

François Rastier

Interpretative Semantics

[To appear in Riemer (ed.), *Routledge Handbook of Semantics*]

DRAFT TRANSLATION – ANY COMMENTS GRATEFULLY RECEIVED

nick.riemer@sydney.edu.au

*Our attempt will seem radical. We are /
sure that one day it will be blamed for not having been radical enough.*
Benveniste, *Baudelaire*, 14, f°1 / f°80.

Abstract. — Saussurean linguistics initiated a tradition of semantic investigation characterized by a non-referentialist and non-compositional conception of language (Hjelmslev) and a systematic description of contexts and texts (Greimas, Coseriu, Pottier). This stream of research was unified in Interpretative Semantics, whose programme was formulated in the mid-1980s. In recognizing the deficits of the logico-grammatical paradigm, Interpretative Semantics offers a unified theory extending from the word to the text and then up to the corpus. Since the global determines the local, the corpus being described bears on the meaning of the text, which in turn determines the meaning of its units right down to the morpheme. As in construction grammars, the problem of semiosis (the matching of content and expression) therefore takes on a criterial status.

Applications of interpretative semantics extend from the description of Amerindian languages to computational linguistics. In parallel with the development of the concept of text, interpretative semantics offers a new articulation between text linguistics, document-based philology, and the hermeneutics of finished works.

Since meaning is made from differences, the methodology adopted is historical and comparative, like that of anthropology and most of the social sciences. Recognizing the complexity of languages – defined as cultural formations – interpretative semantics is therefore a stakeholder in cultural semiotics.

1/ *Some context*

Interpretative Semantics (henceforth IS) first appeared as a research programme in France in the middle of the 1980s. Part of the Saussurean tradition, it rests in particular on a synthesis of European structural semantics as developed by such authors as Louis Hjelmslev, Eugenio Coseriu, Émile Benveniste, Klaus Heger, Kurt Baldinger, Horst Geckeler, Bernard Pottier, and Algirdas-Julien Greimas (see Rastier 1987 for an initial presentation of the theory).

In the 1970s the generative perspective was dominant in both Linguistics (in the form of the Chomskyan paradigm) and semiotics (in the form of the Greimasian paradigm). This perspective inherited the legacy of the philosophical grammars that predate linguistics' constitution as a science: at issue was the explanation of 'surface' linguistic phenomena through cognitive operations of a logical nature applying to deep structures, a process which received an axiomatic presentation. The interpretative problematic represents a break with this traditional dualistic approach.

Since languages' expression and content are inseparable, semantics cannot be an autonomous discipline: it only describes a methodologically determined *perspective* on signs and therefore must be complemented by a *perspective* on expression : syntax (partly), morphology, phonology and graphemics describe a complementary perspective on the same signs. Linguistics is thus defined as the semiotics of languages – quite independent of the semiotics found in logical positivism and the syntax/semantics/pragmatics division, which cannot be applied to natural languages.

Relative to the dominant paradigms in competition on the international scale, cognitive and logical semantics, IS opens up a third way. Indeed, it rejects both the cognitive and logical forms of dualism, expressed through the separation between ideas and signs, or between signs – or names – and referents. IS does not advance hypotheses about either the theory of knowledge or ontology, and does not deal with either representations or entities in the world. Indeed, it describes linguistic meaning and the meaning of oral texts without any appeal to conceptual or wordly realities, but as the product of *differences* between signs and other units, whether in context or within texts and corpora.

If linguistic meaning does not consist in representations, it still imposes constraints on their formation; thus, within texts, semantic structures promote various referential impressions.

In this difference-based problematic, value is the fundamental concept. (i) Value is the true reality of linguistic units. (ii) It is determined by the position of units in the system (and hence by differences). (iii) Nothing preexists the determination of value by the system. Value therefore is not a sign, but a relation between signifieds. It excludes an atomistic definition of the sign, which would provide it with *a priori* signification – for a signification is a result, not a pre-given fact. It proscribes the compositional definition of meaning, since it establishes the determination of the local by the global as a structural principle. As a result, a sign must be acknowledged not to be a universal concept, but a signified which is relative to a language, or, in fact, to a text and a corpus.

The logical and ontological tradition which has prevailed in grammar and subsequently in the language sciences has isolated the word from its connection with its referent, the phrase from its connection with a situation, and the text from its relation with a world, whether fictional or not. For this paradigm of *signification*, the basis of which is, at the end of the day, metaphysical, I think it useful to substitute that of *meaning* [*sens*], a term originally from the rhetorical and hermeneutic tradition. This allows us to break the triple isolation of sign, phrase and text: the word takes on meaning in the syntagm, the syntagm in the period, the period in the text, the text in the social practice in which it is produced, relative to other texts. As a result, since languages are not denotationally or psychologically ‘transparent’, both their content and their expression constitute an autonomous domain of objectification.

2/ General principles

A. *The semiotic environment.* — Furthermore, to avoid isolating signs and reifying meaning, it seems useful to recall the following principles:

(i) Since the characterization of signs depends on interpretative processes which vary according to context, the “same” sign can function as an interpretative ‘hint’, an index, a symbol, etc. The study of signs is therefore dependent on that of interpretative practices.

(ii) The object of semiotics is not made up of signs, but of complex performances such as opera, rituals, and so on. The complex precedes the simple, and since oral or written texts are the empirical object of linguistics, methodological procedures of a non-trivial kind are needed to delimit signs.

(iii) The differential characterization of texts and other semiotic performances presupposes the constitution and critical analysis of corpora.

(iv) By their nature, signs are neither instruments of thought nor the expression of perceptual impressions. The semiotic realm, which consists of complex performances, is what constitutes the human milieu: this milieu is not an instrument, but the world in which we live and to which we have to adapt ourselves. The interpretative problematic is therefore no longer that of representation, but that of *coupling* in the biological sense, extended to *cultural coupling* with the semiotised environment.

(v) Although pragmatics privileges the here and now, the human environment contains masses of objects which are absent, or which, at least, lack any immediate perceptual

substrate: they populate the *distal zone* of the semiotic environment, which is the intended object of sciences as much as religions. Since signs are not referential, they allow worlds to be created.

B. *Meaning*. — Briefly : (i) Meaning is a level of objectivity which is reducible neither to reference nor to mental representations. It can be analysed into *semantic features* (or *semes*), which are moments of stability in the interpretative process. (ii) The typology of signs depends on the typology of the interpretative processes which have them as their objects. (iii) Since it is made up of differences that are perceived and characterized in actual practices, meaning is a property of texts and not of isolated signs – which have no empirical existence. (iv) The meaning of a unit is determined by its context. Since the context is the entire text, *microsemantics* depends on *macrosemantics*. (v) The elementary textual units are not words but *passages*. A passage's expression will be referred to as an *extract*; its content as a *fragment*. (vi) On the semantic level, relevant semantic features are organized in such a way as to constitute *semantic forms*, such as themes, which emerge against the *semantic base*, consisting especially in isotopes (recurring identical semes). Semantic forms are moments of stability within series of transformations, both within a single text, and between different ones.

C. *Signs*. — (i) If the *morpheme* is the elementary linguistic unit, the *text* is the minimum unit of analysis, since the global determines the local.

(ii) Every text is derived from a genre which determines, without completely constraining, its genetic, mimetic and hermeneutic modes. The genetic mode is what regulates the text's production, the mimetic mode is what regulates the referential impression the text produces and the hermeneutic mode is what regulates its interpretation.

(iii) Every genre belongs to a discourse-type. Through its genre every text is therefore connected to a discourse-type.

(iv) Every text depends on a corpus and must be related to it in order to be interpreted.

(v) A text's preferred corpus is composed of texts of the same genre. The genetic and interpretative procedures within the text are inseparable from the interpretative procedures in the intertextual structure of the corpus.

D. *Languages*. — A language is made up of a corpus of oral or written texts and a system. The system as linguists reconstruct it is a rational hypothesis formulated on the basis of regularities observed in the corpus. Between the corpus and the system, *norms* play a mediating role: anchored in social practices, norms of discourse, genre and style testify to the influence of social practices on the texts which depend on them. To avoid the spurious opposition between a language as a system of forms and as the product of a culture, the system should be considered as possessing *rules* and *norms* of varying regulative force; for example, the norms of the French ballad differ from those of the English one.

Rules and norms presumably only differ in the regularities of their diachronic evolution. As is well known, words (lexemes, then morphemes) develop from the crystallization and erosion of syntagmatic strings; what is true of these linguistic units is no doubt also true of the rules which regulate their relations and thereby constitute them as units: rules are probably frozen norms of discourse.

In synchrony, every rule is contiguous with the norms that accompany – or, in fact, condition – its application: without them, for instance, indefinitely recursive, though grammatically correct, embeddings could not be prevented. The grammaticality of a phrase can therefore only be assessed if the discourse, the genre and the text from which it is extracted are known – in addition, obviously, to the dating and the geographical origin of the text in question. Elementary though this observation is, it reveals the inanity of the discussions of ungrammaticality and asemanticity which cannot fail to arise as soon as we agree to discuss unattested or out of context phrases. Thus, in contrast to that of a formal language, the system

of a *natural* language is in fact plural, unfolding in different structural regimes according to the levels and stages of analysis. Its local or regional arenas of organization are not unified in any hierarchy which attests to the existence of a unique and homogeneous system – as is confirmed by the continuous evolution of languages, whose systemic heterogeneity constitutes the internal driver of their unending alteration through a process of disturbance and readjustment.

Natural language is thus never the only normative system at work; a text (oral or written) is the point of junction, in the context of an actual practice, between a language, a discourse, a genre and a style.

3/ Stages/Levels of description

IS recognizes four levels of complexity, starting with the morpheme and extending to the period, the text and the corpus. Working backwards, the corpus determines the meaning of the text, the text determines the meaning of its parts, down to the period and the morpheme.

3. 1. *Lexical semantics* ('microsemantics')

3.11. *Semes*. — There are two principal levels of lexical complexity:

a) The *morpheme* is the minimal linguistic sign. For instance, the word *rétropropulseurs* ['back thrusters'] contains five morphemes : *rétro-*, *pro-*, *puls-*, *-eur*, *-s*. A word is composed of one or more morphemes. They are divided into a signifier (whether overt or not: e.g. in French, the singular of "substantives" has a zero signifier) and a signified, the *sememe*.

b) The *lexeme* is the integrated group of morphemes which constitutes the unit of signification. A lexeme may be composed of only a single morpheme (e.g. the preposition *to*).

A sememe is a structured whole of relevant features or *semes*. They are defined as relations of opposition or of equivalence within classes of sememes : for instance, 'bistouri' is opposed to 'scalpel' by the seme /for the living/; the opposition /animal/ or /vegetable/ distinguishes 'venomous' from 'poisonous' ; 'mausoleum' is opposed to 'memorial' in virtue of the seme /presence of the body/ but is equivalent to it in virtue of the seme /funerary monument/ (We use the following conventions: cited expressions are italicized, contents – here, sememes – are in quotation marks, semes are enclosed in slashes, and semantic classes by double slashes). Since semes are units particular to specific languages, we do not make any universalist hypothesis about them.

Two types of seme can be distinguished : (a) *Generic* semes index the sememe in semantic classes of higher orders of generality. (b) *Specific* semes distinguish sememes in the context of lexemes belonging to the same minimal class, e.g. 'poir'- and 'pomm'- in the context of 'poire' [pear] and 'pomme' [apple] or of 'poirier' [pear tree] et 'pommier' [apple tree] (but not 'poireau' [leek] or 'pommeau' [pommel]).

Semes can assume two different statuses according to their mode of actualisation, understood as the instantiation of a type by a token.

a) *Inherent* semes are inherited by default from type to token, unless contradicted by the context. Each of the semes in a type is an attribute with a typical value. For instance, in 'crow', the attribute (or semantic axis) <colour> has /black/ as its typical value. /Black/ is therefore said to be an inherent seme for 'crow'. But a contextual determination could very well prevent this inheritance taking place and impose an atypical <colour> value (e.g. *I see a white crow*). No inherent seme therefore appears in every context.

b) *Afferent* semes are divided into two classes. The first designate the relations that apply a minimal class of sememes (a *taxeme*) or of *semies* (taxemie) to another. For instance, in French

the members of the taxeme //‘man’, ‘woman’// are the objects of an application relation whose source is the members of the taxeme //‘strength’, ‘weakness’//. This kind of application explains so called connotational phenomena, as well as certain prototypicality ones.

The distinction between actualization and virtualization has to be specified in degrees of relevance. Four of these can be distinguished, according to whether the seme is neutralised (excluded) or virtualized (but able to be reactualized), actualized, or salient. For instance, in ‘Guillaume was the woman in the household’ (Zola), the seme /feminine sex/ is neutralized in ‘woman’. Furthermore, the seme /human/ is actualized, but not made salient; by contrast, /weakness/, though afferent, is salient.

3.1.2. *Lexical classes.* — Since the definition of semes depends on the structurally and contextually established semantic classes in a language, these classes need to be characterized.

1 - The minimal class is the *taxeme*. The sememe’s specific semes are defined within its scope, just like its least generic (taxemic) seme, e.g. /funerary monument/ for ‘mausoleum’ and ‘memorial’. Taxemes reflect situations of choice; for example, ‘bus’ belongs to the same taxeme as ‘underground’, unlike ‘coach’ (which belongs to the same class as ‘train’). Within a taxeme, different kinds of relation can be found: oppositions between contraries (*male, female*), between contradictories (*possible, impossible*), graded oppositions (*burning, hot, warm, cold, freezing*), implications (*mobilized, demobilized*), complementarity (*husband, wife ; theory, practice ; hunger, thirst ; sell, buy*).

2 - The *field* is a structured set of taxemes; for instance the field //forms of transport// includes taxemes like //‘bus’, ‘underground’// and //‘coach’, ‘train’//. In texts, sememes belonging to different hierarchical levels within the field can be juxtaposed (e.g. ‘Wine or Perrier?’ ‘Beaujolais or water?’).

3 - The class of highest generality is the *domain*. Each domain is linked to a type of specific social practice. Lexicographical indicators like *chem.* (chemistry) or *mar.* (marine) are domain indicators. In the written languages of developed countries, between three and four hundred domains can be counted. Their number, nature and content differ from culture to culture.

4 - Lastly, *dimensions* are classes of high generality, but not superordinate to the preceding ones. Small in number, they divide the semantic universe into broad oppositions like /vegetable/ vs /animal/ or /human/ vs /animal/. They are often lexicalized (cf. ‘poisonous’ vs ‘venomous’ for the first opposition, ‘bouche’ [human mouth] vs ‘gueule’ [animal mouth] for the second).

3.1.3. — *Interpretative operations.* — In context, lexemes’ meanings are determined by three operations which transform the significations available in language : semes’ activation or inhibition, and the propagation of activated semes from one sememe to another. These three operations observe laws of dissimilation or assimilation, which increase or reduce semantic contrasts. Without in any way claiming exhaustivity, we will now exemplify these three operations.

a) *Inhibition* prevents the activation of inherent semes, which are therefore virtualised. Phraseological uses provide excellent examples of this process. ‘Step up’ includes the inherent seme /spatiality/, ‘battlement’ includes the inherent semes /architecture/ and /verticality/. Both are activated in *Sir Bayard, noble knight, steps up to the battlement* but virtualized in *Sarkozy steps up into the breach*. If these semes are not fully suppressed, their perceptual salience is still reduced. The content ‘Sarkozy’, indexed in the domain //politics// prompts a generic allotomy with ‘battlement’, indexed in the domain //war//. Governed in this case by the principle of assimilation, interpretation inhibits certain semes which index the sememe in the //war// domain, throwing into relief those compatible with //politics//.

The law of dissimilation can also prevent semes’ activation. For instance, in the menu

formula *fromage ou fromage blanc* [‘cheese or fromage blanc’ (a yoghurt like soft creamy cheese, literally ‘white cheese’)] the first occurrence of *fromage* receives a restricted sense relative to the one it takes on in *fromage ou dessert* [‘cheese or dessert’]: all the inherent semes in *fromage* which are specific to *fromage blanc* are inhibited. Contrastively, it signifies ‘fermented cheese’, and the seme /fermented/ is salient.

b) *Activation* allows afferent semes to be activated — these are present in the semie-type in the form of categories rather than that of specified features (or, in terms of frame theory, that of attributes of unknown value). For example, the seme /standing upright/ does not belong to the meaning of ‘shepherdess’: it is simply one of the virtual features which can be inferred from the inherent seme /human/. Nevertheless, in the context *Bergère ô tour Eiffel* [‘Shepherdess, O Eiffel Tower’] (Apollinaire, *Zone*), /standing upright/ is actualised by the presence of the inherent seme /verticality/ belonging to ‘tower’. (When readers represent the shepherdess to themselves, they imagine her standing upright.) The law of assimilation thus applies in an equative syntactic construction.

Interpretative operations are not put into effect without any conditions. In each case, in order to trigger the interpretative process, it is necessary to distinguish: (i) the *problem* whose solution it results in; (ii) the *interpretant* which selects the inference to be made; (iii) the *reception condition* which lowers the activation threshold and allows or facilitates the process.

3. 2. *The semantics of the period, or mesosemantics*

Mesosemantics explains the intermediate level between the lexeme and the text and thus covers the phrase or, more precisely, the region extending from the syntactically functioning syntagm up to the complex phrase and its immediate connections.

A. *Case theory* distinguishes *agency zones* (Pottier, 1974; 1992, pp. 124-127; Rastier, 1997): an event zone of *primary agency*, two zones of *secondary agency*, one anterior to the event, the other subsequent to it; and lastly a zone of *dependence* where the event’s *associated circumstances* are located. We suggest distinguishing two forms of primary agency, which we will refer to as *intrazone* or *interzone agency*, depending on whether the content which they articulate is located within a single zone or between two different ones. Interzone agency is distributed across three pairs: identity-proximal, proximal-distal or identity-distal.

The primary agents are the nominative and the attributive cases, the ergative and the accusative, the sender and the receiver. Secondary agency includes the agents which are not engaged in the process underway in the primary agency, such as the initial agents – the final and the causal – and the final ones – the benefactive and the resultative.

(i) *Attribution*. — The distinction between the three anthropic zones allows us to specify various forms of attribution, distinguishing the situation where two zones are put in relation to each other from that in which two contents of a single zone are. Intrazone attributive predications correspond to so-called analytic propositions, and interzone attributive predications to synthetic propositions.

(ii) *Non-attributive predications*. — If transitivity in the strong sense is defined – respecting its etymology – as the crossing of a boundary, three kinds of transitivity can be distinguished: identity-proximal, proximal-distal and identity-distal, each of them open to two perspectives depending on whether the source of the action is located in one zone or in another.

(iii) *The associated circumstances*. — Associated circumstances situate the utterance and its agents with respect to the zones (on the axes of time, space, mode and evaluation), taking into account the fact that the zone in which the utterance is situated obviously cannot be the one where it takes place.

B. IS privileges agreement relations – so much so, in fact, that it describes dependence relations in terms of agreement, since agreement relations bind successive syntagms without

any *a priori* limit and thus do not break up the syntagmatic order of the text. It defines morphosyntax as a system regulating the spreading of semantic features: their reiteration in bundles establishes semantic bases; their combination in structures establishes semantic forms which develop throughout the text. Semantic analysis consists in describing the mechanisms which regulate the spreading of semantic features. Certain syntactic structures on the level of the syntagm encourage the spreading of semantic features, while others inhibit it. Beyond the period, other syntagmatic structures take over. Starting from the level of the syntagm, they are superimposed over syntactic structures, but have not yet been described by Linguistics, since they are a matter of norms rather than of rules.

The concept of isotopy can now be invoked – a notion in principle independent of syntactic structures and putative phrase boundaries. An isotopy can extend over two morphemes, two words, a paragraph, or a whole text. It is possible to distinguish between isotopies triggered by the recurrence of a specific feature (e.g. /inchoative/ in *The dawn lights up the spring* (Éluard), where this feature is recurrent in ‘dawn’, ‘light up’ and ‘spring’; or by that of a generic feature (like /navigation/ in *the admiral ordered the sails to be furled*). It is also necessary to distinguish isotopies prescribed by the functional system of language (Pottier’s *isosemies*), as well as those which are optional because they belong to other systems of norms. The problem of the connections between syntax and semantics therefore concerns the relations between the isotopies prescribed by the functional system of language and those governed by other norm systems. Five noteworthy cases can be highlighted.

(i) Absence of both optional isotopy and isosemy, e.g. *That uselessly to the but I Bianca cardinal the* (a sequence obtained by random excerpting from *The lovers of Venice*, by Michel Zévaco). This sequence is neither a phrase nor an utterance.

(ii) Isosemies without any optional isotopy, e.g. *the vertebral silence indisposes the licit veil* (Tessnière) or *a paved pupil paraded presbytrally* (Martin). Such utterances, which are syntactically correct, refer to no semantically identifiable domain; a logician could therefore call them absurd.

(iii) Two interlaced domain isotopies, e.g. *Shepardess, O Eiffel Tower, the herd of bridges is bleating this morning* (Apollinaire). The utterance triggers a complex referential impression, and remains indeterminable. (According to several authors, cognitive *blending* theories take up the principles of the analysis of isotopies presented in Rastier 1987).

(iv) An optional isotopy, but with the violation of isosemies, e.g. : *Once the train disappears, the station leaves laughing, looking for the passenger* (René Char). This utterance triggers a referential impression by referring to the domain //forms of transport// : ‘train’, ‘station’, ‘passenger’, ‘leaves’ all include a generic seme which indexes them to this domain. As a result, the utterance appears to refer to a counterfactual world: it remains determinable, but logically false.

(v) An optional isotopy, accompanied by isosemies, e.g. *The green signal shows the track is clear* (Tessnière); *Every woman, even the ugliest, has made their lover suffer* (Apollinaire); *without tacking, and with the wind behind it, Eric Loiseau’s catamaran won the race*. These utterances refer to the domains //forms of transport//, //love// and //navigation// respectively. In this kind of utterance, several sememes or semies are indexed to a single domain; no other sememe stands in contradiction to this domain. The utterance triggers a referential impression, and is therefore determinable.

C. While what is described on the microsemantic level is the activation of semes, the mesosemantic level deals with the problem of how they are incorporated into semantic forms or bases. Considered in isolation, syntagms and periods appear to be sites in which semantic forms or their components are constituted. Reinserted into the continuity of the text, however, they are sites in which bases and forms are maintained, a process which consists in their

continuation, reiteration or deformation. From this point of view, the syntactic structures in which they participate are means by which semantic features are channelled and semantic activity distributed: they mutually govern the course of semantic and expressive activity and constitute semiosis on their own level of complexity. Formalisms with their origin in unification grammars allow this course of activity to be described without resort to the hierarchical and ontological categorizations bequeathed to linguistics by school grammars, first and foremost the traditional inventory of parts of speech, which is inadequate for most languages (cf. Vaillant forthcoming on Caribbean creoles).

The articulation between the expression (determining the identity of the magnitude) and the content (determining its value) occurs within the *passage*, a privileged site of local semiosis. In the interpretative perspective, this local magnitude corresponds variously to a sign, a phrase or a paragraph. On the level of the signifier, the passage is an *excerpt*, placed between two blank spaces in the case of a minimal string of characters, or between two pauses or punctuation marks in the case of a period. On the level of the signified, the passage is a *fragment* which points to its immediate and remote left and right contexts. This is true for both the sememe and for the content of the syntagm or the period.

3.3. *Textual semantics – or macrosemantics*

A text is an attested empirical linguistic sequence, produced within a determinate social practice, and fixed on some kind of support. The production and interpretation of texts can be conceived of as a non-sequential interaction of various independent components – specifically, the thematic, dialectic, dialogic and tactic ones.

(i) *Thematic*. — IS describes the theme as a structured grouping of semes (a *semic molecule*). It is not necessarily dependent on any particular lexicalisation; in technical texts, however, themes have a privileged, or even exclusive lexicalisation.

(ii) *Dialectic*. — As it deals with intervals of represented time and the developments which occur within it, dialectic particularly includes theories of narrative. It is defined on two levels. The first, called the *event level*, appears in all texts structured by a dialectic component. Its basic units are *actors*, *roles* and *functions* – in the sense of types of action represented. *Functions* are typical interactions between actors: they are classes of process. Like actors, they are defined by a semic molecule and generic semes: thus, the *gift* is a pacifying function (a transmission function of ternary valency), the *challenge* is a polemical function (a confrontational function of binary valency). Functions correspond to actorial *valencies*. Functions can be grouped in functional syntagms; for instance, an *exchange* is made up of two transmissions, a *confrontation* of an attack and a counter-attack.

The agonistic level, which is hierarchically superior to the event level, has *agonists* and *sequences* as its basic units. An *agonist* is a type which constitutes a class of actors, a sequence is a ratification of functional syntagms. In general, the dialectic component of practical texts only includes an event level, while fictional or mythical texts add an agonistic level to it.

(iii) *Dialogic*. — Dialogic accounts for the modalisation of semantic units at all the levels of complexity of the text. A *universe* is the set of textual units associated with an actor or an enunciatory focus: every modality is relative to a site (a universe) and a reference point (an actor). For instance, when the narrator of Balzac's *Cousin Betty* speaks of a 'good bad deed', 'good' refers to the universe of the two characters, and 'bad' to his own universe.

(vi) *Tactic*. — This final element accounts for the linear arrangement of semantic units on all levels.

Each semantic unit, on the different levels of analysis, can therefore be characterised in virtue of the four components. Only a methodological decision can isolate these four components, which interact with each other simultaneously and non-hierarchically.

On the textual level as on others, units result from segmentations and categorisations over semantic forms and bases, which can be given the general name *morphologies*. Their study is divided into three sections: links between bases, for example in the case of genres which include several generic isotopies, like the parable; links between forms; and above all links between forms and bases, crucial for the study of *semantic perception* (cf. Rastier, 1991, ch. 7).

Depending on the components, semantic morphologies can be the objects of different descriptions. For example, in connection with the four components, a stable grouping of semantic features (a semic molecule) can be described as theme, actor, goal or source of a modal point of view, and as a position in the linearity of the text. Further, different types of productive and interpretative operation correspond to each component.

Description must reproduce the dynamic aspect of the production and interpretation of texts. The first step consists in describing the dynamics of these bases and forms: for example, the construction of semic molecules, their development, and their potential dissolution. These dynamics and their optimisations are parameterized in different ways according to genre and discourse, because the forms and bases are constituted and recognized in them as a function of different norms: the perception of semantic bases seems to be linked to rhythms, and that of forms to contours, in a way reminiscent of prosodic contours.

Accordingly, the meaning of a text is not something that is deduced from a series of propositions, but rather something that results from the perusal of macrosemantic forms, each with their own signifying capacity, and through the particular way they unfold and the valuations which are conferred on them. One thus encounters analogous problems in the understanding of texts to those posed by the *recognition* of simulated or incomplete forms.

The morphosyntactic conception of the text can be modelled by dynamical systems theory, in which case the semantic bases appear as series of regular points, and the semantic forms are individuated by their singular points (cf. Rastier, 1999).

Accordingly, beyond a concatenation of symbols, the text can be conceived as a semiotic *course of action* (any text, as the semiotic part of a social practice, takes part in a codified set of actions). The genre codifies the way in which the action is conducted, but what could be called the *ductus* specifies an utterer, and allows the semantic style to be characterized by the particular rhythms and outlines of the forms' contours.

The generation of a text consists in a series of metamorphisms (transformation relations between forms) and transpositions (transformations of relations between forms and bases), which can be revealed in speech by the study of reformulations and in writing by the study of drafts. Its interpretation consists in the main in the identification and evaluation of metamorphisms: for instance, the meaning of a narrative is articulated through thematic and dialectic transformations.

3.3. *Corpus semantics* – or megasemantics

An isolated text has no more existence than an isolated word or phrase : to be produced and understood, it must be related to a genre and a discourse, and, via them, to a type of social practice.

A corpus is a structured grouping of whole documented texts, potentially enriched through labelling and assembled (i) reflexively theoretically, taking discourses and genres into account (ii) practically, with a view to a range of applications.

Several details need to be clarified at this point. (i) The *archive* brings together the entirety of documents available for a particular application or descriptive task. It is not a corpus, because it is not put together for a particular investigation. (ii) The *reference corpus* is constituted by the entirety of the texts. It will stand in contrast to the study corpus. (iii) The *study corpus* is delimited by the needs of the application. (iv) Finally, the working sub-corpora vary from one

stage of the research to the next and may only contain the relevant passages from the text or texts under study.

Corpora are thus not merely reservoirs of citations, nor even collections of texts. As long as they are set up critically, taking into account genres and discourses, and protected by the necessary philological guarantees, they can become sites for the description of the three regimes of textuality – the genetic, the mimetic and the hermeneutic. In point of fact, a text has its sources in a corpus, it is produced from this corpus and has to be maintained or reinserted in it in order to be correctly interpreted: the genetic and the hermeneutic regimes are thus synchronized with each other. As for the mimetic regime, which determines the referential impression, it also depends on the corpus and especially on the doxa for which it gives evidence.

The well-confirmed correlation between global variables like discourse, generic field, or genre and local variables (e.g. morphosyntactic ones as much as graphic or phonological) leads us to raise the problem of *textual semiosis* (cf. Rastier, 2001, 2011). Semiosis is usually defined on the level of the sign, as a connection between signified and signifier; however, the effect of a genre is precisely in defining a normative connection between signifier and signified on the textual level: for instance, in the genre of the short story, the first paragraph is most often a description, not an introduction, as it is in a scientific article. The local and conditional semiosis that language structure provides on lower levels of complexity, from the morpheme to the lexeme, is only realized if it is compatible with the generic or even stylistic norms which make textual semiosis possible.

4. Methodology of an instrumental semantics

The essentially lexical character of currently available semantic theories and their adoption of mentalist postulates (cognitive semantics, prototype theory, etc.) make the methodological requirements of semantics even greater. Corpus semantics, however, may advance reflection on the methodological (practical) and epistemological (theoretical) levels alike. An instrumental semantics may render new data objectively observable and hypotheses of different kinds testable.

Since IS has the ambition of *applicability*, its methodology proposes to reconcile three requirements: the principles of corpus choice, the definition of hypotheses, and the choice of descriptive concepts (certain distinctions being able to be neutralized as a function of their applications). It uses software programmes as experimental instruments, particularly since certain functions (in Hyperbase and Txt) have been developed in light of the problematic to which SI is addressed.

a) Since meaning consists of differences, the methodological incorporation of instrumental processes allows differences to be constructed – between words, passages, texts, authors, genres and discourse. Relevance emerges not from quantitative data, but from the meeting of two horizons: ‘subjective’ relevance as determined by the task itself, and ‘objective relevance’ specific to the qualitative differences within and between texts.

b) On the epistemological level, the recourse to experiment allows objectivity to be attained (i) by undermining or confirming hypotheses, (ii) by allowing the object’s structural regularities to emerge, when different instrumental procedures obtain congruent results despite differences in experimental material, scale, etc.

Corpus analysis allows polysemy to be relativised and reduced, and ambiguity to be checked for; it also allows the values of grammatical forms to be determined: for instance, the future does not have the same values in legal discourse as in novels.

A renewed relation to empirical data entails a new relation to theory: to articulate the connections between theory and practice more clearly, IS provides for simplifications as a

function of different applications. Lastly, corpus semantics, from the moment it adopts a reflexive point of view with respect to its own procedures, allows us to make a break from candid objectivism: it does not practise the automatic analysis of *data*, in so far as this must first be taken *as data*, then interpreted after processing.

New facts that remained unnoticed until not long ago, and that have now been made objectively observable, take on scientific import, because they are inconceivable for the most widespread linguistic theories. Most of these theories are based on the entirely suspect tripartite division between semantics, syntax and pragmatics, and can only consign the phenomena in question to rhetorical or stylistic research outside linguistics itself. Within this programme, corpus-based IS emphasises two general complementarities: that of linguistic levels or planes of description (morphology, syntax, semantics) and that of the levels of organisation and complexity (word, phrase, text, intertext).

5. Applications

Applications of IS concern disciplinary domains that deal with texts, just as much in the domain of the humanities (Latin and medieval corpora) as in contemporary corpora, whether literary scientific or journalistic.

i/ *Descriptive linguistics.* — IS has been enlisted in work on Romance and Amerindian languages (see especially Enrique Ballón-Aguirre and colleagues' work on Quechua agrarian vocabulary and on Chipaya (1992, 2002, 2009)).

ii/ *Textual semantics.* — Since IS is originally a text-based semantics, it can be applied in multiple domains, for instance to ancient and modern literary corpora (Amiri (2004), Ballón-Aguirre (2001), Canon-Roger and Chollier (2009), Choi (2006), Botchkarev (1999), Mézaille (2003), Gérard (2007)), to philosophical corpora (Loiseau and Rastier on Deleuze, 2010) and to scientific ones (Valette on Guillaume (2003), Djaoud, Hazem, Belghanem on Bourdieu (dans Rastier et Valette forthcoming), Poudat (2006) on linguistics articles). Pedagogic developments privilege the use of digital corpora in the teaching of grammar and literature (Mézaille forthcoming).

iii/ *Automated processing and corpus linguistics.* — In its application to corpus linguistics, IS has the task of renewing the domains of information retrieval and knowledge representation (cf. Pincemin (1999), Tanguy (1997), Thlivit (2000), Beust (1998), Perlerin (2004), Roy (2007)). In particular, IS can further applications which are increasingly the objects of social demand: recognizing a text type by its lexical or morphological characteristics; detecting a type of website; aiding in thematic analysis; achieving targeted dissemination by defining proximity between texts, etc. Most applications today presuppose characterization tasks : within a corpus, the aim is to single out the components relevant to the application. This reconnects linguistics, via a new path, with the problematic of the description of singularities, specific to the sciences of culture; the description of laws, long deemed the necessary condition of scientificity, is subordinated to the systematic study of actually occurring uses. Corpus linguistics also participates in the programme of cross-linguistic language comparison; but most of all, it allows this programme to be pursued within each language, by comparing discourses, genres and texts with each other (see Bourion, 2001, Rastier, 2011).

iv/ *Non-linguistic semiotics and cultural semiotics.* — The methodological principles governing the critical constitution of corpora are valid for all digital documents, for instance corpora of photos (Kanellos *et al.*, 2000) or web sites (Beauvisage, 2004 ; Trudel, forthcoming). An overview on the languages of icons is presented in Vaillant (1999). Other domains, such as interactive virtual narratives, are also used (Cavazza *et al.*, 2006).

The fact that cultural objects depend on their conditions of creation and interpretation does not alter the fact that the values which they actualize can still be rendered objective in the form of facts. Everywhere now people are dealing with digital corpora, whether they are of

music, images still or moving, dance, or polysemiotic performances like cinema, opera, rituals, etc. The scientific necessity of describing such corpora intersects here with social demand. With digital corpora, the sciences of culture discover new epistemological and methodological perspectives, and even a unifying project that could bring them together.

How to reconcile language and thought, content and expression, the supposed universality of human intelligence and the diversity of cultures? How to describe the human environment in its massive semiotisation? It is necessary to go beyond theories of the origin of language in order to better understand the emergence of the semiotic order, relying in particular on recent results of linguistics and anthropology (Rastier, 2013b). Since languages are human creations more than they are the providential products of evolution, summary oppositions between the innate and the acquired or nature and culture must be relativized. This task falls to cultural semiotics, to avoid our species disappearing before it has been described.

6. Références

General websites

Signo, bilingual semiotic theory website: www.signosemio.com

Texto ! Textes et cultures (online journal) : <http://www.revue-texto.net>

Introductory or summary works

Ablali, Driss, Badir, Sémir, et Ducart, Dominique (2013) *Documents, textes, œuvres*, Rennes, Presses Universitaires de Rennes.

Choi, Yong Ho (2004) *Tekst Umiron Kwanjei (Course in Text Semantics)*, Seoul, Ingan Sarang.

Hébert, Louis (2001) *Introduction à la sémantique des textes*, Paris, Honoré Champion.

Rastier, François (2002) *Semantics for descriptions*, Chicago University Press, CSLI Lectures Notes, 138 (chapter 4 : Marc Cavazza ; ch. 6 : Anne Abeillé).

Other references

Amiri, Bassir (2004) *Chaos dans l'imaginaire antique de Varron à l'époque augustiniennne. Etude sémantique et herméneutique*, ADRA-Nancy, Diffusion : de Boccard-Paris, 400 p.

Belghanem, Ali (2012, octobre). "Sémantique du discours scientifique de Pierre Bourdieu. Construction et classification d'un corpus de travail.", [En ligne], *Texto!*, Volume XVII - n°4 (2012), URL : <http://www.revue-texto.net/index.php?id=3068>.

Beust, P. (1998) *Contribution à un modèle interactionniste du sens. Amorce d'une compétence interprétative pour les machines*. Thèse d'informatique, Université de Caen.

Botchkarev, Andréï (1999) *Le motif végétal dans "A la recherche du temps perdu"*. Villeneuve d'Ascq, Presses universitaires du Septentrion.

Ballón-Aguirre, Enrique (2001) *Desconcierto barroco*. México: Universidad Nacional Autónoma de México.

Ballón-Aguirre, Enrique (2003) *De la semántica componencial a la semántica interpretativa (el léxico agrario andino)*, *Perfiles semióticos*, 1, 17-41.

Ballón-Aguirre, Enrique (2006) *Tradición oral peruana – Literaturas ancestrales y populares*. 2 vol. Lima: Fondo Editorial de la PUC.

Ballón-Aguirre, Enrique y Cerrón-Palomino, Rodolfo (1992) *Vocabulario razonado de la actividad agraria andina – Terminología quechua*. Cuzco: Centro de Estudios Regionales Andinos "Bartolomé de las Casas".

Ballón-Aguirre, Enrique y Cerrón-Palomino, Rodolfo (2002) *Terminología agraria andina - Nombres quechumaras de la papa*. Lima: International Potato Center – Centro de Estudios Regionales Andinos "Bartolomé de las Casas", 2002.

Ballón-Aguirre, Enrique y Cerrón-Palomino, Rodolfo (2011) *Chipaya. Léxico y etnotaxonomía*, Lima, Radboud Universiteit Nijmegen – Fondo Editorial de la PUC.

Beauvisage, Thomas (2004) *Sémantique des parcours des utilisateurs sur le Web*, Thèse de doctorat, Université Paris X.

Beust, Pierre, Ferrari, Stéphane, Perlerin, Vincent (2003) *NLP Model And Tools For Detecting And Interpreting Metaphors In Domain-Specific Corpora, Corpus Linguistics, Proceedings of the Corpus Linguistics 2003 conference*, Editors : Archer D., Rayson P., Wilson A. and T. McEnery, University Centre for Computer Corpus Research on Language, vol. 16, part 1, pp. 114-123.

Bourion, E. (2001) *L'aide à l'interprétation des textes électroniques*, Thèse, Université de Nancy II. Ed. pdf. <http://www.texto-revue.net>

- Cavazza, M. O. and Pizzi, D. (2006) Narratology for interactive storytelling: A critical introduction, in Gobel, S., Malkewitz, R. and Iurgel, I. A. (eds) *Technologies for interactive digital storytelling and entertainment*, Lecture notes in computer science, 4326. Heidelberg: Springer Berlin, pp. 72-83.
- Choi, Yong Ho (2002) Discursive space – an introduction to Text Semantics’, *Semiotic inquiry*, 12.
- Choi, Yong Ho (2004) ‘Text Semantics of François Rastier: a critical reading about the short story of Bernard Werber’, *French Studies*, 30.
- Choi, Yong Ho (2006) *Umi wa Sulwhasung (Meaning and narrativity)*, Seoul, Ingan Sarang.
- Canon-Roger, Françoise et Chollier, Christine (2008) *Des genres aux textes. Essais de sémantique interprétative en littérature de langue anglaise*, Artois Presses Université.
- Canon-Roger, Françoise et Chollier, Christine (2009) “A Comparison of Several Interpretations of ‘Snow’ by Louis MacNeice”, *Imaginaires : l’interprétation au pluriel*, Presses Universitaires de Reims, pp. 155-75.
- Canon-Roger, Françoise (2009) Traduction et réélaboration interprétative, *Revue Française de Linguistique Appliquée*, XIV, 2009-1, pp. 25-38.
- Chollier, Christine (2005) « Essai d’interprétation des rythmes sémantiques dans The Heart Is A Lonely Hunter de Carson McCullers », *Imaginaires* N° 11 (déc. 2005). Reims, PU de Reims. 255-272.
- Chollier, Christine (2009) A Comparison of Several Interpretations of ‘Snow’ by Louis MacNeice ». écrit en collaboration avec Françoise Canon-Roger. *Imaginaires*, 13, L’interprétation au pluriel / The Plurality of Interpretation, pp. 155-175.
- Chollier, Christine (2010) Rôles créateurs des contextes dans les parcours interprétatifs des passages ». *Texte !* [En ligne], Volume XV - n°3, URL : <http://www.revue-texto.net/index.php?id=2662>.
- Djaoud, Smaïl (2009) Quelques processus d’élaboration de concepts sur le Maghreb dans les sciences sociales (Julien, Bourdieu et Tillion), [En ligne], Volume XIV - n°1 (2009, URL : <http://www.revue-texto.net/index.php?id=2165>.
- Duteil-Mougel, Carine *et al.* (2012) Semiotics and semantic : tools for an effective appropriation of information, communication and health technologies, *Medetel 2012, The International eHealth, Telemedecine and Health ICT Forum For Education, Networking and Business*, Luxembourg, 17 – 20 April 2012.
- Gérard, Christophe (2007) Sémantique et linéarité du texte. La place du rythme en sémantique des textes », in M. Ballabriga (dir.), *Rythme et textualités*, Toulouse, Éditions Universitaires du Sud.
- Kanellos, Ioannis (2012) Patrimonial Traditions Meet Educational Preoccupations: The Interpretive Shift of the Accessibility Requirement. *Innovative Methods for Science Education: History of science, ICT and Inquiry Based Science Teaching*, Berlin, Frank & Timme, pp. 203-222.
- Kanellos, Ioannis, Thliviitis, Théodore et Léger, Alain (2000) Indexation anthropocentrée d’images au moyen de textes : arguments théoriques et directions applicatives du projet SEMINDEX. *In Cognito*, n° 17, pp. 33-44
- Loiseau, Sylvain (2010) « Investigating the interactions between different axes of variation in text typology », in Grzybek P. & Kelih E. (éd.) *Text and Language: Structures, Functions, Interrelations*, Wien, Praesens, pp. 109—118.
- Loiseau, Sylvain & Rastier, François (2011) Linguistique de corpus philosophiques : l’exemple de Deleuze, in Patrice Maniglier (dir.), *Le moment philosophique des années 1960 en France*, Paris, PUF, pp. 73-93.
- Mézaille, Thierry (2000) Accès sémantique aux banques textuelles - L’exemple de Balzac, in *Champs du Signe*, 10.
- Mézaille, Thierry (2003) *La blondeur, thème proustien*, Paris, L’Harmattan.
- Missire, Régis (2010), « Unités linguistiques à signifiant discontinu, du morphème au texte – une approche néo-saussurienne », in J.-P. Bronckart, C. Bota, E. Bulea (Eds.) *Le projet de Ferdinand de Saussure*, Genève, Droz, pp. 289-312.
- Missire, R. (2007), « Rythmes sémantiques et temporalité des parcours interprétatifs », in M. Ballabriga, P. Mpondo-Dicka (dir.) *Rythme, Sens et textualités, Linguistique, sémiotique du discours, sémantique des textes, rhétorique, stylistique, poétique*, Éditions universitaires du sud, pp. 75-115.
- Missire, R. (2004), « Norme(s) linguistique(s) et afférence sémantique : une lecture de *Sémantique interprétative* à partir d’Eugenio Coseriu (sistema, norma y habla) », in *Texte ! Textes et cultures*, vol. IX, n° 4, pp. 1-28.
- Missire, R. (2005), « Une larme baudelairienne, essai de description morphosémantique de *Tristesses de la lune* », in *Champs du signe*, 20, Toulouse, Éditions Universitaires du Sud, pp. 87-114.
- Perlerin, Vincent (2004) *Sémantique légère pour le document* Thèse d’informatique, Université de Caen.
- Pincemin, Bénédicte (1999) *Diffusion cible automatique d’informations : conception et mise en œuvre d’une linguistique textuelle pour la caractérisation des destinataires et des documents*. Thèse de linguistique, Université de Paris IV.
- Poudat, Céline (2006) *Étude contrastive de l’article scientifique de revue linguistique dans une perspective d’analyse des genres*, Thèse de doctorat, Université d’Orléans.
- Rastier, François (1971) *Idéologie et théorie des signes*, La Haye, Mouton, 1971.
- Rastier, François (1987) *Sémantique interprétative*, Paris, Puf, 1987 ; 2^e éd. augmentée, 1996 ; 3^e, 2009.
- Rastier, François (1993) Problems of cognitive semantics, in Franson D. Manjali, éd. *Language and Cognition*, New Delhi, Bahri Publications.
- Rastier, François (1997) *Meaning and Textuality*, Toronto University Press, tr. anglaise de *Sens et textualité*, Paris, Hachette, 1989, par Frank Collins et Paul Perron. Texte revu et augmenté.

- Rastier, François (1998) Sign and symbol — Semiotics and Cognitive science, in Manjali, F. éd. *Cognitive science*, Bahri Publications, New Delhi. Réédition en ouvrage de 1997 h.
- Rastier, François (1998) Prédication, actance et zones anthropiques, in Forsgren, M., Jonasson, K., et Kronning, H. édés, *Prédication, Assertion, Information*, Acta Universitatis Uppsaliensis, coll. Studia Romanica Upsaliensia, 56, pp. 443-461.
- Rastier, François (1999a) Representation or interpretation ? *Linguistics in the Morning Calm*, 4, Séoul, The linguistic society of Korea, pp. 115-135.
- Rastier, François (1999b) Cognitive Semantics and Diachrony, in Andreas Blank et Peter Koch, édés., *Historical Semantics and Cognition*, Mouton de Gruyter, Berlin (Cognitive Linguistics Research), pp. 109-144.
- Rastier, François (1999c) Sign and Symbol : Semiotics and Cognitive Science, in H.S. Gill, et G. Manetti, édés. *Signs and Signification*, II, Bahri Publications, New Delhi, pp. 199-208.
- Rastier, François (2000) On Signs and Texts : Cognitive Science and Interpretation, in Perron P. et coll., édés, *Semiotics as a Bridge between the Humanities and the Sciences*, New-York — Toronto, Legas Press, pp. 409-450.
- Rastier, François (2005) On signs and texts, *Applied Semiotics/Sémiotique appliquée*, vol. II, n° 4-5, pp. 195-244.
- Rastier, François (2009a) Interview, in F. Stjernfeld et P. Bundgaard, édés., *Signs and Meaning — Five Questions*, s. 1., Automatic Press, pp. 139-152.
- Rastier, François (2009b) Passages and Paths within the Intertext, *Belgian Journal of Linguistics*, 23, pp. 7-29.
- Rastier, François (2010) Web Semantics vs the Semantic Web : the Problem of Keyness, in Marina Bondi and Mike Scott, *Keyness in Texts*, Studies in Corpus Linguistics, 41, Amsterdam-Philadelphia, Benjamins, pp. 93-112.
- Rastier, François (2011) *La Mesure et le Grain. Sémantique de corpus*, Paris, Champion.
- Rastier, François, (2012) Text semiotics: Between philology and hermeneutics – from the document to the work, *Semiotica*, 192, pp. 99-122.
- Rastier, François (2013) *L'Homme de signes*, Paris, Éd. du Cerf, 2013.
- Rastier, François, et Bouquet, Simon, eds (2002) *Une introduction aux sciences de la culture*, Paris, PUF, 290 p.
- Rastier, François et Floetum, Kjersti, eds (2003) *Academic Discourse, Multidisciplinary Approaches*, Oslo, Novus.
- Rastier, François et Valette, Mathieu, eds (à paraître) *Concepts en contexte*, Paris, AFK.
- Roy, Thibault (2007) *Visualisations interactives pour l'aide personnalisée à l'interprétation d'ensembles documentaires*. Thèse d'informatique, Université de Caen.
- Tanguy, Ludovic (1997) *Traitement automatique de la langue naturelle et interprétation : contribution à l'élaboration d'un modèle informatique de la sémantique interprétative*, Thèse, Université de Rennes I
- Thlivit, Théodore (1998) *Sémantique interprétative intertextuelle : assistance anthropocentrée à la compréhension des textes*. Thèse d'informatique, Université de Rennes I.
- Vaillant, Pascal (1999) *Sémiotique des langages d'icônes*, Paris, Éditions Honoré Champion.
- Vaillant, Pascal (à paraître) La syntaxe, c'est de la sémantique, in Ablali et coll., édés, *Documents, textes, œuvres*, Rennes, PUR.
- Valette, M. (2003) « Conceptualisation and Evolution of Concepts. The example of French Linguist Gustave Guillaume », *Academic discourse – multidisciplinary approaches*, Kj. Fløttum & F. Rastier, eds., Novus Press, Oslo, pp. 55-74.
- Vaxelaire, Jean-Louis (2005) *Les noms propres – Une analyse lexicologique et historique*, Honoré Champion, Paris.