The Interaction of Corpus, Text and Genre in
"Structural Semantics" of Greimas

Driss ABLALI
Université de Lorraine

Abstract
This paper focuses exclusively on the specific nature of the semiotic project developed in Greimas's "Structural Semantics," a project which needs to be distinguished from the rest of the semiotician's work. The aim, in short, is to establish that the driving force of this text, from both the epistemological and heuristic viewpoints, is based on three descriptive categories: text, genre and corpus, which open up the perspective, not of textual or discourse semiotics, but of corpus semiotics.

Keywords: Semiotics, structural semantics, text, genre, corpus

1. Preamble

What I will try to show in this paper relates to the specific nature of the semiotic project contained in the three-hundred pages of "Structural Semantics" (henceforth referred to as SS). This was an uncompleted project—unlike the one which would be developed from "Du Sens," accompanied by theoretical renunciations or ‘acts of force’—which involved corpus semantics, intersected by ontology, perception and the physical world, and was rooted in reflection about texts as being an integral part of discourse genres.

To describe the intellectual coherence of the SS project from the basis of three descriptive categories, corpus, text and genre, I will proceed in three stages: the first part will be mainly focused on the question of the corpus, a concept that is rarely, or never, associated with Greimas; the second will deal with the holistic concept of meaning as it is used in SS thanks to the text/corpus linking; the third, in the wake of the two preceding parts, will show that there exists in SS, contrary to what is still thought, a semiotic reflection on the category of genre, which is indispensable to the category of narrative.
2. The question of data: variables in the corpus

As will be seen in what follows, *SS* is a semiotic project not of text or discourse, but of the corpus. Before dealing with the fate reserved by Greimas for these three notions (text, discourse, corpus), it would probably be advisable to take a look at how they fare in *SS*, and by this means venture a way of appreciating his semantic theory of corpus. Despite many criticisms which are not always founded, Greimas’s thinking about the nature of observables, to describe signification, was frequently linked in *SS* to corpora and the conditions under which they were compiled. Nowadays, however, it seems that work on corpora readily bypasses any reference to *SS*. This statement should nevertheless be qualified: although Greimas is neglected by specialists in linguistics or linguistics of or on corpora, we should, to put the record straight, recall that the word ‘corpus’ is used 109 times in *SS*. It is not a minor issue. This at least is what stands out, in my eyes, from comments of the following type: from the first occurrence of the word on page 43 (“We will take the liberty of dividing up this corpus as we see fit”), one clearly sees the direction that Greimas wished the semiotic project to take, namely to construct an objective semantic theory. Greimas judged that the data to be analysed had to be organised in accordance with well-founded linguistic and extra-linguistic criteria, among them criteria of “homogeneity” and “representativeness”. We must here insist on the fact that, in the 1960s, the word “corpus” was not yet a preoccupation among linguists. In France, it was not until the late 1970s that a new trend appeared in linguistics, notably in the context of discourse analysis—a trend which was officialised by a crop of works explicitly aiming to extract the discipline from the phrasal straitjacket. Before attempting to understand the reasons for favouring corpus over other types of data, like sentence, text and discourse, let us recall the delicate issues linked to the difficulties of deciding whether to keep or reject texts in a corpus. For Greimas, collecting a corpus was not about considering just any set of texts with no relation to one another. Aware of this fact, Greimas instituted what he called the “non-linguistic homogeneity of the corpus” (p. 94) which was necessary for data collection. The corpus was not a simple juxtaposition of independent and disparate fragments; according to Greimas, its collection had to meet very specific selection criteria in order to constitute a “homogeneous” entity that he defined as follows: “One clearly sees what should be understood, in this particular case, by non-linguistic homogeneity of the corpus; what allows fifty or so non-individual responses to be gathered together in a collective corpus is a set of characteristics shared by those being tested: the fact that they belong to the same linguistic community and the same
age range; and also that they share the same level of culture, the same ‘situation of examinees’”. (1966: 94). Certain forms of expression provide indisputable evidence of this—statements such as: “The homogeneity of the corpus appears to depend on a set of non-linguistic conditions, on a situational parameter relating to perceptible variations either at the level of the speakers, or at the level of communication volume”. (1966: 144).

Greimas is aware here of the impact of collected data on the analysis results; he therefore draws attention to the fact that analysis of the textual data is supremely dependent on the initial choices of what makes up the corpus, which guarantee an intrinsic basis to the analysis. He indicates this in plain language:

“The idiolectal character of individual texts does not allow us to forget the eminently social aspect of human communication. It is therefore necessary to widen the problem by introducing the principle that a certain number of individual texts, on condition they are chosen according to non-linguistic criteria guaranteeing their homogeneity, may be formed into a corpus and this corpus may be considered as sufficiently isotope”. (1966: 93). A set of semiotic objects is therefore assembled, clearly announcing that the SS intellectual project is limited neither to signs, nor to phrases, nor even to the text. A close reading shows us that the semiotician must go beyond the text in order to be able to construct a corpus, corresponding to the widest semiotic world in which each element finds its semantic value. Among “The conditions of a scientific semantics”1 set up by Greimas, he considers the sign as part of the phrase, the phrase as a passage in a text, and the text as sample of a corpus. On this point, we can only partly share F. Rastier’s criticism when he reproaches Greimas for his conception of the text as a sign: “The simplest way of evading the question is to consider the text as a sign. It is the solution chosen by Peirce, as it is by Greimas and Eco (Cf. 1988, p. 32: “the Message is equivalent to the Sign”). This evasion obviously takes little account of the difference in level of complexity between the sign and the text, but it especially avoids thinking about the effect of the global upon the local, in this instance of the text upon each of the signs which go to make it up”. (Rastier 1997: 147). Why “partly share”? Because this remark requires a few clarifications. Yes, it is quite true that Greimas, and what is called the “Paris School”, accord no place within the transformational progression to levels that are above the text; but, with respect to analysis categories, the SS project is not the same as Greimas’s later work. Greimas is completely Rasterian in SS, and Rastier,

1 This is the title of the first chapter in Structural Semantics.
in relation to the SS project, is entirely Greimassian, as we shall see later. They are both in complete agreement here.

To return to SS, we will follow it a little further to emphasise that Greimas never thinks of the text as the sole product issuing from use of the linguistic system, but as interaction between this system and other norms. He expresses it very lucidly:

As far as linguistics is concerned, on the other hand, what allows us to gather fifty or so individual responses into a collective corpus is a set of characters shared by those being tested: the fact that they belong to the same linguistic community and the same age range; and also that they share the same level of culture, the same ‘situation of examinees’. (Greimas 1966: 93-94).

One cannot therefore say that there is no semiotic thinking about corpus in SS. Greimas even devotes a section in it to what he entitles “Forming the corpus”, where we can read explicitly how this semiotician conceives the question of relationship between global and local dimensions in textual analysis—by introducing, for the first time in the history of textual and discursive theories, complexification factors of signification which show that the digital did not invent the corpus. Four semiotic categories are introduced here to define signification in its trans-phrasal dimension (within a single text) and its intertextual dimension (from one text to another within a corpus): “corpus”, “discourse”, “text” and “genre”. Four levels are set up, with the need to distinguish between them in order to better articulate them. I have chosen this passage from among the more explicit: “The procedure which, logically, follows upon the formation of a corpus consists in the transformation of the corpus into text. The text, in fact, is a determined sequence of the discourse and, as such, can only be a manifestation of logomachy, from which only one of the chosen isotopies should be retained. We will therefore understand text (and, what comes to the same thing, meta-text) to mean all the signification elements situated on the chosen isotopy and enclosed within the limits of the corpus”. (1966: 145).

First of all, one fact is blatant here: Greimas is opposed to the idea of a compositionality of signification. Signification is not constructed phrase by phrase, still less sign by sign, as the whole of formal tradition makes out, but by associative links formed by different linguistic operations allowing different aspects of the semantic contents to be revealed. There are effects of semes, and of isotopies of narrative actions, not only within a single text, but between different texts in the same corpus. An intertextual
dimension, hidden by linearity, appears here, and on it rests the holistic aspect of signification which is constructed as a totality and is more than the sum of its parts. In fact, for Greimas, no text is seen in isolation. Access to signification emerges against a background of other texts in locality zones directly contributing to the construction of significance. As we can read in several passages of *SS*, the linguistic system is insufficient to account for the semantic regularities of any phenomenon. What unequivocally stands out here is the establishing of the corpus over the text, of the “global” over the “local”. Greimas therefore admits that the procedures for describing the modes of existence and manifestation of the semantic world must take account of the relationships between one text and another, something which is only possible within a corpus. Greimas explains this clearly: “This means that, if the actors can be set up within a story-occurrence, the actants, who are classes of actors, can only be so if based on a corpus of all stories: an articulation of actors constitutes a particular story; a structure of actants constitutes a genre” (1966: 175). My underlining).

This redundant correlation in *SS* between local textual descriptions and global description in a corpus enables us to understand the issue of articulation between the problematic of the sign and the problematic of the corpus, by always linking the first to the second. We should remember here that Bernanos is not the only corpus described by Greimas. Other types of discourse are subject to the scalpel of semiotic analysis, an analysis which always favours the global dimension of signification. Before any semantic analysis, a prior phase is necessary for the semiotician, namely the formation of the corpus. In other words, for Greimas, no text is seen alone: there is not the meaning of the text but the meaning of texts. This is the most valuable epistemological implication of *SS*, linking all the observables to be analysed to the corpus as global authority. It is the case for the “popular tale” (1966: 147), the “psychodrama narrative” (p. 214) the “Littré dictionary” (p. 43), the “play” (p. 177), the “game of chess” (p. 184), and a “questionnaire” collected from philology students at the University of Poitiers (p. 93)—all these show that textual corpora is the semiotician’s true activity. And this conception of the corpus or corpora is very far removed from what Rastier calls a “logico-grammatical” conception where “the corpus boils down to a sample of language, a pool of examples or attestations”. (2010: 35). Here, Greimas develops a “rhetorical-hermeneutic” conception which “takes account of the relations between text and text, which is not possible within a discourse”. (Ibid.) The passage which appears below makes Greimas a sort of indisputable precursor of corpus semantics, at the time of 1960s semantic and linguistic research, when the
corpus was relegated to a back seat in favour of examples and phrases, cut off from their contextual and discursive terrain. Above all, these questions show that the term corpus, for Greimas, was the only language observatory where the signification of semes, actants and figures could be constructed. And Rastier clearly has this passage in mind in the distinctions he makes today between “reference corpus”, “study corpus” and “virtual corpus” (2010: 16), hence the plurality of realities he points to in Greimas’s writing: “The practical question thus raised is knowing what signification should be attributed to the three possible corpora respectively: the corpus having the dimensions of a novel, the corpus of all the writings of Bernanos and, finally, the corpus of all the novels in a given society and historical period, and what structural correlations one can reasonably hope to find between the models that can be made explicit from such corpora. (1966: 148. My underlining).

For the heuristic bases of SS’s scientific project, the consequences of these choices are evident: to affirm that “signification does not pre-exist discourse” (p. 33) is to affirm that there can be no semantics without a corpus. “No seme or semic category, even if its designation is borrowed from the French language, is identical in principle to a lexeme manifested in discourse” (1966: 34). Or, a little further on: “Discourse, considered as a manifestation of language, is, as we have seen, the unique source of information about the significations immanent in this language”. (1966: 39).

But Greimas does not only insist upon the impact of the corpus for access to signification. He also ensures that the delimitations and defining criteria of textual corpora as entities are rethought. Modes of cohesion are therefore necessary for a set of texts to be able to constitute a textual corpus. This is what the following passage allows us to glimpse:

“A certain number of precautions and practical advice should therefore accompany this choice, so as to reduce, as far as possible, the element of subjectivity manifest in it. We shall say that a corpus, to be well formed, should satisfy three conditions: it should be representative, exhaustive and homogeneous”. (Greimas 1966: 143). And a few lines earlier in SS, this is confirmed: “Forming a corpus does not therefore simply mean preparing a description, because the value of the description depends, in fact, on this prior choice, and, conversely, one can only judge the value of the corpus once the description has been completed” (1966: 142-143). This conception of corpora, which would remain exclusive to SS, prefigures what we would read half a century later in Rastier’s La mesure et le grain. The same criteria advanced by Greimas in 1966 are now essential when configuring
a mass of data as a corpus. The following passage is clearly in the wake of, and an extension to Greimas’s thinking on the status and weight of the data gathered. Without taking into account here the development of Greimas’s thinking, this decisive fragment by Rastier should be quoted: “The very notion of corpus must be refined, for a corpus is not a set of data, still less a collection with no defined principle, embellished with the name of linguistic resource: as always in cultural sciences, the point of view which governs the formation of a corpus naturally conditions subsequent research. If a corpus’s representativeness has nothing objective about it and depends on the type of use planned, its homogeneity also depends on the type of research” (Rastier 2010: 80).

And to return to the fourth category of genre, mentioned above, still in relation to the levels of globality above the text, we might recall that in all the corpora described or mentioned in SS, Greimas does not forget to link genre and signification.

3. Textual genres and levels of textual complexity

Following in the wake of L. Hjelmslev, the figure of Greimas has always been linked to the question of text. However, as I have attempted to show above, in SS there is no text without a corpus. At the level of the analysis data, the text in itself does not exist. It is being inscribed in a corpus which gives it observable status, and this status is often referred to by Greimas in SS in direct association with the notion of “genre”. Semiotics has always been reproached with having made genre the unthought element in its signification theory, as indeed Greimas and Courtés wrote in the Dictionary of Semiotics, claiming that it was “founded on implicit ideological premises” (Greimas and Courtés 1979: 164). Here, once again, we must clearly distinguish the semiotic project specific to SS from the new perspectives taken by the semiotic project as a whole after 1966. Why make this distinction? For two reasons directly related to the genre category. Before examining them, let us first stress that this category crops up a hundred or more times in SS, while being completely absent from Greimas’s other semiotic analyses (from Maupassant to Sémiotique des passions, with “la soupe au Pistou” and L’imperfection in between).

The first reason concerns the place of the “genre” variable in compiling corpora which, in order to satisfy the homogeneity principle, must be built up according to certain variables allowing different sets of texts to be compared. In this conception of the corpus, as shown by the diagram below, which distinguishes several levels of complexity, the genre variable
is essential to a holistic description of meaning:

\[
\begin{array}{ccc}
\text{meta-genre} & \text{genre} & \text{genre roman du XX\textsuperscript{e} siècle} \\
\text{X} & \text{style de la personnalité} & \\
\end{array}
\]

Gide, Malraux and Bernanos did not only write novels, so if Greimas includes them together in this diagram, it is first of all because they all wrote in the same genre, the novel, and because secondly, they wrote novels dating from around the same period. By solely targeting twentieth-century novels, and not the novel alone, Greimas is clearly showing he is sensitive to the diachronic nature of genres, which are continually evolving, both synchronically and diachronically. This grouping of their texts to form a corpus has to take account of a more encompassing variable than the “author” one, which is judged to poorly represent the genre variable—hence the need to contrast texts by different authors which relate to the same generic category. A second point about the diagram, on which I would like to insist here, concerns the cultural dimension of genres. Greimas considers that analysing a corpus of texts belonging to the same genre enlightens us in two respects: the first relates to the types of language activities peculiar to each novelist, what he calls “personality ‘style’ genre”; and the second to the novel genre in twentieth-century French society. Here, Greimas starts from three texts (Gide’s *Les Faux-Monnayeurs* [The Counterfeiters], Malraux’s *La Condition Humaine* [The Human Condition] and Bernanos’s *Le Journal d’un curé de campagne* [Diary of a Country Priest]) in order to describe what he calls a “meta-genre,” which is none other than “an inventory of genres typical of a given linguistic or cultural community”. The question of genres is so crucial that any semantic modelling of the story’s structure depends on it. To put it another way, Greimas relies on analysis of the genres in a corpus for the purpose of showing the specific way each genre becomes a narrative, a specificity necessary to the setting up of the narrative from all textual genres. By relating the generic specificities of texts to actantial, thematic and figurative invariants, Greimas’s analyses attempt to fix the rules of the story within a body of epistemological principles; to start out from the “story-tale” genre as “story-occurrence” (1966: 217) and move towards a “meta-genre”, defined as a concatenation of generic micro-stories.
On these two points, one might reproach the Greimas of SS with anything other than forgetting the genre category, given the decisive role this plays in modelling the story’s overall structure, as can be seen from this passage:

“If an inventory of models is a stage on the way to constructing a genre of models, description can just as well be applied to the establishment of a text which is an inventory of genres. Insofar as one succeeds, for example, in defining the popular story as a genre, an inventory of all comparable genres can give rise to the description of a common meta-genre which would be the story, considered in its generality, or else some sort of sub-set of stories”. (1966: 147).

It is not the structure of the text’s narrative alone which counts for Greimas, but the structure of the genre. Analysis of the genre must lead to setting up the story, and the story, like “a unique signification structure” (1966: 213-214), is the consequence of the contrastive analysis of the different genres. In short, the genre of texts takes precedence over all the other variables with a view to making the descriptive models capable of being generalised—hence the multiplicity of genres described, evoked or quoted in SS: “popular Russian tales” (p. 147), “detective stories, Chinese tale, tales of espionage”, “analytical psychodrama” (p. 213-214), “theatre dramas” (p. 175), “longer stories, in verse or prose, de Vigny’s Moïse, Camus’s La Peste” (p. 97-98), “play” (p. 176), “poetry” (p. 58), and “literary genre” (p. 70).

A story’s structure must take into account the effect of genres on the semantic, actantial and figurative codifications. Because, Greimas insists, “semantics, which aims to be a social science, seeks to describe values and not to postulate them”. (1966: 58). And in this description, analysis of the story is conducted in such a way as to be able to formulate the rules whose application is conditioned by the effect of genre on the text. One remark is necessary here: when Greimas refers to the texts analysed, or to be analysed, he often calls them by their genre and not by their title. For example, he speaks of the popular Russian tale, rather than naming a specific tale. With the novel, he speaks of “Bernanos’s novel, and not the novel in general, just as, in his descriptions, he always insists on the novel’s subject, on the “novels of a society”, of “a given historical period”, “twentieth-century novels” and “novels-occurrences”. Genre is at the heart of the modelling of the story’s structure. It is the partial modellings of each genre which form a coalition and thus define the story’s global structure as “meta-genre”. The heuristic power of this procedure is due to the fact that
it enables the researcher to start from the regularities observed in a corpus to then unify multi-level invariants through defined norms within a class of genre: the genre here is what retains and qualifies the collected data to restore the story’s complexity in all genres. At the level of mass data collection, Greimas cannot do without genre, just as he cannot do without it to construct the “the achronic organisational model of contents, that we thus meet with in areas that are very far distant from one another”. (1966: 233).

As we saw earlier, in different passages of *SS*, analysis of the story is not envisaged without the complementary upper levels like text, genre and corpus. To construct a theory of the story, the text alone, cut off from the other levels, makes no sense in Greimas’s eyes. This, inevitably, raises a question: what became of this complementarity in Greimas’s other writings after *SS*? The answer, without further ado, is: in the case of Maupassant, up to “La soupe au pistou”, there is no further place for the complexity of levels above the text in the analysis of a story’s structure. This is the point at which to recall that the work on Maupassant involves only one short story, “Les deux amis”, and the analysis of a recipe involves only one text, “La soupe au pistou”. This method of analysing texts heralds a quite different approach on the part of Greimas, an approach which posits that the text is the only upper level of complexity, governing the lower levels like the narrative, the figurative, the thematic and the enunciative. By focusing the analysis on a single text, Greimas abandons the narrow articulation between text and corpus, which he himself had implemented and justified by reasoned argument, in *SS*. If this present article was concerned with giving a detailed history of the development of Greimas’s thinking on these points, attention would be drawn to the substantial change in his view of the relationship between text, genre and corpus, and consequently of the semiotic project as a whole.

In *Du Sens*, the work published five years after *SS*, and which brings together articles written from 1966 onwards, there appears a conception of the text showing that the text’s meaning no longer has need of textual configuration on several levels in order to be described. As we read in the introduction to *Du Sens*, formalisation has become the true path to analysing texts:

“It is through a narrow gate, between two indisputable skills—philosophical and logical-mathematical—that the semiotician is obliged to conduct his investigation into meaning” (1972: 12). And this investigation of meaning is only defined in *Du Sens* in order to describe the text, the text alone, and not the texts grouped according to whether they belong to
a particular genre, as we saw before in SS. Greimas now situates meaning in the text itself and not in the interaction between different texts in the corpus. This hypothesis is confirmed in the Maupassant, which involves analysis of a single short story. We have thus moved from the signification of an actant, of a figure or seme described from a corpus, to a conception of signification which understands the text as a “closed world”. To quote Greimas himself:

“It nevertheless remains true that certain figurative values which we shall seek to describe can be apprehended thanks to their recurrence in the closed text” (1976: 55).

Only the text makes sense, therefore, in Greimas’s eyes. The other levels of complexity, evoked in SS, no longer have a legitimate place in the analysis of stories, recipes or passions, such as defiance or anger. The central consideration, of data relating to the text alone, in any analysis by Greimas of literary, anthropological, architectural or pictorial discourse, provides an instant and unambiguous sign of a particular conception of meaning and signification: it is one that is strictly interested only in the close relations which the actants and actors, the semes and figures have with one another, within a closed textual configuration, independently of any other attachment, whether of corpus or genre. Because if Structural Semantics imposes on the analysis an overall unity which is broader than the text, in Maupassant’s Deux amis, Greimas follows the semes and figures step by step in their immediate environment, favouring a return to the text which ignores the links establishes between Deux amis and Maupassant’s other short stories. This allows us to discern the concept of meaning which Greimas chooses to adopt by wishing to consider this notion within the close relations observed in a narrow, limited context. This premise is not explicitly formulated in Greimas’s analyses, but it is presupposed in the way the problem is dealt with. This should be explained further. It was in the Dictionary of Semiotics that the Greimassian concept of textual analysis appeared in its entirety. With regard to the notion of genre, a problem then arose: unlike other entries in the Dictionary, genre was not defined, by Greimas and Courtés, as a key concept of semiotics. This, at least, is how it appears in the entry in question, as it does in many others—because, unlike the premises in SS, which linked text, genre and corpus, from the 1970s onwards primacy was clearly accorded to the specific structure of the text. Moreover, the dictionary entry under
“genre” makes a complete tabula rasa of SS’s most significant premises, with the entry under “genre” being defined in a general way which excludes it from the conceptual system of semiotics. The main lines are developed of what is called “the theory of genre in the European cultural context” (164), but without it being linked to the global unit levels of description of meaning, which are the corpus and discourse, as was the case in SS. This also holds true for the entry under “corpus”: the authors, having recalled its different meanings in linguistics, wish to show that it is an operating notion, not only for syntactic theories, but also semantic ones: “Thus it is possible to speak of syntagmatic corpora (all of an author’s texts) or paradigmatic corpora (all the variants of a story), while taking account of the fact that they are never closed or exhaustive, but only representative, and that the models which help to try and explain them will be hypothetical, projective and predictive”. (1979: 74).

Now, as regards Greimas’s analysis, from the 
Maupassant to the 
Soupe au pistou, only the correlations in a single text have meaning. The text is therefore perceived in isolation, cut off from other texts in the corpus in relation to which it acquires meaning. Furthermore, in the sub-entries of “corpus”, Greimas and Courtés refer neither to discourse, nor to genre, nor to text. It is rather the notions of “generation”, “lexicon” and “verification” which figure. This is of no small importance, because it is about isolating the text from its contours and upper levels in favour of an immanentist conception of meaning. Texts are therefore confined within their internal logic, cut off from exogenous determinations. This means that all the theoretical and epistemological advances of SS have remained, for Greimas, at the stage of virtualisation. For Greimas and his followers, it led to the idea of a total autonomy of signs pleading for uniqueness of meaning, as Rastier rightly reminds us in this passage:

> It would seem that the text in itself does not exist, and that textuality is an abstraction (as, in fact, is language). […] The universals in the matter are only […] universals of method which appear useful for describing discourse, genres and texts.

(1994: 6)

This is the stance which Rastier takes when, in the late 1980s, and in line with his work over the preceding decades,2 he proposes studying the

---

meaning of texts, in the SS tradition of Greimas, Pottier and Coseriu, while
taking account of global determinations (corpus, discourse, genre, text)
over the local (sign). It was therefore Rastier’s semantics which were des-
tined to build on the initiatives and legacy of Greimas’s *Structural Seman-
tics*, a work “whose programme has not yet been sufficiently developed”
(Rastier 2008).
Bibliography


