Interpretative semantics

François Rastier (translated by Nick Riemer)

1 Some context

Interpretative semantics (henceforth IS) first appeared as a research programme in France in the middle of the 1980s. As 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 that received an axiomatic presentation (see Chapter 3). 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 (in part), 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 (see Chapters 10 and 11), which cannot be applied to natural languages.

Relative to the dominant paradigms in competition on the international scale, cognitive (see Chapter 5) and logical semantics (see Chapter 4), 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 worldly 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: see Saussure, *Course in General Linguistics*, Part II, chapter 4). (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 pregiven 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”, their content and their expression jointly constitute an autonomous domain of objectification.

### 2 General principles

#### 2.1 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 “clue”, 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 semiotized 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.
2.2 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, micro-semantics 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 background, consisting especially in isotopies (recurring identical semes). For example, in the phrase *The admiral ordered the furling of the sails*, the generic feature /navigation/ is recurrent in “admiral”, “furling”, and “sails” and therefore constitutes an isotopy (see section 3.2 below). Semantic forms are moments of stability within series of transformations, both within a single text, and between different ones. A form is a cluster of semantic features: for example, in *Dawn lights the source* (Eluard) the cluster /inchoative+/flow+/clear/ is recurrent in “dawn”, “lights”, and “source”.

2.3 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 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.

2.4 Languages

A language is made up of a corpus of oral or written texts and a system (cf. Chapter 6). 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 (compulsory) rules and (non-compulsory) 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 scales 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 Scales of description

IS recognizes four scales 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, and the text determines the meaning of its parts, down to the period and the morpheme.

3.1 Lexical semantics (“microsemantics”)

3.1.1 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, substantif, the singular of substantifs “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] and “pommier” [apple tree]
(but not “poireau” [leek] or “pommeau” [pommel]).

Semes can assume two different statuses according to their mode of actualization, under-
stood 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) to another. For instance, in French the members of
the taxeme //“man”, “woman”// are the targets 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 neutralized
(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 seman-
tic 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 “mau-
soleum” 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 (**demobilized**
**presupposes mobilized**), 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?”).
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, 300–400 domains can be counted. Their number, nature and content differ from culture to culture.

Last, *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 seme 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 virtualized. 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 (créneau)* but virtualized in *Sarkozy steps up into the breach (créneau)*. If these semes are not fully suppressed, their perceptual salience is still reduced. The content “Sarkozy”, indexed in the domain //politics//, prompts a generic allotopy (domain discrepancy) 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 core meaning in the form of categories rather than 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 actualized 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, 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: 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 last, a zone of dependence where the event’s circumstantial complements 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 peripheral ones – the benefactive and the resultative.

In Mary lends a book to Paul for Tom, Mary and Paul are the primary agents, and Tom is a secondary one, located in a distal zone.

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 circumstantial complements. Circumstantial complements situate the utterance and its agents with respect to the zones (on the axes of time, space, mode and evaluation).

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 backgrounds; 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, as well as those that 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 (Tesnière) ou a paved pupil paraded presbyterally (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. Shepherdess, 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 compulsory isotopies, e.g. Once the train disappears, the station leaves laughing to look 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 truth-functionally determinable, but logically false.

v An optional isotopy, accompanied by isosemies, e.g. The green signal shows the track is clear (Tesnière); 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 backgrounds. 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 backgrounds and forms are updated, 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 scale 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 (2014) on Caribbean creoles).

The articulation between the expression and the content occurs within the passage, a privileged site of local semiosis. In the interpretative perspective, this local unit can variously correspond, among other things, 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 (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 lexicalization; in technical texts, however, themes have a privileged, or even exclusive lexicalization.

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 connection between analogous 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 modalization 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.

iv 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 characterized 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 categorizations over semantic forms and backgrounds, which can be given the general name morphologies. Their study is divided into three sections: links between backgrounds, 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 backgrounds, crucial for the study of semantic perception (cf. Rastier (1993: 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 backgrounds 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 backgrounds are constituted and recognized in them as a function of different norms: the perception of semantic backgrounds 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 incomplete forms.

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

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 backgrounds), 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 (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
Interpretative semantics

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, or 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 to be correctly interpreted: the genetic and the hermeneutic regimes thus mutually adjust themselves to each other. As for the mimetic regime, which determines the referential impression, it also depends on the corpus and especially on the commonsense beliefs for which it gives evidence.

The well-confirmed correlation between global variables like discourse, generic field, or genre and local variables (e.g. morphosyntactic as much as graphic or phonological ones) leads us to raise the problem of textual semiosis (cf. Rastier (2000, 2010)). 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 the application). 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 IS 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 and (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 relativized 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. Last, 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 emphasizes two general complementarities: that of linguistic levels or planes of description (morphology, syntax, semantics) and that of the levels of organization 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 the work on Quechua agrarian vocabulary and on Chipaya by Enrique Ballón-Aguirre et al. (1992, 2002, 2011)).

ii *Textual semantics.* Since IS is originally a text-based semantics, it can be applied in multiple domains, for instance ancient and modern literary corpora (Amiri 2004; Ballón-Aguirre 2001; Canon-Roger and Chollier 2009; Choi 2006; Botchkarev 1999; Mézaille 2003; Gérard 2007), philosophical corpora (Loiseau and Rastier (2011) on Deleuze) and scientific ones (Valette (2003) on Guillaume; Djaoud et al. (2012) on Bourdieu (in Rastier and Valette forthcoming); Poudat (2006) on linguistics articles). Pedagogic developments privilege the use of digital corpora in the teaching of grammar and literature.

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); Thlivitis (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 tasks of characterization (identification through contrast): 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 the semiotics of cultures. 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 2013). An overview on the languages of icons is presented in Vaillant (1999). Other domains, such as interactive virtual narratives, are also used (Cavazza and Pizzi 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 the human mind and the diversity of cultures? How to describe the human environment in its massive semiotization? It is necessary to go beyond theories of the origin of language to better understand the emergence of the semiotic order, relying in particular on recent results of linguistics and anthropology (Rastier 2013). 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 the semiotics of cultures to avoid our species disappearing before it has been described.

Further reading

Signo (www.signosemio.com) is a bilingual semiotic theory website.

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


References


François Rastier


François Rastier


Vaillant, Pascal forthcoming. La syntaxe, c’est de la sémantique. In Ablali et al. (eds) *Documents, textes, œuvres*. Rennes: PUR.


**Related topics**

Chapter 3, History of semantics; Chapter 6, Corpus semantics; Chapter 12, Lexical decomposition.