

Revue des arts
et médiations humaines

Journal of arts
and Human Mediations

Journal *Hybrid*, no. 7

The Creative Web of Languages

For an anoptical literature

Jean-Pierre Balpe

Jean-Pierre Balpe, Professor Emeritus of the University of Paris 8, was the Head of the Hypermedia Department and of the Paragraphe research laboratory at the same until September 2005. A researcher and theorist of the relations between the computer and literature, he has written various scientific and technical publications on the subject. His last theoretical book is *Contextes de l'art numérique* (Hermès, 2000). He is also a poet (*Bleus*, 1984, Action poétique; *Le Silence*, 1989, Action poétique; *101 poèmes du poète aveugle*, 2000, Farrago) and a novelist (*La Toile*, 1999, Cylibris). He has also published numerous short stories in various French journals. Since 1985, he has created or participated in numerous exhibitions of art and new technologies, including his interactive scenario *Shangai-Paris*. He has also had various interactive and generative shows including, in 1997, *Trois mythologies et un poète aveugle* with the French musical institute IRCAM. Currently he has various interactive and generative shows in progress. One of his most recent generative works is the e-novel *Trajectoires*, which can be read at www.trajectoires.com. In 2000, he created a mail-novel which was sent daily to the recipients agreeing to participate in the project. In February-April 2002, he created *MeTapolis*, a multimodal installation-show for the Marco museum in Monterey (Mexico), conceived with the Italian composer Jacopo Baboni-Schilingi and the French artist Miguel Chevalier. His latest works include *...nographies*, an electronic oratorio using generated poetry, also with Baboni-Schilingi, produced at the Molière Theatre in Paris in November 2003, *Babel poesie* for the P0es1e exhibition in Berlin (January-April 2004), *Fictions d'Issy* an urban

multimodal novel in 2005, and a three-day hypermedia poetry performance in Switzerland (2006), *Interdit au moins de seize ans*, *The Herbarium* (2009), generative and interactive installation with Miguel Chevalier in, *Capture* (2010), exhibition, show around lyrics rock automatic generation; *Moments carolingiens* (2014), generative installation for the CCR de l'abbaye Royale de Saint-Riquier, and *Monstrueux monstres* and *L'Esprit humain* (2015), generatives installations for the CCR de l'Abbaye Royale de Saint-Riquier.

Translated by Saskia Brown

Abstract

Based on Jean-Pierre Balpe's text generation software and experiments, this paper addresses the difference between "langue" and "langage" in French, how this difference plays out in the automatic literary text generation, and how it is possible to have literary texts automatically generated in French.

Keywords

automatic generation, "language" and "langue", French, literature

Published: 15 june 2021

Full text (PDF file)

“in the #post_digital era, the human mind must confront the challenge of managing several heterogeneous creatures simultaneously, which inhabit and belong to the same body moving in different spaces.”

Khaldoun Zreik (message on Twitter)

We must start by mentioning a problem in the English version of this issue’s theme, the term “language.” When translated into French, or other languages such as Italian, Spanish and Portuguese, the term is always ambiguous. These Latin languages make a clear distinction between, in French, “la langue” and “le langage.” In English, however, as in other derivatives of Anglo-Saxon (German, Swedish, etc.), the term “language” does not seem to be charged with ambiguity. Yet, as a French speaker, I feel obliged either to address both senses of “language” or to choose only one of them, and therefore, in any case, I must always define what I am talking about. This article will refer now to one sense, now to the other.

To understand this better, imagine a situation where the French word *langue* is used (translated as “language” or “tongue”). For exemple, the expression “cet homme emploie une langue de cochon.” This cannot be translated as “This man speaks a pig’s tongue,” but “this man uses bad language”. Of course, the context makes a difference, but this example clarifies why the title “The creative Web of languages” is problematic for a French ear. We shall consider both aspects of “language” here, since they are fundamental to any artistic approach to machine-generated literary texts. But for that we require clear definitions. Especially since there is a further ambiguity when we are considering computer-generated objects, and therefore – at their heart – programming languages. Now programming languages are nothing like a language as tongue, a system of signs which enables written or spoken exchanges between the members of a human community. They are, rather, rigorously structured organisations of symbols and formal combinatory rules with a technical function: they are codes.

Code: a conventional system of symbols, whose structure is governed by formal rules.

Tongue [*langue*]: a system of written or spoken exchanges between the members of a specific human community.

Language [langage]: a system of non-verbal signs fulfilling the function of open communication. It is thus possible to talk of “cinematic, pictorial, etc., language.” We should note that what Saussure

calls “parole,”¹ the individual act of will and intellect, is one, but not all, of the possibilities of language [langage]. This act can mobilise any means of expression, as long as it remains in the realm of inter-human verbal exchange, and so does not include all of what is considered to be language.

So we have no choice really: since digital literary works, and particularly the automatic generation of literary texts, call on these three levels of language for specific needs, we must discuss them all. Machine-generated texts raise the question of how indeed we can conceive of a machine, a program, which uses only code to begin with, yet creates language, or more precisely, text, in a particular tongue [*langue*]. Further, we must ask in what ways the text produced by this sort of machine is something other than ordinary text, that is, how it uses the diverse possibilities of human language to make this text a literary one, and, moreover, a literary text that can exist not on paper but digitally.

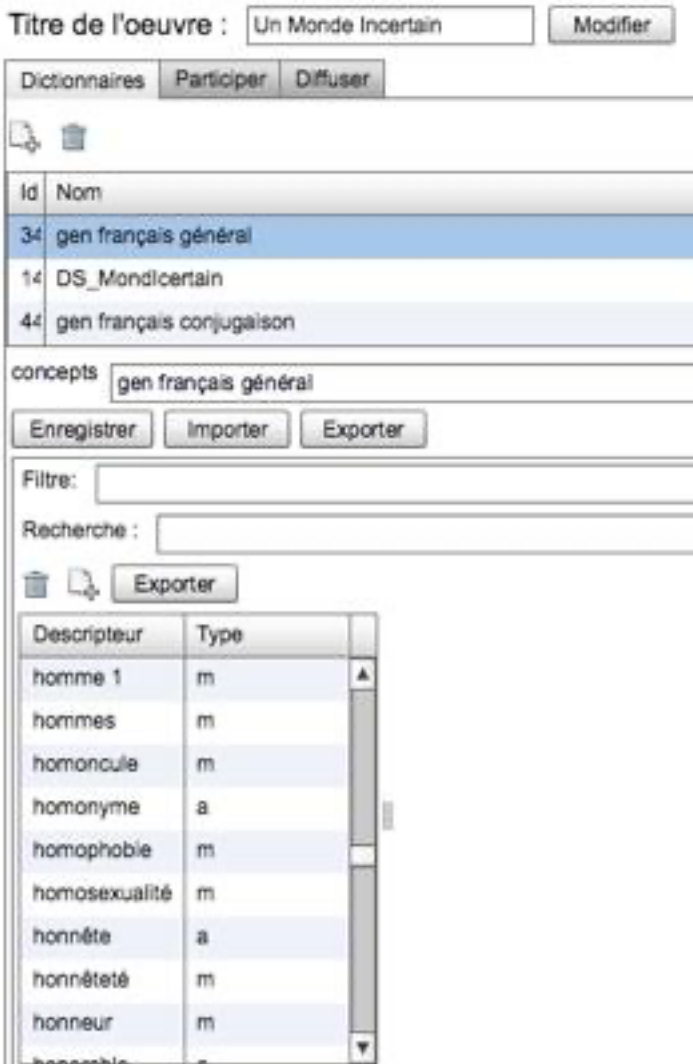
As everyone knows, the different layers of programming used in this process actually depend on one single code by which electronic machines function, the binary code, which in turn depends on two values, the “binary digits” 0 and 1. These values are structured into a group of 8 bytes, which build up to the level of the various programming languages. These, in turn, establish sets of instructions accepted by other programming languages, and so forth.

In this article, I shall simply present how I manage to get a computer to write a literary text from the deepest layers of programming I use. Unfortunately, this program is not open-source, so I cannot give its Internet address, but only describe how it functions.

First, to create an acceptable text in a given tongue, I must have the right dictionary, or at least a set of elements from a dictionary in the tongue I am using, in this case French. The English, German and Spanish dictionaries that I used occasionally, which function in the same way, proved to be much too small. Below, an example:

¹ Ferdinand de Saussure, *Cours de linguistique générale*, Paris, Payot, 1995.

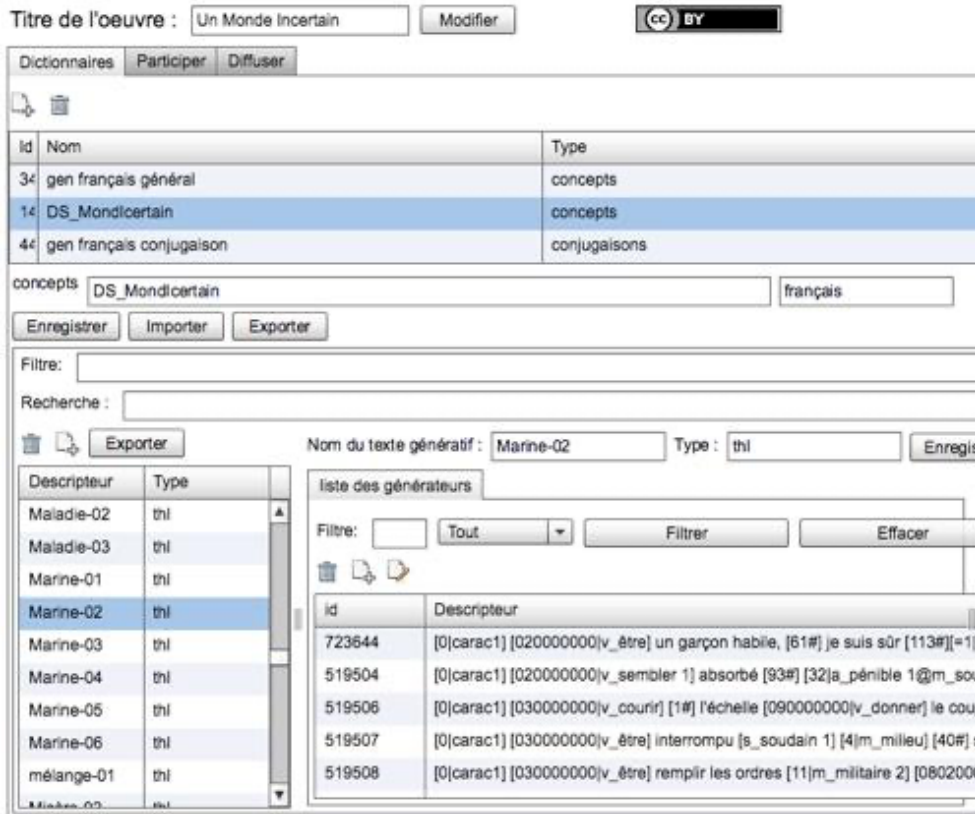
Fig. 1



Screen shot showing the architecture of dictionaries: each literature-generating program has two layers, 1. The layer of literary generation of a particular type of text, here, at the top, “DS_Mondincertain,” 2. The layer of a language [*langue*] described through a linguistic dictionary (here, French, “gen français general”), applicable to generating any text in French. A fragment appears in the bottom half of the screen shot. Under “Type,” ‘m’ means substantive, ‘a’ means adjective, etc.

One also needs to have a grammar of the language which corresponds to the dictionary definition of each word:

Fig. 2



Screen shot of the semantic dictionary of the text generator “Un monde incertain.” The lower, left-hand part of the screen shows the description of a set of semantic sub-sets, Marine-01, Marine-03, etc. The right-hand part displays some of the semantic structures of Marine-02, which play a role in the surface generation of a text. Each element in square brackets [] is a variable referring to entries in the general dictionary.

In brief – we cannot go into further detail here, for lack of space – all of these features make up a first-level metalanguage, which enables the generator to build grammatically correct (in this case, French) sentences. A second level of metalanguage takes us beyond the level of grammatically correct linguistic units, but with the risk of absurdity most of the time, if one leaves it to chance, for example: fear [sub: fear] [verb: to fear] [sub: fear] [adj: feared], which could give something like “fear fears feared fear” or “fears fear a feared fear” The problem is how to introduce a higher level of meaning from terms whose meaning is only embryonic. Meaning is defined basically by the relations words have with the real world, which are reflected in structures of relations. For example, in a particular tongue, “horse” and “neigh” are related, but no such link exists between “cow” and

“neigh” or “sparrow” and “neigh.” Yet despite many attempts at descriptions using semantic models (Schank, Meehan, Chomsky, Greimas,² etc.), none of them has worked, so ultimately it must be admitted that such relations exist as conventions within a particular language, and that they are at least as arbitrary as the relations between signifiers and signifieds in a language. The best way to turn this arbitrariness to our advantage is to focus on the set structures of a language. The automatic text generator behaves like a huge parasite which imports millions of data units and sentences from the Internet, to be processed almost automatically in order to obtain what can be defined as basic cells of meaning. We are somewhere between code and language [*langue*], in a layer of codes specific to the programmes that are able to use them.

Here is an example:

```
[0|carac2] [020000000|v_to have] [32|m_allié< all-powerful><, [103#] treacherous [55#]>
[122#] [020100000|v_to help]<, [8|m_night]>< [38#] [70|m_appetite 1] [10|m_flesh]
[=1|a_human]>
```

The surface realisation is about 2 000 possible sentences, of which three are given below:

Theresa had an all-powerful ally who helped her, every night, in her voracity for human flesh

Marie Anne had an all-powerful ally who helped her, every night, in her bingings on flesh

Hirimalda had an even more treacherous ally, who helped her in her greed for human flesh.

Etc.

Similarly, a text is defined by a set of basic cells described by one or several semantic markers

```
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]
```

² Roger C. Schank et Robert P. Abelson, *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*, Hillsdale, Lawrence Erlbaum Associates, 1977; James R. Meehan, “Tale-Spin, an interactive program that writes stories,” in *Proceedings of IJCAI*, 1977. [Online] <https://www.cs.utah.edu/nlp/papers/talespin-ijcai77.pdf> [accessed 8 April 2021] and Id., “The Metanovel: Writing Stories by Computer,” *Outstanding Dissertations in the Computer Sciences*, 1976; Noam Chomsky, *Aspects de la théorie syntaxique*, Paris, Seuil, 1971; A. J. Greimas, *Sémantique structurale. Recherche et méthode*, Paris, Larousse, 1966.

[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]

[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]

[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01][s_ponctuer]

[thl-FPalancy-01][s_ponctuer] [thl-FPalancy-01].

By recurrently applying the contextual data [thl-Palancy-01], the following text, among many others, can be generated:

The newspapers mention Gundrade more often than Chahan, and write about how she died at Diéval, of the plague.

– I saw the abyss beneath our feet, once again a war for this kingdom that no longer exists!

– if it was true that the divinities lie beyond the Seventh Gate, why did you not pass through it to destroy them.

– the night pulsates at the moment with the flight of the condors, Anral went on. Seeing its riches, the men of the earth, we had stuffed our pockets and walked along this beach; Theresa has never managed to invent anything. Did not want to be bothered with feelings... Anral despised writers. Before performing an action, Anral mostly counted to 16 – Anral told us how Françoise had found the young woman of the kingdom, our ancestor, and this is what he later wrote about this dead woman – at first nothing distinguished Anral and Anar. Anral was afraid to see the king, Dajan, and at the same time it was his hope – and that of the men with us in the bedroom.

– I will massacre anyone who crosses the threshold of the Sixth Gate, were he the tallest, the strongest, or the best warrior. The wild beasts attack stray dogs, animals in herds, and, since I became aware of it, two children of Aidan. Was more interested in the functioning of the rockets than in Orry la Ville.

– be that as it may, the insects were everywhere, whatever speed you travel at you will never be far enough away. Anral feared the hubbub and whistling of feelings like the plague... Finches chirped out their calls

– you will keep the key hanging on a chain which will be hung around your neck. Anral tried to find Ukam.

At this point, we have dealt with part of the code, and the generator is capable of writing in a programmed natural language, mainly French at present. That could be enough. However, if we do not want the generator just to produce miles of text, we need it to have the capacity to give the text shape, which is where language use [*langages*] intervenes. Poetry, for example, uses language [*langue*] in a very specific way, which we can call poetic language. This has nothing in common with the highly codified way in which plants are described for a herbarium, with formalisable rules. The same is true of a novel, or a lawyer's or policeman's report. In each of these areas particular rules of language are applied to the language [*langue*], an operation we used to think could be defined by the vague term "style." Many authors – Gertrude Stein, James Joyce, Georges Perec, etc. –

shape their national language into very particular forms. This has been occurring ever since a literature I would call “erudite” (to distinguish it from 99% of literary production) began. Here is an example from a herbarium generated for the interactive program “L’Herbier” [“Herbarium”] that we created with the sculptor and artist Miguel Chevalier (<http://www.balpe.name/Herbier-anglais>):

Psychotria hypnagogica “Ken Kesey”

Several types of Arrowwood grow in France but not all would be suitable for use in construction. One of the most common is *Scrophularia Canadensis*. It is impossible to ignore this American Dogwood (*Geum Tricolor*); so many lines were written on the subject. The closest plant is the scented aloe. Who knows that this Digit Gold (or *Cicuta Digitalis*) is an European plant, scented, arrowheadshape, with numerous uses and often described. The two upper petals are stiff, and in the shape of diamonds. Its complex stems are precious to composers by the evil nights, and the landscape draws its brackets. According to Marie Anne Alisson de La Tour that grassy plant grows abundantly around the great palace of Paris or the strait of Gibraltar; José Vasconcelos says it would rather be close to the apartment of the English queen. There are seventeen species of *Melampyrum*. *Thaumatococcus* call it diaphanous aconite. Rootstock is very rich in g-coniceines. *Sansévéria* bloom primarily in spring. This *Pixellia* in parallel planes seeds, long employed to stop cerebral fevers, have made Pithiviers’s fortune. When the weather is favorable, its fruits mature in February. Get four seeds of this wine, add three leaves of *Scrophularia Zacathechichi* and three grapefruits, cook them all in some orange juice, boil it : you get a sure poison. You had no right to notice your rafters in the eyes of others: that's that rare plant.

And here, for comparison, is a counter-haiku produced by the same program for generating French, but whose language was specifically programmed (<http://www.balpe.name/Contre-Haikus>):

Shutters shaking in the wind

Sky almost white

Put speech on slow

The different levels of language are not the only elements to be used in digital works. Language strata which have visual and often sound or musical effects are also operative, but only if each has its own layer of code. It is very rare for the language of digital literature to seek to produce only text. Most of the time it is a multimedia production using the expressive capacities of other media such as images, movement, videos, etc., as in the many works I am involved in on www.balpe.name. So every new experimental digital work using the same French-language text generation program requires linguistic codifications of different complexity. For example, among others:

<http://www.balpe.name/Capture-Subversion>

<http://www.balpe.name/L-Esprit-Humain-2-200>

<http://www.balpe.name/Monstrueux-Monstres-2-197>

All these layerings of code and of language [*de code, de langue et de langage*], are a complex system compared to the simple natural writing of a text by an author. And it is legitimate to ask why it seems necessary to create digital literature when it requires such an abstruse instrumentarium for aesthetically focused written language. Yet the possibilities are vast, as is clear in contemporary literary activity in digital form. The many original perspectives opened up depend solely on the choice and development of new languages.

The specific potential of digital linguistic production, and hence of digital literature, includes uchronia (the work can be displayed simultaneously in different times, and its temporalities are programmatic); utopia (in the Greek sense of *u-topos*, that is, without place, since it can be produced simultaneously, or not, on many different devices; interactivity; generativity; multiplicity. These are but some of the features of the language of digital literature on which authors can draw for their creative works. All these capacities are condensed in a concept I borrow from Olivier Auber's recent book, *anopticality*.³

Anopticality denotes a historical change in perspective in our understanding of things. Olivier Auber explains that the way we perceive the world has changed throughout history. One of the most significant of these cultural changes was our adoption (invention?) of perspective, during the Italian Renaissance. To which I would add the – almost simultaneous – invention of printing. These two “inventions” produced a new vision of the world, which was centred on one point alone: rigid monarchic hierarchies, visual representations focused on a single “vanishing point,” the conception of the literary work as converging towards a single end (in which the novel's predominance originated), and so forth. In the early twentieth century, another of these cultural changes occurred, with Cubism's introduction of non-optical (anoptical) representation, and abstract art. Here is not the place to analyse in detail the transformations made possible by these changes. Suffice it to say that our relation to the world has changed, and we are no longer in a linear and powerfully organised vision of the world, but in a diffuse, fragmented, and immediate apprehension, on many different screens and in other media. This new conception of perspective implies that individual behaviour takes on greater importance, adapting to constant change, and readapting in constant exchange and open collaboration with others.

This is what, ever since I started working on machine-generated texts, I have attempted to represent in my many generative writings, through all my installations, and above all since *La Disparition du général Proust* [*The Disappearance of General Proust*], which can only be understood through the thirty-odd writing spaces I have scattered across the Internet. These can be accessed only through the reciprocal links they contain. (One of the possible entry points is the *Écrits* de Marc Hodges [*Writings* of Marc Hodges], at <http://sensdelavie.canalblog.com/> and the links this site contains). Due to how spaces have developed on the Internet, some of these links will have disappeared, or be

3 Oliver Auber, *Anoptikon. Une exploration de l'internet invisible*, Roubaix, FYP, 2019.

frozen. Since *La Disparition...*, I have been attempting – but the term is too metaphorical – to create a post-Cubist literature where every textual element implies a change of reading perspective – a perspective that is therefore constantly shifting – and where, even if every textual element taken alone can seem to have internal unity, the text itself can only be understood by tracing random pathways through the whole. Some readers will, of course, consider that this represents something between wastage and chaos, a sort of illegible maze (whose labyrinthine forms are called in French *dédales*, which is the name I give the work's parts in one of my main entries). These readers feel completely lost, and they refuse to make the effort needed to enter into it and move through it, to abandon themselves to a reading process which is constantly searching for an equilibrium. The reader will find no familiar markers, and the narration obeys a sort of scattered, Brownian motion, which is always looking for new stabilities. Maybe – it's not that important. What I seek to do is to imagine the potential of an anoptical literature, by showing in its texts and even its functioning how massively the possibilities opened up by the digital age have transformed our relations to texts. To create a literature that presents how human tongues and languages [*des langues et des langages humains*] are constantly changing.

And if the technical possibility existed of improving how my codes and programs work, I would try, even now, to achieve generative texts whose every word and every sentence welcomed interactivity, that is, where each element would be modifiable by the reader at any time and at any point on the globe. Each element would immediately be translated and adapted into any language, to form ultimately a constant movement of changing texts.

You will find some experiments around this in the following examples:

```
http://www.balpe.name/Le-monde-selon-Rachel  
http://www.balpe.name/Litanie-eternelle-de-voeux-infinis  
http://www.balpe.name/Participez-a-Un-Monde-Incertain  
http://www.balpe.name/Un-Monde-Incertain-par-un
```

Each reader request elicits a different text, although each text is linked, through different sorts of links, and is dependent on all the other possible ones. Thus the reader can – a possibility refused by some – read them in very different ways depending on the potential and virtual pathways through the whole. This is what we call a universal anoptical literature, where texts cannot be read as classical narratives, but require a radically new mode of reading, nearer to a form of writing, in the sense that the reader is immersed in a world of texts, a bit like in a video game. Here, the reader can encounter a new vision of the text and, consequently, of the totality of a language [*langue*], its literature and, perhaps, the world of that language. Something like always-renewed experiments in the writing of literary texts. Something like a permanent snapshot of human literature. And if this dream could be programmed in today's new and powerful programming languages, it could certainly become a reality.

Bibliographie

Auber, Olivier, *Anoptikon. Une exploration de l'internet invisible*, Roubaix, FYP, 2019.

Chomsky, Noah, *Aspects de la théorie syntaxique*, Paris, Seuil, 1971.

Greimas, Algirdas Julien, *Sémantique structurale. Recherche et méthode*, Paris, Larousse, 1966.

Meehan, James R., "The metanovel: Writing stories by computer," *Outstanding Dissertations in the Computer Sciences*, 1976.

Meehan, James R., "Tale-Spin, an interactive program that writes stories," in *Proceedings of IJCAI*, 1977. [Online] <https://www.cs.utah.edu/nlp/papers/talespin-ijcai77.pdf> [accessed 21 July 2020].

Saussure, Ferdinand de, *Cours de linguistique générale*, Paris, Payot, 1995.

Schank, Roger C. and Abelson, Robert P., *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*, Hillsdale, Lawrence Erlbaum Associates, 1977.