. Dautriche (LSCP, Paris, France),
– E. Dupoux (LSCP, Paris, France),
– T. Francois (Cental, UCL, Belgium),
– B. Gaume (CLLE-ERSS, Toulouse, France),
– J. Ginzburg (CLILLAC, Univ. Paris-VII, France),
– M. Grant (McGill University, Canada),
– J. Hale (Cornell University, USA),
– B. Hemforth (LLF, Univ. Paris-VII, France),
– P. Logacev (Potsdam University, Germany),
– F. Pellegrino (DDL, Lyon, France),
– D. Reitter (Penn State University, USA),
– E. Shutova (U. Berkeley, USA),
– J. Thuilier (CLLE-ERSS, Univ. Toulouse Jean-Jaurès, France),
– A. Villavicencio (UFRGS, Brazil),
– T. von der Malsburg (UC San Diego, USA),
– M. Zock (LIF, Marseille, France).

## 4. References

Alishahi A., Stevenson S., "A computational model of early argument structure acquisition", *Cognitive Science*, vol. 32, n° 5, p. 789-834, 2008.

Bach E., Brown C., Marslen-Wilson W., "Crossed and nested dependencies in German and Dutch: A psycholinguistic study", *Language and Cognitive Processes*, vol. 1(4), p. 249-262, 1986.

Blache P., "Chunks et activation : un modèle de facilitation du traitement linguistique", *Actes de la conférence Traitement Automatique du Langage Naturel (TALN)*, Les Sables-d'Olonne, 2013.

Boston M. F., Hale J. T., Patil U., Kliegl R., Vasishth S., "Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus", *Journal of Eye Movement Research*, vol. 2, n° 1, p. 1-12, 2008.

Boston M. F., Hale J. T., Vasishth S., Kliegl R., "Parallel processing and sentence comprehension difficulty", *Language and Cognitive Processes*, vol. 26, n° 3, p. 301-349, 2011.

Briscoe T., "What can formal or computational models tell us about how (much) language shaped the brain?", *Biological Foundations and Origin of Syntax*, MIT Press, Cambridge, USA, 2009.

Buttery P., Korhonen A., "I will shoot your shopping down and you can shoot all my tins – Automatic Lexical Acquisition from the CHILDES Database", *Proceedings of ACL 2007 Workshop on Cognitive Aspects of Computational Language Acquisition*, Association for Computational Linguistics, Prague, 2007.

Chomsky N., *Knowledge of Language: Its Nature, Origin, and Use*, Praeger Publishers, 1986.

Christiansen M. H., Kirby S., "Language evolution: Consensus and controversies", *Trends in Cognitive Sciences*, vol. 7, n° 7, p. 300-307, 2003.

Church K., "A pendulum swung too far", *Linguistic Issues in Language Technology*, vol. 7, p. n/a, 2011.

Dehaene S., *Les Neurones de la lecture*, Odile Jacob, New York, 2007.

Demberg V., Keller F., "Data from eye-tracking corpora as evidence for theories of syntactic processing complexity", *Cognition*, vol. 109, n° 2, p. 193-210, 2008.

Dillon B., Mishler A., Sloggett S., Phillips C., "A computational cognitive model of syntactic priming", *Journal of Memory and Language*, vol. 69, n° 4, p. 85-103, 2011.

Dreyfus H. (ed.), *Alchemy and AI*, RAND Corporation, Santa Monica, 1965.

Dreyfus H. (ed.), *What Computers Can't Do*, MIT Press, Cambridge, 1972.

Ferguson A., Craig H., Spencer E., "Exploring the potential for corpus-based research in speech-language pathology", *Proceedings of the 2008 HCSNet Workshop on Designing the Australian National Corpus*, Somerville, 2008.

Fitch T. W., Huber L., Bugnyar T., "Social cognition and the evolution of language: constructing cognitive phylogenies", *Neuron*, vol. 65, p. 795-814, 2010.

Frank S. L., "Uncertainty reduction as a measure of cognitive load in sentence comprehension", *Topics in Cognitive Science*, vol. 5, n° 3, p. 475-494, 2013.

Frank S. L., Trompenaars T., Vasishth S., "Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?", *Cognitive Science*, 2015.

Gibson E., "Dependency locality theory: A distance-based theory of linguistic complexity", *in* A. Marantz, Y. Miyashita, W. O'Neil (eds), *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*, MIT Press, Cambridge, MA, 2000.

Givón T., *On Understanding Grammar*, Academic Press, New York, 1979.

Grasemann U., Kiran S., Sandberg C., Miikkulainen R., "Impairment and Rehabilitation in Bilingual Aphasia: A SOM-Based Model", *in* Laaksonen, J. and Honkela, T. (ed.), *Proceedings of WSOM11, 8th Workshop on Self-Organizing Maps*, vol. 16, Springer Verlag (LNCS 6731), p. 207-217, 2013.

Haarmann H. J., Just M. A., Carpenter P. A., "Aphasic sentence comprehension as a resource deficit: A computational approach", *Brain and Language*, vol. 59, n° 1, p. 76-120, 1997.

Hale J., "The information conveyed by words in sentences", *Journal of Psycholinguistic Research*, vol. 32, nº 2, p. 101-123, March, 2003.

Hale J. T., "A probabilistic Earley parser as a psycholinguistic model", *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, PA, 2001.

Hale J. T., "Uncertainty about the rest of the sentence", *Cognitive Science*, 2006.

Husain S., Vasishth S., Srinivasan N., "Strong expectations cancel locality effects: Evidence from Hindi", *PLoS ONE*, vol. 9, nº 7, p. 1-14, 2014.

Husain S., Vasishth S., Srinivasan N., "Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus", *Journal of Eye Movement Research*, vol. 8(2), p. 1-12, 2015.

Hutchins J. (ed.), *Machine Translation: Past, Present, Future*, Ellis Horwood (*Ellis Horwood Series in Computers and their Applications*), Chichester, 1986.

Hutchins J. (ed.), *Early Years in Machine Translation: Memoirs and Biographies of Pioneers*, John Benjamins, Amserdam, 2000.

Jäger L., Chen Z., Li Q., Lin C.-J. C., Vasishth S., "The subject-relative advantage in Chinese: Evidence for expectation-based processing", *Journal of Memory and Language*, vol. 79-80, p. 97-120, 2015.

Joshi A. K., "Processing crossed and nested dependencies: An automaton perspective on the psycholinguistic results", *Language and Cognitive Processes*, vol. 5(1), p. 1-27, 1990.

Joshi A. K., Becker T., Rambow O., "Complexity of scrambling: A new twist to the competence-performance distinction", *in* A. Abeillé, O. Rambow (eds), *Tree Adjoining Grammars*, CSLI, p. 167-181, 2000.

Just M., Carpenter P., "A theory of reading: From eye fixations to comprehension", *Psychological Review*, vol. 87, nº 4, p. 329-354, 1980.

Kiran S., Grasemann U., Sandberg C., Miikkulainen R., "A computational account of bilingual aphasia rehabilitation", *Bilingualism: Language and Cognition*, vol. 16, p. 325-342, 2013.

Kirby S., Griffiths T., Smith K., "Iterated learning and the evolution of language", *Current Opinion in Neurobiology*, vol. 28C, p. 108-114, 2014.

Konieczny L., "Locality and parsing complexity", *Journal of Psycholinguistic Research*, vol. 29(6), p. 627-645, 2000.

Kutas M., Hillyard S., "Brain potentials during reading reflect word expectancy and semantic association", *Nature*, vol. 307, p. 161-163, 1984.

Legate J., Yang C., "Empirical re-assessment of stimulus poverty arguments", *Linguistic Review*, vol. 19, nº 1/2, p. 151-162, 2002.

Levy R., "Expectation-based syntactic comprehension", *Cognition*, vol. 106, p. 1126-1177, 2008.

Levy R., Fedorenko E., Gibson E., "The syntactic complexity of Russian relative clauses", *Journal of Memory and Language*, vol. 69, nº 4, p. 461-495, 2013.

Levy R. P., Keller F., "Expectation and locality effects in German verb-final structures", *Journal of Memory and Language*, vol. 68, nº 2, p. 199-222, 2013.

MacWhinney B., "A multiple process solution to the logical problem of language acquisition", *Journal of Child Language*, vol. 31, p. 883-914, 2004.

MacWhinney B. (ed.), *The CHILDES Project: Tools for Analyzing Talk*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1995.

Morgenstern A., Parisse C., "The Paris Corpus", *Journal of French Language Studies*, vol. 22, p. 7-12, 2012.

Nicenboim B., Vasishth S., Kliegl R., Gattei C., Sigman M., "Working memory differences in long distance dependency resolution", *Frontiers in Psychology*, 2015.

Oudeyer P.-Y., *Aux sources de la parole: auto-organisation et évolution*, Odile Jacob, Paris, 2013.

Peters G., Epstein R. (eds), *The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, Kluwer, Dordrecht, 2004.

Poibeau T., Villavicencio A., Korhonen A., Alishahi A., "Computational modeling as a methodology for studying human language learning", *in* A. Villavicencio, T. Poibeau, A. Korhonen, A. Alishahi (eds), *Cognitive Aspects of Computational Language Acquisition*, Springer Verlag ("Theory and Applications of Natural Language Processing"), p. 1-26, 2013.

Pullum G., Scholz B., "Empirical assessment of stimulus poverty arguments", *Linguistic Review*, vol. 19, n° 1/2, p. 9-50, 2002.

Rambow O., Joshi A. K., "A processing model for free word order languages", *in* C. C. Jr., L. Frazier, K. Rayner (eds), *Perspectives on Sentence Processing*, Lawrence Erlbaum Publisher, Hillsdale, NJ, 1994.

Roy D., "New horizons in the study of child language acquisition", *Proceedings of Interspeech 2009*, Brighton, England, 2009.

Rumelhart D. E., McClelland J. (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, 1986.

Safavi M. S., Vasishth S., Husain S., "Locality and expectation in Persian separable complex predicates", *Proceedings of the CUNY Sentence Processing Conference*, University of Southern California, Los Angeles, CA, 2015.

Sagae K., Davis E., Lavie A., MacWhinney B., Wintner S., "High-accuracy annotation and parsing of CHILDES transcripts", *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition*, Association for Computational Linguistics, Prague, Czech Republic, p. 25-32, 2007.

Sagae K., Davis E., Lavie A., MacWhinney B., Wintner S., "Morphosyntactic annotation of CHILDES transcripts", *Journal of Child Language*, vol. 37, n° 03, p. 705-729, 2010.

Seyfarth R. M., Cheney D. L., "The evolution of language from social cognition", *Current Opinion in Neurobiology*, vol. 28, p. 5-9, 2014.

Socher R., Manning C., Bengio Y., "Tutorial: Deep learning for NLP (without magic)", *North American Association for Computational Linguistics*, vol. 236, p. 433-460, 2013.

Staub A., "Eye movements and processing difficulty in object relative clauses", *Cognition*, vol. 116, n° 1, p. 71-86, 2010.

Steels L., "What triggers the emergence of grammar?", *Proceedings of AISB'05: the Second International Symposium on the Emergence and Evolution of Linguistic Communication (EELC'05)*, Hertfordshire, 2005.

Tomasello M., *Constructing a language: a usage-based theory of language acquisition*, Harvard University Press, 2003.

Turing A., "Computing machinery and intelligence", *Mind*, vol. 236, p. 433-460, 1950.

Van Trijp R., “Not as awful as it seems: explaining German case through computational experiments in fluid construction grammar”, *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Avignon, 2012.

Vasishth S., Lewis R. L., “Argument-head distance and processing complexity: Explaining both locality and antilocality effects”, *Language*, vol. 82, n° 4, p. 767-794, 2006.

Villavicencio A., Poibeau T., Korhonen A., Alishahi A. (eds), *Cognitive Aspects of Computational Language Acquisition*, Springer Verlag (“Theory and Applications of Natural Language Processing”), Berlin, 2013.

Wagers M., Lau E., Phillips C., “Agreement attraction in comprehension: Representations and processes”, *Journal of Memory and Language*, vol. 61, n° 2, p. 206-237, 2009.