

# Towards a computational approach to literary text analysis

Antonio Roque

Computer Science Department  
University of California, Los Angeles  
antonio@roque-brown.net

## Abstract

We consider several types of literary-theoretic approaches to literary text analysis; we describe several concepts from Computational Linguistics and Artificial Intelligence that could be used to model and support them.

## 1 Problem Statement

Consider the first sentence of the novel *Finnegan's Wake* (Joyce, 1939):

riverrun, past Eve and Adam's, from swerve of shore to bend of bay, brings us by a commodus vicus of recirculation back to Howth Castle and Environs.

To computationally analyze this sentence as literature, we must understand that its meaning is more than the combination of its semantic components. The rubric of "who did what to whom, when, where, and why" will at best lead us only to understand that somewhere, probably in Ireland, a river is flowing.

Some obvious low-level tasks to improve our reading include: exploring the meaning of non-standard capitalization and spacing, as in "riverrun"; resolving allusions, such as "Eve and Adam's," and considering the significance of variations from common phrasings<sup>1</sup>; identifying alliterated phrases such as "swerve of shore" and "bend of bay" and considering their effect; recognizing tone shifts such as "commodus vicus of recircula-

tion," and resolving any allusions they may indicate; identifying the significance of named entities such as "Howth Castle and Environs"<sup>2</sup>; exploring the effect of the line's syntax on reception, as described by writing scholars (Tufte, 2006).

But becoming absorbed in these admittedly interesting questions threatens to distract us from the larger questions that literary theorists have been studying for over a century. Those questions include:

- what interpretation is the "gold standard" by which others should be judged? Is it the meaning intended by the author? Is it the significance of the text to the readers (and if so, which readers?) Or is the meaning of a literary text inherent in how it takes part in a system and process of language use?
- what metrics can tell us whether one interpretation is better than another?
- how should we model the literary text as it relates to the larger body of language use, which includes both literary and nonliterary texts as well as everyday situated language use by intelligent agents? What features are necessary and sufficient to represent the way meaning (both literary and non-literary) is created and established among language-using populations? How is this meaning tied both to an intelligent

<sup>1</sup> For example, the quotation-delimited phrase "Adam and Eve" returns over 12 million Google results but "Eve and Adam" only returns around 200,000 (as of March 28, 2012.)

<sup>2</sup> For example: do they have an appearance or other attribute that would commonly be brought to mind? Are there associations that would normally be suggested to members of a given community of language use? *cf.* the significance of the Watergate office complex in American communities of political discourse.

agent's abstract beliefs as well as that agent's moment-to-moment understanding of its environment?

The wording of these questions is slanted to suggest their utility to computational linguistics. First, we may want to know how much of the meaning of a literary text comes from the author as opposed to from our situated interpretation of the text or from a language system<sup>3</sup>. Second, evaluation metrics would help us determine whether or not the performance of an automated literary system is improving. Finally, we would benefit from the explanations of a computational model of a literary text's meaning as it emerges from the situated reading of an authored artifact in the context of a multi-agent language system; if nothing else, it would tell us how to design algorithms that both consume and produce literary artifacts in human-like ways.

## 2 Approach

Computationally, the questions in Section 1 are likely to be answered over the course of decades rather than years. Contemporary relevant research from the fields of Computational Linguistics (CL) and Artificial Intelligence (AI) includes: semantic analysis of narratives (Elson and McKeown, 2009, Finlayson, 2011); summarizing fiction (Mani, 2005; Kazantseva and Szpakowicz, 2010) and performing information-extraction on fiction (Elson et al., 2010); modeling affect and reader-response in narrative (Mani, 2010; McQuiggan, 2010; Mohammad, 2011; Francisco et al., 2011); properties of narrative such as novelty (Peinado et al., 2010) and irony (Utsumi, 2004); models of discourse in narrative (Polanyi et al., 2004; Goyal et al., 2010); computational models of aesthetic creativity (Gervás et al., 2009); and the automatic generation of prose (Callaway and Lester, 2002) and poetry (Manurung, 2003; Gervás, 2007; Greene et al., 2010).

However, these disparate research traditions consider questions closer to the low-level tasks described in Section 1 than to the theoretical questions of interpretation ranking, evaluation, and computational modeling of meaningful human lan-

guage use. This is possibly because of the empirical methods which have become dominant in AI/CL in recent history (Cohen, 1995). A field whose methods are tuned to empirical evaluation will naturally shy from an area with few clear empirical tasks, whose humanities practitioners are likely to indulge in analyses assuming human levels of knowledge and language-processing capabilities.

Because of this we will turn instead for inspiration from the **digital humanities** (Schreibman, 2004). With its roots in humanities computing (Hockey, 2004) which constituted the earliest use of computers in the humanities, digital humanities took shape with the advent of the Internet. Digital humanities researchers currently apply computers to research questions such as authorship attribution (Jockers and Witten, 2010), statistical word-use analysis (Burrows, 2004), and the development of resources for classical lexicography (Bamman and Crane, 2009), often collaborating with statisticians or computer scientists.

Digital humanities has always had detractors among more traditional humanities scholars, but scholars sympathetic to the overall goals of digital humanities have also critiqued some of its practices. Consider the technological constraints imposed by projects in which texts are digitized, annotated, and statistically analyzed. Those constraints make tacit assumptions about the objectivity of knowledge and the transparency of its transmission (Drucker, 2009). Those assumptions may be contrary to a literary theorist's understanding of how literary text analysis actually works.

For example, in the case of scholar/artist Johanna Drucker, knowledge is seen as partial, subjective, and situated. Subjectivity in this context has two components: a point of view inscribed in the possible interpretations of a work, and "inflection, the marked presence of affect and specificity, registered as the trace of difference, that inheres in material expressions" (Drucker, 2009). To Drucker, subjectivity of knowledge is evident in the fact that interpretation occurs in modeling, encoding, processing, and accessing knowledge.

Drucker's focus is on humanities tools in digital contexts rather than digital tools in humanities contexts. We will proceed in a similar spirit, considering the tasks and approaches of literary text analysis as practiced by literary theorists and considering what kinds of models and approaches from contemporary AI/CL research they might find useful,

<sup>3</sup> We may be interested in user modeling of the author, versus modeling our own interpretative techniques, versus performing sentiment analysis on a particular community of language use, for example.

rather than starting with the tasks and approaches that AI/CL researchers are most familiar with and asking how they can be applied to literary text analysis.

As a specific goal to guide our thought, we will adopt a statement from another scholar who emphasizes the importance of keeping the humanities central to computational text analysis. In *Reading Machines: Toward an Algorithmic Criticism*, Stephen Ramsay develops the notion of adapting the constraints imposed by computation to intentionally create "those heightened subjectivities necessary for critical work" (Ramsay, 2011). While doing so, Ramsay states that from a humanist's perspective:

Tools that can adjudicate the hermeneutical parameters of human reading experiences - tools that can tell you whether an interpretation is permissible - stretch considerably beyond the most ambitious fantasies of artificial intelligence.

The rest of this paper will attempt to respond to Ramsay's claim by developing such ambitious fantasies. We will strive to consider literary text analysis as it is understood by literary theorists of recent history, and we will describe how representative processes from each of these theories could be modeled computationally using techniques from the AI/CL research communities.

### 3 Literary Text Analysis

#### 3.1 Expressive Realism

Human judgments on the nature of literature and the way literature is best read have changed frequently since classical times. The last century in particular has provided numerous, often contradictory, notions of how we should determine the meaning of a story, leaving us with no consensus. Even within a school of thought there may be significant differences of opinion, and evaluation is typically no more empirical than how persuasive the interpretation of a given text may be. Still, we may identify certain key ideas and use them to imagine ways they could involve computation.

We may begin by considering **expressive realism**, an approach to literary theory which developed in the late 19th and early 20th centuries, and is a combination of the classical Aristotelian

notions of art as *mimesis* (reproducing reality) and the Romantic-era view of poetry as an outpouring of strong emotions produced by an artist whose percepts and affective processing are unusually well-tuned<sup>4</sup> (Belsey, 1980). The task of the reader in this formulation is to faithfully create in their minds the realities being represented by the work, and to enrich themselves by following the thoughts and feelings that the artist experienced.

Computationally, we may frame this as a knowledge engineering task: the writer is a subject matter expert in perceiving the world, and has developed knowledge about the world and innovative ways of emotionally relating to the world. The literary critic's task is to identify which writers have produced knowledge and affective relationships that are most worth adopting. The reader's task is to be guided by the critics to the best writers, and then strive to adopt those writers' knowledge and affective relations as their own.

It may seem difficult to perform such a task with a text such as *Finnegan's Wake*, which is not easy to translate into propositions. But consider a writer's understanding of what happens when reading expressive realist fiction (Gardner, 1991):

If we carefully inspect our experience as we read, we discover that the importance of physical detail is that it creates for us a kind of dream, a rich and vivid play in the mind. We read a few words at the beginning of a book or the particular story, and suddenly we find ourselves seeing not only words on a page but a train moving through Russia, an old Italian crying, or a farmhouse battered by rain.

Gardner describes fiction as producing an immersive experience in which the reader's sensations are empathically aligned with those of the writer. This alignment produces an understanding unlike that of propositional knowledge:

[The writer] at the very least should be sure he understands the common objection summed up in the old saw "Show, don't tell." The reason, of course, is that set beside the complex thought achieved by drama, explanation is thin gruel,

<sup>4</sup> Belsey, who is critical of this approach, quotes the poet William Wordsworth's view of artists as "possessed of more than usual organic sensibility." In fact, Wordsworth believed a Poet was "endowed with more lively sensibility; more enthusiasm and tenderness, who has a greater knowledge of human nature, and a more comprehensive soul, than are supposed to be common among mankind..." (Wordsworth, 1802.)

hence boring. ... After our [reading] experience, which can be intense if the writer is a good one, we know why the character leaves when finally she walks out the door. We know in a way almost too subtle for words...

The subtly described by Gardner's explains how a text such as *Finnegan's Wake* may be read without recourse to a detailed exegesis producing propositional content. The reader need only become suggestible to the text, and allow themselves to experience the "complex thought" suggested by the writer. Of course, this "intense" experience may lead one to a further study of the writer's mind-set, which would then create an even fuller understanding of that writer's approach.

Such a description may seem like an unlikely candidate for computational modeling, but consider the neurolinguistic implications of models of the mirror neuron system (Rizzolatti and Craighero, 2004): hypothetically, a reader's neural structure might literally copy that of the writer's, provided the stimulus of the text. In this way we might model the transmission of knowledge "almost too subtle for words."

### 3.2 New Criticism

Later literary theories found expressive realism problematic in various ways. For example, the Anglo-American **New Criticism** defined the *intentional fallacy*, which states that "the design or intention of the author is neither available nor desirable as a standard for judging the success of a work of literary art" (Wimsatt and Beardsley, 1954)<sup>5</sup>. Wimsatt and Beardsley proposed to avoid "author psychology" by focusing on the *internal evidence* of the text, which they defined as

public evidence which is discovered through the semantics and syntax of a poem, through our habitual knowledge of the language, through grammars, dictionaries, and all the literature which is the source of dictionaries, in general through all that makes a language and culture...

The language knowledge and resources were used to identify the "technique of art". New Critic

<sup>5</sup> Note that Wimsatt and Beardsley did not deny the scholarly value of "literary biography," and New Critic John Crowe Ransom stated "Without [historical studies] what could we make of Chaucer, for instance?" (Ransom, 1938) New Critics merely believed that *close readings* of the text should take precedence during literary text analysis.

John Crowe Ransom provided examples of what *devices* should be used in analyzing poetry (Ransom, 1938):

its metric; its inversions; solecisms, lapses from the prose norm of language, and from close prose logic; its tropes; its fictions, or inventions, by which it secures 'aesthetic distance' and removes itself from history...

However, these devices were not studied for their own sake. Ransom continued: "the superior critic is not content with the compilation of the separate devices; he suggest to him a much more general question." The question in this case is "what [the poem] is trying to represent" and why it does so using those particular devices. This was worth understanding because the New Critics believed that "great works of literature are vessels in which humane values survive" (Selden and Widowson, 1993) and which reinforce those values in the diligent reader.

Computationally, the list of language resources described by Wimsatt and Beardsley recalls the corpus- and knowledge-based resources promoted by textbook approaches to CL (Jurafsky and Martin, 2000). The low-level tasks in analyzing *Finnegan's Wake* described in Section 1 align with the New Critical identification of literary devices. Much of the CL/AI research described in Section 2 is in this vein.

However, to create a complete computational model of literary reading from this perspective we would also need a model of the types of "humane values" that New Critics revered. Unfortunately, the New Critics themselves did not explicitly provide such a model, as doing so was considered irrelevant. But we ourselves could adapt a computational model of culture to develop a representation of the New Critic's cultural values. AI researchers develop computational model of culture by, for example, implementing Cultural Schema Theory and Appraisal Theory in cognitive architectures to describe how culture emerges from an individual's cognitive processes (Taylor et al., 2007). There is room here to adapt the system of perceived affordances (Gorniak and Roy, 2006) in which language understanding is represented as the process of filtering real-world devices in a way analogous to how the New Critics filter literary devices.



### 3.3 Russian Formalism

The New Criticism developed independently of **Russian formalism**, which similarly focused on the text and the literary devices present, rather than the author's intentions or the context of the text's production. Because of this, most of the computational representations used in discussion of the New Critics could also be applied to the Russian formalists.

However, unlike the New Critics, the Russian formalists believed that art existed to create a sense of *defamiliarization*:

art exists that one may recover the sensation of life; it exists to make one feel things... The technique of art is to make objects 'unfamiliar,' to make forms difficult, to increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged. *Art is a way of experiencing the artfulness of an object: the object is not important.*<sup>6</sup>

The defamiliarizing force of literature is easy to see in a text such as *Finnegan's Wake*, whose second sentence reads:

Sir Tristram, violer d'amores, fr'over the short sea, had passencore rearrived from North Armoriga on this side the scraggy isthmus of Europe Minor to wielderfight his penisolate war: nor had topsawyer's rocks by the stream Oconee exaggerated themselfe to Laurens County's gorgios while they went doublin their mumper all the time: nor avoice from afire bellowsed mishe mishe to tauftauf thuartpeatrick: not yet, though venissoon after, had a kidscad buttended a bland old isaac: not yet, though all's fair in vanessy, were sosie sesthers wroth with twone nathand-joe.

This is not a text that can easily be read rapidly. A more methodical reading is most obviously rewarded by the portmanteaux (which are created by combining words in new ways) along with the other literary devices. Computationally, as before this can be seen as another set of devices to be automatically processed. However it may be more productive to see this as an example of how writers strive to invent new devices and combine devices in new ways, which may be resistant to automated

analyses. From this perspective, defamiliarization has its effect on the computational linguist who is developing the algorithms. The perception of the researcher is thus shifted and prolonged, creating in them a recovery of the sensation for language.

### 3.4 Structuralism and Post-Structuralism

Linguist Roman Jakobson was central figure in both Russian formalism and **structuralism**, two mutually influential schools of thought. A key difference between the two is their understanding of the relation between aesthetic products and their cultural context. To Russian formalists (as well as to New Critics), literary text existed apart from other cultural phenomena, whereas structuralism provided a formal framework which studied systems of arbitrary signs which could be built at different *levels*, (Schleifer, 1993) so that literary structures could be built with reference to cultural structures.

With roots in the semiotics of linguist Ferdinand de Saussure and of philosopher Charles Sanders Peirce, structuralism aimed at systematically uncovering the way that meaning arises from systems of signs forming linguistic elements such as sentences and paragraphs as well as higher levels of narrative discourse.

Continued scholarship on structuralism exposed a number of difficulties. Besides its lack of interest in individual cases or in the way systems change over time, the arbitrary nature of structuralist signs contradicted its aspirations to systematic representation (Schleifer, 1993). This was leveraged by philosopher Jacques Derrida, who argued that one could not study structures from "outside," in the way that an objective study requires.

Derrida was a **post-structuralist**, who used structuralism as a starting point but did not limit themselves with structuralism's constraints. Another post-structuralist, literary theorist Roland Barthes, used the phrase *death of the author* in a way reminiscent of the New Critics' intentional fallacy. Barthes, however, used the the arbitrariness of signs to go beyond the New Critics and reject the existence of any "ultimate meaning" of a text. Barthes saw the source of understanding as the reader:

[A] text consists of multiple writings, issuing from several cultures and entering into dialogue with each other, into parody, into contestation;

<sup>6</sup> First published in 1917, this translation is from (Shlovsky, 1988). Emphasis from the original.

but there is one place where this multiplicity is collected, united, and this place is not the author, as we have hitherto said it was, but the reader... (Barthes, 1967)

To Barthes, readers are not important in terms of their personal history or their state of mind, but rather that they are the one who "holds gathered into a single field all the paths of which the text is constituted." (Barthes, 1967) In other words, the text's meaning is dependent on the structures of signs in which the reader exists. And because signs are arbitrary, the reading produced by any reader must likewise be arbitrary, at least in terms of any objective measure of quality.

Another post-structuralist, psychologist Jacques Lacan, maintained that humans entered systems of signs in which they found or were provided roles, such as child/parent or male/female (Selden and Widdowson, 1993). This process is directed by the unconscious, and the only way it is able to take on comprehensible meaning is in expression through a system of language signs.

These are but a few of the influential structuralist and post-structuralist scholars, but they suffice to consider applicable computational techniques.

We begin by considering the concept of language as a complex adaptive system (Beckner et al., 2009). This provides a model that brings together language, interpretation, and intelligent agents (Steels, 2007) in a way that allows experiments with both sets of software agents and language-using robots (Steels, 2006). As in the structuralist view, meaningful language use is dependent on complex systems involving signification.

But this complex system is made up of language-using agents, who must work together to determine norms as well as actually use language for real-world tasks and abstract reasoning. The model must work not only at the system level, but also at the individual level. CL/AI research in societal grounding (DeVault et al., 2006), dialogue grounding (Traum, 1994), semantic alignment (Pickering and Garrod, 2004), and relational agency (Bickmore and Picard, 2005) provide ways of representing how populations of agents use language meaningfully, and how pairs of human-like intelligent agents coordinate language in situated dialogues, while developing social relationships. As in the Lacanian subject, these agents are created or trained in terms of their difference or similarity

from the other agents, adopting and being defined by their roles in the structured societies of agents.

When considering *Finnegan's Wake*, an intelligent agent would bring with it an algorithm for identifying features in stories, as well as resources such as language model data and its model of the role it fits in its social structures. Reading the text, the agent identifies literary devices that it uses as affordances to react with its emotions and its social perceptions, as well as to weigh the semantics of the text. When reading the text, the agent's interpretation of the story will be based on its gendered identity and personal history. In this way, the literary analysis of the agent is highly dependent on its sense of identity, as well as the localized nature of its language resources.

## 4 Conclusions

We began by describing some of the larger questions that literary theorists have been working with for over a century. We described some ideas from the digital humanities, including an expressed skepticism in artificial intelligence's ability to model human-like readings of literary texts. In response to that skepticism, we have described several major approaches to literary text analysis, and for each we have suggested ways in which state-of-the-art CL/AI techniques could be applied to model or support their approach.

Of course this is by no means an exhaustive survey of either literary theoretical approaches or applicable CL/AI techniques. Rather, we are suggesting that a great number of possibilities remain unexplored between the two.

## References

- David Bamman and Gregory Crane. 2009. Computational Linguistics and Classical Lexicography, *Digital Humanities Quarterly*, Volume 3 Number 1.
- Roland Barthes. 1967. The Death of the Author. *Aspen*. No. 5-6.
- Clay Beckner, Nick C. Ellis, Richard Blythe, John Holland, Joan Bybee, Jinyun Ke, Morten H. Christensen, Diane Larsen-Freeman, William Croft, Tom Schoenemann. 2009. Language Is a Complex Adaptive System: Position Paper. *Language Learning*, 59:Suppl 1, December 2009, pp 1-26.
- Catherine Belsey. 1980. *Critical Practice*. Routledge. London, UK.

- Timothy Bickmore and Rosalind Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (ToCHI)*.
- John Burrows. 2004. Textual Analysis. In *A Companion to Digital Humanities*, ed. S. Schreibman, R. Siemens, and J. Unsworth, Oxford: Blackwell Publishing.
- Charles B. Callaway and James C. Lester. 2002. Narrative Prose Generation, Artificial Intelligence. Volume 139 Issue 2, Elsevier Science Publishers Ltd. Essex, UK
- Paul R. Cohen. 1995. *Empirical Methods for Artificial Intelligence*. Bradford Books. Cambridge, MA.
- David DeVault, Iris Oved, and Matthew Stone. 2006. Societal Grounding is Essential to Meaningful Language Use. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*
- Johanna Drucker. 2009. *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. University Of Chicago Press.
- David K. Elson, Nicholas Dames, Kathleen R. McKeown. 2010. Extracting Social Networks from Literary Fiction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, Uppsala, Sweden
- David K. Elson and Kathleen R. McKeown. 2009. Extending and Evaluating a Platform for Story Understanding. *Papers from the 2009 AAAI Spring Symposium: Intelligent Narrative Technologies II*. The AAAI Press, Menlo Park, California.
- Mark A. Finlayson. 2011. Corpus Annotation in Service of Intelligent Narrative Technologies, *Proceedings of the 4th Workshop on Intelligent Narrative Technologies*, Stanford, CA.
- Virginia Francisco, Raquel Hervás, Federico Peinado, and Pablo Gervás. 2011. EmoTales: creating a corpus of folk tales with emotional annotations. *Language Resources & Evaluation*.
- John Gardner. 1991. *The Art of Fiction: Notes on Craft for Young Writers*. Vintage, New York, NY.
- Pablo Gervás. 2009. Computational Approaches to Storytelling and Creativity. *AI Magazine*, Fall, p 49-62.
- Pablo Gervás, Raquel Hervás, Jason R Robinson. 2007. "Difficulties and Challenges in Automatic Poem Generation: Five Years of Research at UCM". in *Papers presented at e-poetry 2007*, Université Paris8.
- Peter Gorniak and Deb Roy. 2007. Situated Language Understanding as Filtering Perceived Affordances. *Cognitive Science*, Volume 31, Issue 2, pages 197-231.
- Amit Goyal, Ellen Riloff, Hal Daumé, III. 2010. Automatically producing plot unit representations for narrative text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*.
- Erica Greene, Tugba Bodrumlu, and Kevin Knight. 2010. Automatic Analysis of Rhythmic Poetry with Applications to Generation and Translation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 524-533.
- Susan Hockey. 2004. The History of Humanities Computing. In *A Companion to Digital Humanities*. Blackwell, Oxford, UK.
- Matthew L. Jockers and Daniela M. Witten, 2010 "A comparative study of machine learning methods for authorship attribution", *Literary and Linguistic Computing* 25(2):215-223
- James Joyce. 1939. *Finnegan's Wake*. Faber and Faber, London, UK.
- Daniel Jurafsky and James H. Martin. 2000. *Speech and Language Processing*. Pearson Prentice Hall. Upper Saddle River, New Jersey.
- Anna Kazantseva and Stan Szpakowicz. 2010. Summarizing Short Stories. In *Computational Linguistics*, 36(1), pp. 71-109.
- Scott W. McQuiggan, Jennifer L. Robison, and James C. Lester. 2010. Affective Transitions in Narrative-Centered Learning Environments. In *Educational Technology & Society*. 13 (1): 40-53.
- Inderjeet Mani. 2005. Narrative Summarization. *Journal Traitement automatique des langues (TAL)*: Special issue on Context: Automatic Text Summarization.
- Inderjeet Mani. 2010. Predicting Reader Response in Narrative. In *Proceedings of the Intelligent Narrative Technologies III Workshop*.
- Hisar Maruli Manurung. 2003. *An Evolutionary Algorithm Approach to Poetry Generation*. PhD thesis, University of Edinburgh. College of Science and Engineering. School of Informatics.
- Saif Mohammad. 2011. From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales, In *Proceedings of the ACL 2011 Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH)*, June, Portland, OR.
- Federico Peinado, Virginia Francisco, Raquel Hervás, Pablo Gervás. 2010. Assessing the Novelty of Computer-Generated Narratives Using Empirical Metrics. *Minds & Machines*, 20:565-588.
- Martin J. Pickering and Simon Garrod. 2004. Towards a mechanistic Psychology of dialogue. *Behavioral and Brain Sciences*, 27:169-22.
- Livia Polanyi, Chris Culy, Martin van den Berg, Gian Lorenzo Thione, David Ahn, 2004. Sentential Structure and Discourse Parsing. In *ACL2004 - Workshop on Discourse Annotation*.

- Stephen Ramsay. 2011. *Reading machines: Towards an Algorithmic Criticism*. University of Illinois Press, Urbana, IL
- John Crowe Ransom. 1938. Criticism, Inc. Anthologized in *The Norton Anthology of Theory & Criticism*. 2010. WW Norton & Company, New York, NY.
- Giacomo Rizzolatti and Laila Craighero. 2004. The Mirror Neuron System. In *Annual Review of Neuroscience*. 27:169–92.
- Ronald Schleifer. 1993. Structuralism. in *The Johns Hopkins Guide to Literary Theory and Criticism*. Michael Groden and Martin Kreiswirth, eds. The Johns Hopkins University Press. Baltimore, USA.
- Susan Schreibman, Ray Siemens, John Unsworth, eds. 2004. *A Companion to Digital Humanities*. Blackwell, Oxford, UK.
- Raman Selden and Peter Widdowson. 1993. *A Reader's Guide to Contemporary Literary Theory*. University Press of Kentucky. Lexington, KY.
- Luc Steels. 2006. How to do experiments in artificial language evolution and why. *Proceedings of the 6th International Conference on the Evolution of Language*. pp 323-332.
- Luc Steels. 2007. Language Originated in Social Brains. *Social Brain Matters: Stances of Neurobiology of Social Cognition*, pages 223-242, Editions Rodopi. Amsterdam NL.
- Glenn Taylor, Michael Quist, Steve Furtwangler, and Keith Knudsen. 2007. *Toward a Hybrid Cultural Cognitive Architecture*. CogSci Workshop on Culture and Cognition.
- David Traum. 1994. *A Computational Theory of Grounding in Natural Language Conversation*, TR 545 and Ph.D. Thesis, Computer Science Dept., U. Rochester, NY.
- Virginia Tufte. 2006. *Artful Sentences: Syntax as Style*. Graphics Press, Cheshire, CT.
- Akira Utsumi. 2004. Stylistic and Contextual Effects in Irony Processing. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*.
- W.K. Wimsatt, Jr., and Monroe C. Beardsley. 1954. The Intentional Fallacy. From *The Verbal Icon: Studies in the Meaning of Poetry*. University of Kentucky Press, Lexington, KY.
- William Wordsworth. 1802. Preface to *Lyrical Ballads*. Anthologized in *The Norton Anthology of Theory & Criticism*. 2010. WW Norton & Company, New York, NY.