



**Michel Pêcheux**

**Automatic Discourse  
Analysis**

**edited by Tony Hak and Niels Helsloot**



# Michel Pécheux: Automatic Discourse Analysis

This One



GC3X-Y4W-KSXX

---

---

**Utrecht Studies in Language and Communication**

**5**

---

---

**Series Editors**

**Paul van den Hoven  
Wolfgang Herrlitz**

# MICHEL PÊCHEUX

## Automatic Discourse Analysis

Edited by  
Tony Hak and  
Niels Helsloot

Translated by David Macey



Amsterdam - Atlanta, GA 1995

© The paper on which this book is printed meets the requirements of "ISO 9706:1994, Information and documentation - Paper for documents - Requirements for permanence".

ISBN: 90-5183-645-7

©Editions Rodopi B.V., Amsterdam - Atlanta, GA 1995

Printed in The Netherlands

# Contents

Acknowledgments	vii
Part I            INTRODUCTIONS	1
1        Pêcheux's contribution to discourse analysis <i>Niels Helsloot and Tony Hak</i>	3
2        Theoretical issues in Pêcheux's Automatic discourse analysis (1969) <i>Paul Henry</i>	21
3        Pêcheux's theory of ideology and the subject <i>Françoise Gadet, Jacqueline Léon, Denise Mالدیدیر          and Michel Plon</i>	41
4        Pêcheux's linguistic references <i>Françoise Gadet, Jacqueline Léon, Denise Mالدیدیر          and Michel Plon</i>	49
Appendix: Pêcheux's approach to formalization <i>Françoise Gadet, Jacqueline Léon, Denise Mالدیدیر          and Michel Plon</i>	57

<b>Part II</b>	<b>AUTOMATIC DISCOURSE ANALYSIS</b>	<b>61</b>
<b>5</b>	<b>Automatic discourse analysis</b> <i>Michel Pêcheux</i>	<b>63</b>
<b>6</b>	<b>Overview and prospects</b> <i>Michel Pêcheux</i>	<b>123</b>
 <b>Part III</b>	 <b>EXPERIENCES AND PERSPECTIVES</b>	 <b>187</b>
<b>7</b>	<b>Theory, procedures, findings and perspectives</b> <i>Michel Pêcheux, Jacqueline Léon, Simone Bonnaïfous</i> <i>and Jean-Marie Marandin</i>	<b>189</b>
<b>8</b>	<b>Strategies for textual description</b> <i>Alain Lecomte, Jacqueline Léon</i> <i>and Jean-Marie Marandin</i>	<b>215</b>
<b>9</b>	<b>Three stages of discourse analysis</b> <i>Michel Pêcheux</i>	<b>235</b>
 <b>Bibliography</b>		 <b>243</b>
<b>Name index</b>		<b>251</b>
<b>Subject index</b>		<b>255</b>



## Acknowledgments

The editors and the publisher wish to thank Mrs. Marie-Germaine Pêcheux and the publishers who have kindly given permission for the translation and publication of the following texts: Dunod for *Analyse automatique du discours*, 1969 (Chapter 5 of this volume); Larousse for 'Mises au point et perspectives à propos de l'analyse automatique du discours', *Langages* 37, 1975, pp. 7-80 (Chapter 6); and Presses de la Fondation Nationale des Sciences Politiques for 'Présentation de l'analyse automatique du discours (AAD69)', *Mots* 4, 1982, pp. 95-123 (Chapter 7) and 'Analyse du discours: stratégie de description textuelle', *Mots* 9, 1984, pp. 143-165 (Chapter 8). We thank Gene Diaz, Joke Haafkens, Gerhard Nijhof and Brian Torode for their support during the preparation of this volume.

Rotterdam and Amsterdam, May 1995



## **PART I**

### **INTRODUCTIONS**



## Pêcheux's contribution to discourse analysis

Niels Helsloot and Tony Hak

In the 1970s and the beginning of the 1980s, Michel Pêcheux was the leading figure in French discourse analysis. Initially, his renown was based on his first major work, *Analyse automatique du discours* (1969), which was received enthusiastically upon its publication. This work on automatic discourse analysis and on its theoretical underpinnings – henceforth ADA69 – was only the beginning of a continuous process of revising and reformulating the principles of discourse theory and discourse analysis. Through this work, Pêcheux became a reference point for virtually all work on discourse analysis published in France in the 1970s and 1980s. To date, ADA69 and Pêcheux's theorizing are an integral part of French textbooks on discourse analysis (e.g. Ghiglione and Blanchet 1991, and Maingueneau 1991). The significance of Pêcheux's work is also evident from the commemorative issues of the journals *Mots* and *Langages*<sup>1</sup>, and from some recent anthologies<sup>2</sup>.

Outside of France, ADA69 found ready reception in Italy, Spain, Portugal and several Latin-American countries. As early as in 1972, an extensive summary of ADA69 was published in Italian (Cipolli 1972). In 1978, a Spanish translation of ADA69 was published. A similar Portu-

---

<sup>1</sup> *Mots* 9 (1984) and *Langages* 81 (1986). See also Mالدیدیر (1984). The latter text opens the volume *Histoire et linguistique* (Achard, Gruenais and Jaulin 1984), which is dedicated to the memory of Michel Pêcheux.

<sup>2</sup> Mالدیدیر (1990) offers an extensive overview of Pêcheux's work, including a complete bibliography (cf. Achard 1991, and Helsloot and Hak 1992 for short reviews). Hak and Helsloot (1991) focus on Pêcheux's later works.

guese edition appeared in Brazil in 1990.<sup>1</sup>

In the English-speaking world, Pêcheux is mainly known as the author of *Language, Semantics and Ideology*, the 1982 translation of his second book, *Les Vérités de la Palice* (1975). At that time, French discourse analysis was usually equated with 'post-structuralist' philosophical discourse, presented by such authors as Jacques Lacan, Michel Foucault and Jacques Derrida. Though more technically linguistic forms of French discourse analysis constitute the empirical counterpart of this philosophical tradition, these have not reached the English-speaking world to the degree they deserve. As a result, Pêcheux is known as a philosopher and a theorist of ideology rather than as a discourse analyst.<sup>2</sup> In this way, his significant contribution to the development of empirical strategies of discourse analysis has been overlooked. The aim of the present volume is to correct this one-sided reception of Pêcheux's work by offering an English translation of some key texts on his method of discourse analysis. These texts show how Pêcheux's philosophical and theoretical thinking is framed by problems of analysis.

## Epistemology

Michel Pêcheux was born in 1938. He studied philosophy from 1959 to 1963 at the *École Normale Supérieure*, where Louis Althusser and Georges Canguilhem were among his teachers. Like Michel Foucault and Jacques Derrida before him, he received training in Bachelardian epistemology. Gaston Bachelard had developed a critique of current philosophies of science by confronting them with the results of historical studies of the development of sciences (plural). On the basis of detailed historical analysis, he had shown that the constitution of sciences, such as physics, was the result of theoretical, rather than empirical, breakthroughs. Because

---

<sup>1</sup> The Spanish volume (Pêcheux 1978) includes the texts gathered in Part II of the present volume (Chapters 5 and 6). The contents of the Brazilian volume (Gadet and Hak 1990) are largely identical to the present volume.

<sup>2</sup> Examples of publications in English that discuss Pêcheux's work in isolation from his contribution to discourse analysis are MacCabe (1979), Cousins (1985), Macdonell (1986) and Williams (1992). Some of the few publications in English that discuss Pêcheux's work in relation to empirical discourse analysis are Woods (1977) and Thompson (1984).

of the importance given to theory in the constitution of sciences, Bachelard's epistemology was considered anti-empiricist.

Epistemology formed the background for, among others, the work of Althusser and Foucault in the 1960s, and provided these authors with their concepts, such as 'episteme' (Foucault) and 'epistemological break' (Althusser). The latter concept was coined by Bachelard in reference to the distinction between 'common' knowledge and scientific knowledge, i.e. between everyday experience and experimental technique (Bachelard 1949: 102). He calls the distinction between these two modes of knowing 'philosophically decisive' (Bachelard 1949: 103):

What is at stake is nothing less than the primacy of reflection over perception. [...] What mankind *makes* by means of a scientific technique does not exist in nature and is not even a *natural* result of *natural* phenomena.

Experimental evidence is constituted or produced:

[P]henomena must be selected, filtered, purified, shaped by instruments; indeed, it may well be the instruments that produce the phenomenon in the first place. And instruments are nothing but theories materialized. The phenomena they produce bear the stamp of theory throughout. (Bachelard 1984: 13)

Canguilhem, who took over Bachelard's chair in epistemology at the *École Normale Supérieure*, showed that the process of conceptualization and 'scientification', of breaking away from 'error', is not confined to the transition from common sense to science, but goes on within science. His notion of a continuous breaking away from 'error' became an important element of Althusserianism. According to Althusser, in order to establish itself as the science of history and society, Marxism must liberate itself from 'idealism' in a continuous struggle. In this context, Althusser and his students worked on a study of the 'conditions of production' of Marx's *Capital* (cf. Althusser and Balibar 1970).

Pêcheux's scientific career began within this intellectual climate, which introduced a feel for the historical and practical context of scientific concepts and 'discoveries'. Among his first publications are a summary of the concepts of Bachelardian epistemology (Pêcheux and Balibar 1969), a

lecture on the different effects of the 'Galilean break' in the fields of physics and biology (Pêcheux 1969c), and two articles on the theoretical conjuncture in the social sciences (Pêcheux 1969b, 1969d).<sup>1</sup>

Pêcheux and Balibar (1969)<sup>2</sup> describe an 'epistemological break' as a *point of no return* from which a science begins to exist. They emphasize that

the concept of break has nothing to do with the voluntaristic idea of a jump from ideology into science, with its inevitable implicit religious connotations. The *name* of Galileo in 'the Galilean break' is in fact a misnomer, because science is not the product of a single human being: Galileo represents the effect rather than the cause of the epistemological break that carries his name. (Pêcheux and Balibar 1969: 10-11)

Pêcheux and Balibar outline three effects of the epistemological break that constitutes a science. One of them is that specific pre-existing ideological and philosophical discourses are explicitly dismissed: the new science breaks away from them. A second effect is that retroactively some philosophies are validated and others invalidated: the break redefines the values within a conflict-laden philosophical field. A third effect is the *relative autonomy* of the new science: its further development depends primarily on its experimental practices and on theoretical events rather than on influences from outside. The latter effect implies that an epistemological break is a unique event in the history of a science. Subsequent theoretical

---

<sup>1</sup> Bachelard developed his epistemology by studying the history of natural sciences such as physics. Canguilhem widened this approach by studying life sciences such as biology. Subsequently, Althusser and Foucault applied epistemological concepts to the history of the social sciences. Pêcheux was the first in this tradition who studied the present conjuncture in the social sciences. He had been working in this field since 1966, when he obtained, with the support of Georges Canguilhem, a research position in the laboratory for social psychology of the *Centre National de la Recherche Scientifique* (C.N.R.S.).

<sup>2</sup> In fact, Pêcheux and Balibar do not claim to be the 'authors' of this text. It is presented as a summary of a lecture presented by François Regnault, which, for unclear reasons, could not be published (Pêcheux and Balibar 1969: 7). Regnault's lecture was part of a 'philosophy course for scientists' organized by Althusser at the *École Normale Supérieure* in the winter of 1967-1968 (cf. p. 7, note 1).



displacements cannot be designated with the term 'break' [*rupture*] (Pêcheux and Balibar 1969: 12).

In his contribution to Althusser's 'philosophy course for scientists'<sup>1</sup>, Pêcheux (1969c) analyzes how a conflict between two theoretical ideologies develops differently within the fields of physics and biology. In physics the 'Galilean break' led to a triumph of mechanistic explanations of magnetism and electricity over the animistic experience of wonder. Pêcheux draws special attention to the use of the word 'dynamic' as opposed to 'static', which inaugurates the new field of electrodynamics. Contrastively, in a 'transversal analysis', he shows that in biology the opposition between 'dynamic' and 'static' became connected with the distinction between physiology and anatomy. Here the introduction of the term 'dynamic' did not lead to mechanistic interpretations, but to the development of vitalistic perceptions of 'forces', which are still conceived as animate. Pêcheux links this divergence between physics and biology to differences between the social practices connected with these sciences: physics is applied mainly to the field of the means of production (e.g. machines), whereas biology, through medicine, is applied to the field of labour (humans).

In two other publications of the same year, Pêcheux does not make explicit use of epistemological concepts, but his approach is clearly Bachelardian. In an article on the conjuncture of social psychology, he states that this discipline is haunted by (the struggle between) two kinds of ideology – a biopsychological one on the 'biological' side of psychology and a set of political, religious and moral ones on the 'social' side of psychology (Pêcheux 1969b: 291). These ideologies rely on an unexplicated, ideological concept of the 'subject', which should be transformed theoretically by a critical application of psychoanalysis and Marxism. Pêcheux conceives this transformation not as something merely philosophical but rather as a difficult enterprise whose realization implies profound transformations in research practice. It is a "theoretical inter-

---

<sup>1</sup> This course consisted of the following contributions (Althusser 1990: 71; Wolf 1985: 155-156): Althusser, Introduction (5 lectures); Macherey, The empiricist ideology of the 'scientific object' (3); Balibar, From the 'experimental method' to the practice of scientific experimentation (3); Regnault, What is an epistemological break? (1); Pêcheux, Ideology and the history of the sciences (2); Fichant, The idea of a history of the sciences (2); Badiou, The concept of the model (2). More lectures had been announced but the series ended when the 'events' of May 1968 occasioned Badiou to not continue his lectures. Part of the lectures have been published: Althusser (1990), Badiou (1969), Fichant and Pêcheux (1969).

vention from the philosophical terrain into the practical work that is done by researchers in social psychology" (Pêcheux 1969b: 297).

In a contribution to a discussion among communists on how to influence the development of the human sciences, Pêcheux emphasizes the Bachelardian theme of the conceptual difference between (pre-scientific) 'experience' and (scientific) 'experimentation': "We must take issue with the idea that there are primary 'givens' from which theories could start: it must be stressed that a scientific problem can only exist within a *conceptual and instrumental-experimental field*" (Pêcheux 1969d: 75). In the same article we find a new theme as well, namely a critique of the ideological uses of formalization such as in Chomskyan linguistics:

We must make a distinction between those cases in which mathematics is applied technically and other cases in which it intervenes at a theoretical level between concepts and experimental devices. (An example of the latter would be generative and transformational grammars.) (Pêcheux 1969d: 76)

In generative and transformational grammars, formalization is not merely a technical application of mathematics. Formalization intervenes at a theoretical level. It represents the ideological conception that language *is* formal. According to Pêcheux, the critique and abandonment of empiricism and formalism is not merely a philosophical task – though Marxist philosophy will play a part –, it is rather a practical matter of political intervention in the institutions of research and education.

### **Discourse theory and linguistics**

Pêcheux's work discussed so far clearly demonstrates his immersion in epistemology and his application of Althusserian notions to an analysis of the state-of-the-art of the social sciences, but it does not throw much light on how this Bachelardian background relates to his interest in discourse analysis. At this point, some indications can be found in two articles that Pêcheux published under the pen name of Thomas Herbert. In his very first publication (Herbert 1966), Pêcheux applies Bachelardian and Althusserian concepts to the social sciences, in particular to social psychology. His diagnosis is that these sciences have not established themselves as

proper sciences, because they have not established their own theoretical object. In other words, they have not accomplished the necessary epistemological break with ideology. These 'sciences', therefore, do not produce scientific knowledge but reproduce the ideology of the social system. In line with the epistemology of Bachelard, Canguilhem and Althusser, Pêcheux claims that, in order to become proper sciences, the social sciences must go through a theoretical transformation in which both their objects and their instruments should be redefined. Whereas the current practice of the 'social sciences' consists of transforming (ideological) discourse into other (ideological) discourse, social science proper would transform that discourse into something else, which would be expressed in terms of a new theory. Pêcheux presents discourse analysis as a social scientific instrument that should be defined in terms of the social scientific theory that is yet to be established. For Pêcheux, the development of an instrument of discourse analysis, thus, was explicitly not a merely technical enterprise. On the contrary, it was part of the project of establishing a truly scientific social psychology.

In his 'sequel to the Herbert (1966) article, Pêcheux gives the following succinct summary of its main thesis: "Any science is initially the science of the ideology with which it breaks" (Herbert 1968: 74). In this second Herbert article, Pêcheux outlines the preliminaries of a 'general theory of ideologies', which would make possible a scientific study of how ideologies function. Such a theory is necessary for both an understanding of how ideologies can function as obstacles for the establishment of a science and of how they can be superseded through an epistemological break. For this purpose, Pêcheux makes a distinction between two forms of ideologies: 'empirical' ideologies (which have technical origins) and 'speculative' ideologies (which have political origins). At this point, he introduces linguistic terminology to characterize the different ways in which these two forms of ideologies function:

[T]he *empirical form* refers to the relation between a signification and a reality, whereas the *speculative form* refers to the articulation of significations to each other, under the general form of discourse. When we use terms imported from linguistics, we can say that the empirical form of ideology puts forward a *semantic function* – the coincidence of the signifier and the signified –, whereas its speculative form puts forward a *syntactic function* –

the connection of signifiers to each other. (Herbert 1968: 79)

The ideological process must be understood as a combination of the semantic effect and the syntactic effect. The first effect produces the reality of the signified, whereas the second assigns it its proper place between all other things that can be present in discourse *in the given ideological conjuncture*. The 'social sciences' are blind to the speculative functioning of ideologies because they approach discourses with only empirical means. This approach leads to a technical interpretation of speculation, which overlooks its effects within the empirical approach itself. The two Herbert papers outline the questions to be addressed by Pêcheux's automatic discourse analysis. Discourse analysis is aimed at a description of the functioning of ideologies in general, and of how this functioning is an obstacle for the establishment of a real social science in particular.

In this context, linguistics provides some essential concepts. Pêcheux uses the term 'metaphor' for the semantic production of reality, and he calls the syntactic relations between signifiers 'metonymic'. These terms hint at Pêcheux's indebtedness to structuralism, which had become popular in France through the influence of Jakobson on Lacan and Lévi-Strauss. The reputed founder of structuralist linguistics was Ferdinand de Saussure. A preoccupation with Saussurean linguistics was widespread in France in the 1960s (in particular in 'post-structuralist' circles). However, Saussure's *Course in General Linguistics* was more quoted than actually studied. Saussure was generally seen as the founder of structuralism, which was conceived primarily as a rigid and one-sided method.<sup>1</sup> According to this view, the object of Saussurean linguistics was *langue* (the language<sup>2</sup>)

---

<sup>1</sup> This received interpretation, which was inaugurated by Charles Bally and Albert Sechehaye, the editors of the *Course*, and which still prevails, is disputed in rereadings of Saussure's work by Godel (1957), Engler (1967-1974), De Mauro (1972), Starobinski (1979), and Marinetti and Meti (1986).

<sup>2</sup> Throughout this volume *langue* has been translated as *the* (or *a*) language in order to make a systematic distinction in the English text between the Saussurean concept of *langue* (the language as a system; with an article) and the concept of *language* (language in general; no article). Although in some places the use of the article may seem artificial, it is necessary to stick to it because the theoretical distinction between the two concepts (*langue* and *language*) is absolutely vital for a proper understanding of the discussion on Saussurean linguistics.

as a system of formal relations – in isolation of its use, its history and its content, which were referred to the domain of *parole* (speech). 'Post-structuralists' tried to adhere to the (supposedly Saussurean) structuralist method while counteracting its one-sidedness.<sup>1</sup> In the footsteps of the anthropologist Claude Lévi-Strauss (e.g. 1969), their work was directed to problems in other disciplines than linguistics: Lacan directed his analysis to the structure and workings of the unconscious; Derrida's work focused on philosophical and literary analysis; Foucault's observations on language and discourse were part of his historical analysis of the present.

Pêcheux addressed the question of the theoretical place of 'discourse' within the Saussurean model. His problem was that such questions as 'What does this text mean?' were systematically excluded from linguistic analysis. Their solution was presupposed and, thus, left to the self-evidences of everyday experience. According to Pêcheux, it was precisely this 'leaving uncovered the territory' by linguistics, without its reoccupation by another science, that allowed ideologies to invade that terrain (again). In other words, though linguistics had established itself as a science through a 'Saussurean' epistemological break, it had 'forgotten' to develop an adequate theory of meaning production in discourse.

Within the 'post-structuralism' of that time, Pêcheux's project was unique for two reasons that are intimately related: he wanted to establish a *social scientific* theory of 'discourse', and he emphasized the need of developing an *instrument* for creating experimental (vs. experiential) results (cf. Chapter 2). His emphasis on the need of developing an empirical alternative to linguistic 'speculation', with its implied stress on both theoretical and analytic rigor, induced Pêcheux to a much more intensive study of (Saussurean) linguistics than was common among his fellow-philosophers. This study was necessary in order to find theoretical room within linguistics for the concepts of 'meaning' and 'discourse', and in order to construct a scientific instrument for the analysis of these objects. Pêcheux's 1969 book on automatic discourse analysis (Chapter 5 below) presents the results of this study.

---

<sup>1</sup> Especially Derrida, in much of his work, comes close to a reinvention and radicalization of tendencies that can already be found in Saussure's courses.

### **Automatic discourse analysis**

The ADA69 book starts from a criticism of traditional forms of content analysis and textual analysis (see pp. 63-71 below). Such analyses presuppose a subject (the analyst or his coders) capable of 'reading' the meaning of a text. Pêcheux wanted precisely to avoid this reliance on the reading subject, because it would inevitably produce an ideological reading. However, it must be acknowledged that the role of 'intuitive reading' was a matter of concern for content analysts as well. In the same year 1969, for instance, Krippendorff stated that

traditional content analysis may be said to have presented a technique for reliably *intuiting content* rather than analyzing it. However, when modern content analysts take the word 'analysis' in its literal sense and use, for example, computers for compatible interpretation of text, no part of the procedure can be delegated to the inexplicable process of intuition. [...] The explication of intuitively obvious semantic interpretations and judgments constitutes the most formidable obstacle in computer applications. (Krippendorff 1969: 6)

Though this problematization of intuitive reading was quite similar to Pêcheux's criticism, the solutions proposed were very different. One solution in content analysis was the construction of dictionaries in which entry words were defined with one or more 'tags' representing categories in the investigator's theory. The prime example of this approach is *The General Inquirer* (Stone et al. 1966). A different solution was to restrict the analysis to formal aspects of the studied texts only, i.e. to those kinds of analysis that do not presuppose interpretation. This approach resulted in a proliferation of lexicometric studies. These different solutions have in common that they are blind to the question that is primary to Pêcheux: the question of how 'meaning' and the 'subject' are produced in the discourse.

Central to Pêcheux's approach is the concept of *conditions of production of discourse*. Taking Jakobson's model of communication as a starting point, he 'sociologizes' this model by requiring that the two subject positions in the model – the position of speaker/writer and that of listener/reader – be interpreted as locally and temporally specific *imaginary positions*. What matters is the place that each of them attributes to

itself, to the other and to the 'referent' (the object of which they speak). Such positions are imaginary, not in the sense of being 'unreal' but of being related to *images* which produce material effects. The protagonists are not 'free' in the choice of these images, which depend on structural relations (such as between worker and boss) and on what is said earlier and/or elsewhere. Such restrictions account for the relative stability of discourse through different occasions. This implies that 'meaning' is more or less stable through such occasions, but that it changes when the conditions of production change, which is the case, e.g., when the same speaker speaks to another person or to the same person on a different subject.

Whereas Pêcheux's theory of the conditions of production of discourse may appear as a more or less straightforward application of Althusserian concepts to Jakobson's model, his next step is more daring. This step consists of reintroducing the Saussurean theory of value, which explains the meaning of words by their relations to all other words of the language [*langue*], and applying it to the conditions of production of discourse. The *meaning* of words in a discourse (i.e. in a text or utterance) is explained by their relations to other words that are *not* said: words that could have been said but were not, words that were said previously (either on the same occasion or on other occasions), and words that *could not* be said. This interrelatedness between words is what Pêcheux calls 'metaphoric' relations, and their meaning-effect is called a *metaphoric effect* (see pp. 96-100 below). He outlines a theory of meaning as an effect of metaphoric relations (of selection and substitution) which are specific for (the conditions of production of) an utterance or a text. This theory of meaning is the bedrock of his instrument for (automatic) discourse analysis.

Before outlining the main features of Pêcheux's automatic discourse analysis, it is to be recalled that it was his aim to develop an instrument that would produce *experimental* results (i.e. results which are the product of a theory-driven practice) as opposed to *experiential* results (i.e. results which are based on everyday experience). In order to study the meaning of discourses, he had to construct an instrument that required discourses (not meanings) as its input, and that would have information about the meaning of those discourses as its output. In other words, the instrument must construct metaphoric relations without the analyst 'feeding' it with information about the experiential meaning of the words that build up these discourses.

A method of discourse analysis that was outlined by the American

linguist Zellig Harris (1952) provided Pêcheux with an instrument capable of doing exactly this: "We were fascinated by Harris, because we felt that he could offer something that would allow to escape both merely intuitionist hermeneutic positions and positivistic 'lexicometric' positions" (Pêcheux, interviewed in Woetzel and Geier 1982: 389). It was Harris's aim to extend descriptive linguistics beyond the limits of the sentence. He therefore proposed to focus on a sequence of sentences or a *text*, and to conduct a formal analysis in terms of recurring patterns among its constituent elements. All assumptions about pre-given meanings of the elements, as well as all references to discursive sequences beyond the text in question, were excluded (cf. Thompson 1983: 217). Precisely because of this exclusion of pre-given meanings, Harris's method of discourse analysis fitted Pêcheux's need of a *formal* instrument. But, contrary to Harris, Pêcheux did not exclude discursive sequences occurring in other texts. Because it was Pêcheux's aim to construct a field of metaphors, i.e. of words that *could* have been in the discursive sequence under analysis but were *not*, he used the formal discourse analysis of Harris to relate sequences within a *corpus* of texts instead of sequences within one text.

In order to guarantee the formal character of Harris's method, i.e. to prevent the tacit (re)introduction of 'self-evident' meanings, it had to be made 'automatic'. But in what sense is ADA69 'automatic'? In fact, it is not 'automatic' at all. Or, rather, it is as 'automatic' as, say, instruments in astronomy and in physics. Like such instruments, automatic discourse analysis requires a very specific kind of input (which must be constructed as an input for the instrument) and it produces results that have meaning only within a very specific theoretical frame. The input of ADA69 is constructed in two subsequent steps, which are called the phase of *corpus construction* and the phase of *linguistic analysis* respectively. The third step then is *automatic discourse analysis* proper. But the output of this third phase does not consist of *findings*. The 'findings' of an ADA69 procedure become available only through an *interpretation* of the results of this third phase. It is precisely this requirement of theoretical intervention, before and after the use of the 'automatic' instrument (which itself is defined in terms of the theory), which allows for a scientific (re)occupation of the field of semantics.

The aim of automatic discourse analysis is to provide the analyst with a metaphoric matrix that gives information about the production of meanings under theoretically specified conditions of production. The first phase of ADA69, that of corpus construction, consists of delineating the



object of study ('Which conditions of production are to be studied in this analysis?'), and of selecting the set of texts or utterances that, according to the theory, represent those conditions. This set of texts is called the *corpus*. The metaphoric matrix must be constructed from the elements (words) that constitute this corpus. The second phase of ADA69, the phase of linguistic analysis, consists of rewriting all sentences of the corpus in a standard format, which is required for their being used as input for the phase of discourse analysis. This rewriting is called 'linguistic analysis' because it consists of a form of syntactic parsing and because this parsing is done according to a linguistic theory. The internal cohesion of the separate texts of the corpus is preserved in this phase by attaching symbols to pairs of sentences which represent the connections between these sentences.

The 'linguistic analysis' phase of ADA69 is an application of the 'Saussurean' idea that the language [*langue*] is a system that is shared by a community (a nation or a culture) and that the theory of such a language, linguistic theory, can be treated as a neutral, formal device in an instrument of discourse analysis. This instrument produces metaphoric matrices from a standardized input, irrespective of the specific kind of linguistic analysis that has produced that input. There is no *theoretical* preference for a specific type of linguistic analysis. The theory of ADA69 requires *that* the input to the third – discourse analytic – phase has a standard format, which forms the basis for the construction of metaphoric matrices in this phase. But there is no discourse theoretical reason for the choice of a specific format. This choice is a linguistic one. In *practical* terms, however, the technical details of the phase of discourse analysis proper are dependent on the format of its input. And, more important, the resulting metaphoric matrices, called *semantic domains*, are constructed according to substitutions of words in places (within that format) that are defined by the linguistic theory underlying this format. The choice of linguistic theory, thus, has practical consequences. In other words, although different linguistic formats (resulting from different linguistic theories) require technically different treatments in the phase of discourse analysis which *can* result in partially different outcomes, ADA69 does not provide criteria

that justify the choice of a specific linguistic format.<sup>1</sup>

The term 'automatic', in sum, refers to the formalized character of Pêcheux's method. There are no 'automatic' findings. ADA69 produces outcomes in a formal way, but they remain simply outcomes, 'data' that must be interpreted. In his 1969 book, Pêcheux only mentions this final phase of interpretation of the results (see p. 118 below, where he introduces the concept of 'reading'). Later, the difficulties that Pêcheux encountered in this phase inspired him to make revisions of his theory (which are documented in this volume as well, pp. 175-183 below). The revisions occasioned by difficulties in interpreting the results of ADA69 are an excellent example of the Bachelardian theory regarding the role that theoretically constructed instruments play in the development of a science.

### Later developments

Pêcheux's 1969 book was a *theoretical* book. It presented a theory of discourse and meaning production, and outlined an instrument based on this theory. But the ADA69 book did not present results of the proposed procedure, because the instrument was not 'built' yet.

In the following years, Pêcheux constructed such an instrument. This required further work in two respects. On the one hand, the algorithms that could produce the metaphoric matrix in the phase of discourse analysis were implemented in computer programs (see the Appendix on formalization, pp. 57-59 below). On the other hand, the linguistic phase was refined in a 'manual' (Haroche and Pêcheux 1972a). After the instrument that was proposed in ADA69 had been actually built, the first empirical work with it was done (Gayot and Pêcheux 1971, Pêcheux and Wesselius 1973, Pêcheux et al. 1979). In analyzing pamphlets from the 1968 student movement (Pêcheux and Wesselius 1973) in particular, Pêcheux encountered new problems. At the same time, ADA69 was discussed extensively in the linguistic literature (Provost-Chauveau 1970, Trognon 1972, Fisher and Veron 1973). Both these criticisms and his own experiences in using the instrument resulted in a revision of the theory,

---

<sup>1</sup> Yet, Pêcheux justifies his choice for a specific standard format extensively. His discussion is represented partially in this volume (see pp. 110-113 below). The details of the discussion of 1969 are superseded by later developments within linguistics, among them developments towards automatic parsing that are treated in Chapter 8.

and in changes in the procedures of automatic discourse analysis.

In this period (1969-1975), Pêcheux was working on *Les Vérités de La Palice* (Pêcheux 1975). This book on language, semantics and ideology, was published in the same year as an article in which the revisions of ADA69 are discussed (Pêcheux and Fuchs 1975, henceforth ADA75; Chapter 6 in this volume). These publications were directed to different audiences. The book can be seen as a sequel to the two Herbert articles from the 1960s. It discussed problems of semantics, ideology and Marxist philosophy for a philosophically oriented audience, whereas the article, which discussed the same theory in relation to technical matters of discourse analysis, was written for an audience of empirical researchers. Yet, the two publications presuppose one another: results of automatic discourse analysis are used as examples in the book, which functions in its turn as a theoretical background for the discussion in the ADA75 article.

In this period of 'thematization of uneven entanglement' (cf. pp. 237-238 below), Pêcheux did not radically change the actual procedures of automatic discourse analysis. What changed were the theoretical underpinnings of the whole enterprise and, correspondingly, the interpretation of its results: Pêcheux now explicitly introduced the Althusserian theory of ideology and Lacanian psychoanalysis into his theory, which allowed him to theorize asymmetries *between* discourses. Moreover, he explicitly addressed the problem of the interpretation of the results of automatic discourse analysis (i.e. of metaphoric matrices) by distinguishing between different effects of metaphoric relations: they may be either relations of synonymy or 'oriented' relations (see p. 165 below). In either of the 1975 publications, Pêcheux reconsidered the relationship between ideology and language, and – as a consequence – the relationship between discourse analysis and linguistics. Because ADA69 was based on the assumption of stable conditions of production, it could not theoretically account for 'ideological struggle' in terms of discourse. This problem emerged in interpreting the results of the ADA69 procedure. In 1969, Pêcheux's conception of the metaphoric matrix had included only relations of synonymy (e.g. the president/De Gaulle) and opposition (e.g. De Gaulle/the workers), but the produced matrices contained relations between elements that could not be interpreted that way (e.g. strike/the workers). The latter kind of relation is now defined as being *oriented*: one term is the origin, the source, or the argument of the other (e.g. the workers begin a strike). This relation is not an explicit part of the discourse studied, but tacitly presupposed by it. In other words, we see here the influence of

another discourse within the one studied. By theorizing this influence Pêcheux introduced a true innovation into the theory of discourse: his theory of *interdiscourse*, defined as the 'complex whole in dominance' of conditions of production of discourse.<sup>1</sup>

Neither ADA69 nor ADA75 were designed for studying the workings of interdiscourse, which was constructed as something external to the (stable) conditions of production that were hypothesized. After 1975 it became increasingly clear to Pêcheux that this was a serious drawback of the theory underpinning ADA69. This brought him back to the study of linguistics (Gadet and Pêcheux 1981) and of the analytical tradition of Gottlob Frege and Ludwig Wittgenstein (Pêcheux 1984).<sup>2</sup> These studies resulted in a serious criticism of extant linguistic theories of grammar and semantics. In his last, unfinished article, for instance, Pêcheux discusses differences in meaning-effect between sentences that, according to transformational-generative grammar, have the same deep structure (Gadet, Léon and Pêcheux 1984). It appeared not to be possible to adapt ADA69 to the new theoretical circumstances. At the beginning of the 1980s, Pêcheux virtually abandoned it.<sup>3</sup> This abandonment, however, does not imply a rejection of the possibility of revising automatic discourse analysis according to new theoretical standards. Pêcheux stuck to the Bachelardian idea that the construction of instruments and the production of experimental results are necessary to the development of a science. He was, therefore, very much interested in experimenting with recently constructed software (DÉREDEC) as a tool for parsing French sentences (through a recognition grammar) and in ways of using the results of this linguistic analysis as an input to other (discourse analytic) procedures.<sup>4</sup>

---

<sup>1</sup> Cf. Pêcheux (1982a: 113-118 and 184-193). See on this book Houdebine (1976), Woods (1977), MacCabe (1979), Cousins (1985), and Macdonell (1986: 43-59).

<sup>2</sup> He also started to reread Friedrich Nietzsche (cf. p. 34, note 1).

<sup>3</sup> This process is documented in Chapters 7 and 9 of this volume.

<sup>4</sup> Pêcheux tragically died in 1983 before having written about how DÉREDEC could be used within his theory. Chapter 8 below is included in this volume in order to give a glimpse of its possible uses for discourse analysis.

## **Conclusion**

Many readers of *Language, Semantics and Ideology* (1982), Pêcheux's only book available in English to date, could not know much of Pêcheux's work on discourse analysis. Because of its epistemological problematizations, this book was usually read as the work of a discourse theorist who misleadingly drew impracticable consequences from merely theoretical premises. It was rarely recognized that Pêcheux's theorizing was accompanied by the construction of an instrument and the production of experimental results. The presentation of ADA69 in this volume, therefore, is not only relevant as a contribution to the history of French discourse theory and discourse analysis. It shows, in an exemplary manner, the nature of the obstacles, both in terms of theory and of method, encountered in any form of discourse analysis. The contributions to this volume show how the transition towards a sophisticated theory of discourse and a scientific method of discourse analysis is a matter of both forcing an 'epistemological break' with empiricist conceptions of meaning and of producing the corresponding experimental results.

Pêcheux's texts are presented with extensive annotations and introductions, written by former colleagues and co-workers. In Chapter 2, Paul Henry scrutinizes the ways in which epistemological work pervades Pêcheux's approach to automatic discourse analysis. In Chapter 3, Françoise Gadet, Jacqueline Léon, Denise Maldidier and Michel Plon argue that, at the end of the 1960s, Pêcheux's problematization of the 'subject', and his occasional reference to the unconscious, are to be read in a sociological (or Marxist) rather than in a psychoanalytic (Lacanian) frame of reference. In Chapter 4, they discuss the decisive influence of Ferdinand de Saussure, Zellig Harris, Noam Chomsky, Roman Jakobson, Émile Benveniste, and Antoine Culioli on Pêcheux's work. An Appendix to this introductory part of the volume provides background information on the implementation of the ADA69 algorithms in operating computer programs.

The second part forms the core of this volume. In Chapter 5, Pêcheux's book on ADA69 is presented. After having presented a criticism of methods of content analysis, Pêcheux outlines his 'conceptual guidelines for a theory of discourse', followed by a description of both the linguistic phase and the phase of automatic discourse analysis proper. The chapter is concluded with a discussion of potential applications of ADA69. Chapter 6 presents Pêcheux's critical elaboration of ADA69 published in 1975

(ADA75). After having discussed how the Althusserian theory of ideology and Lacanian psychoanalysis fit in his theory of discourse, Pêcheux deals extensively with the criticisms of ADA69.

The volume concludes with a third part consisting of three chapters that document the theoretical and practical development of automatic discourse analysis after 1975. Chapter 7 presents an application of ADA69 to a corpus of three proposals to the 1979 congress of the French Socialist Party. This chapter illustrates the ADA69 procedure and presents Pêcheux's evaluation of ADA69 in the beginning of the 1980s. Chapter 8 discusses the DÉREDEC software. Pêcheux believed that DÉREDEC would allow a different formalization of the phase of linguistic analysis. This chapter demonstrates how Pêcheux's aim of providing a scientific instrument for the study of discourse and ideology can be achieved in different ways, depending on developments in related fields such as computer linguistics. Finally, Chapter 9, written in 1983, gives Pêcheux's retrospective overview of his work in automatic discourse analysis. It evaluates the earlier stages of automatic discourse analysis in the light of the questions he deemed relevant for future development.

We expect that the availability of these texts in English will bring about a renewed confrontation with the problem of how to develop a *scientific* practice of discourse analysis – a problem acutely defined by Pêcheux in the 1960s and 1970s and not satisfactorily addressed to date.

## Theoretical issues in Pêcheux's *Automatic Discourse Analysis* (1969)

Paul Henry

In 1966, a paper entitled 'Réflexions sur la situation théorique des sciences sociales et, spécialement, de la psychologie sociale' (Herbert 1966) appeared in *Cahiers pour l'analyse*, the journal of the *Cercle d'Épistémologie* of the *École Normale Supérieure* in Paris. The paper was signed 'Thomas Herbert', but was in fact Michel Pêcheux's first publication. During the summer of 1968, a second paper, 'Remarques pour une théorie générale des idéologies' (Herbert 1968), was published under the same assumed name. Between writing the two 'Thomas Herbert' papers, Pêcheux published two articles on discourse analysis under his own name (Pêcheux 1967, 1968). At first sight, there is no clear or obvious connection between the papers signed 'Thomas Herbert' and those on discourse analysis. Similarly, if we look at *Automatic Discourse Analysis*, which appeared in 1969, we might be forgiven for thinking that Michel Pêcheux and Thomas Herbert were indeed two different people with quite different preoccupations and backgrounds.

As a matter of fact, the key concepts and notions to be found in Thomas Herbert's papers, which make explicit reference to 'historical materialism' and to psychoanalysis, are almost completely absent from Michel Pêcheux's book on automatic discourse analysis. The book contains only one reference to the 'theory of ideology' and the 'theory of the unconscious' (p. 118 below), and it appears in a footnote where Pêcheux merely says that the theory of discourse (as he conceives it) cannot take the place of those theories, but can intervene in their field. Similarly, the criticisms addressed to the social sciences, and to social psychology in particular, in Herbert's first paper do not really appear in the book. It must be recalled that the book was published in a series devoted to the behavioural sciences and edited by two renowned psychologists, and that its

content was first presented as a doctoral thesis in social psychology. It might, then, be thought that Pêcheux used a pen name and masked the undercurrents of his thought in his academic publications for purely opportunist reasons, or in order to avoid the problems that a clear and explicit statement of his real theoretical position – which was outside mainstream academic psychology – might have caused him in his academic career. In fact, however, Pêcheux's attitude was an expression of a carefully calculated strategy.

Pêcheux had always been anxious to intervene within the field of the social sciences, and within that of social psychology in particular. At the time of the publication of *Automatic Discourse Analysis*, he said that this was his main professional concern. In making his intervention, he relied upon two theoretical discourses which had, it seemed to him, already overthrown the idealist problematic which dominated the social sciences: historical materialism, as reformulated by Louis Althusser, and psychoanalysis, as reformulated by Jacques Lacan's 'return to Freud'.<sup>1</sup> He further relied upon certain aspects of the broad movement known, somewhat ambiguously, as *structuralism*. The late 1960s marked the apotheosis of structuralism in France. The common denominator between Althusser and Lacan had something to do with structuralism, even though neither of them could really be termed a structuralist. What interested Pêcheux in structuralism was that element within it that implied a non-reductionist attitude towards language. Later we will see why (pp. 32-35).

### A 'scientific' instrument

Pêcheux's first publication concerned the 'theoretical situation' within the social sciences. We will not attempt to give a complete account of that text here. This publication is, however, of fundamental importance if we wish to understand the precise nature of Pêcheux's aim in developing an

---

<sup>1</sup> Althusser's 'Freud and Lacan' appeared in the December 1964-January 1965 issue of *La Nouvelle Critique*, the theoretical journal of the *Parti Communiste Français* (PCF) (see Althusser 1971b). Its appearance signals the end of the PCF's official hostility to psychoanalysis. The membership of the *Cercle d'Épistémologie* at the *École Normale Supérieure* (which Pêcheux joined under the name 'Thomas Herbert') included both Marxists who were close to the PCF (if not members of it) and philosophers who had been greatly influenced by Lacan.



automatic discourse analysis: to provide the social sciences with a *scientific* instrument, which they urgently needed, an instrument which would be a counterpart to a theoretical and scientific breakthrough in their field. For Pêcheux, this implied that:

- (1) the state of the social sciences was in some sense pre-scientific;
- (2) the establishment of a science implies instruments.

The first point relates to Pêcheux's critique of the state of the social sciences at the moment he was writing. But it is also linked to the second. Here, we see Pêcheux's interest in and commitment to the field of epistemology and the history of the sciences.<sup>1</sup> He states that two mistakes are to be avoided: "regarding any use that is made of instruments as scientific, and forgetting the role of instruments in scientific practice" (Herbert 1966: 161). The paper in fact develops a precise analysis of what a scientific instrument is, and it is on the basis of this conception that Pêcheux develops his system of automatic discourse analysis.

What, according to Pêcheux, is a scientific instrument? His point of view is based primarily on the history of science and technology. He follows the analyses of Bachelard and Canguilhem closely,<sup>2</sup> but he also introduces elements derived from a Marxist analysis of the consequences of the division of labour (especially of the division between intellectual and manual labour) and of the contradictory character of the combination of forces of production and social relations of production in a class-divided society.

At the beginning of Herbert's second paper we find a résumé of the 'findings' of the earlier paper. Two fundamental points are made in a very condensed form. The first concerns the conditions under which a science establishes its field. The second concerns the 'methodical reproduction' of its object, or the process whereby a science undertakes an internal investigation of its own discourse and tests its consistency and

---

<sup>1</sup> Pêcheux worked particularly on the historical development of scientific theories of magnetism (cf. Pêcheux 1969c).

<sup>2</sup> Bachelard and Canguilhem were the leading proponents in France of a non-positivist and non-empiricist approach to epistemology and to the history and philosophy of science (as opposed to the approach taken by Duhem, for example). They argue that epistemology cannot be divorced from the history of science, and argue against the view that sciences advance in a continuous way by stressing the importance of discontinuities (see Canguilhem 1977). Foucault adopts a similar approach, albeit within a rather different perspective (cf. Lecourt 1972).

necessity.

(1) "*Every science*", writes Pêcheux-Herbert, "is produced by a conceptual mutation within an ideological conceptual field; by distancing itself from that field, it simultaneously acquires a knowledge of its previous odyssey and a guarantee of its own scientificity. In that sense, any science is initially the science of the ideology with which it breaks" (Herbert 1968: 74). The object of a science is therefore not an empirical object but a construct. Such an object cannot emerge from an almost random investigation of nature which gradually delineates the object and makes both it and its features visible.

(2) Within the history of any science a distinction should be made between two moments: the moment of the transformation that produces its object (which is dominated by the theoretical-conceptual work of subverting the ideological discourse with which it breaks), and the moment of the 'methodical reproduction' of its object (which is both conceptual *and* experimental).

Instruments and tools play different roles in these two moments, or phases, within the history of a science. This point is made mainly in the first of the two Herbert papers. The first moment has been described as being essentially theoretical and conceptual, but this does not mean that instruments or tools (which can be 'material' and/or 'abstract') have no role to play in it. It is, however, in the second phase – of the 'methodical reproduction' of the object – that instruments appear to have a decisive function. But that function cannot come into play until the transformation that produces the object has taken place. The founding moment of a science is also the moment of the re-invention of the tools and instruments it needs, and it finds them where it can: in other established scientific practices, but also in 'technical practices', that is, in practices bound up with the process of production. Pêcheux gives several instances of tools and instruments being developed within 'technical practices' long before they were transferred to 'scientific practices'. Examples include stills, scales and fieldglasses. Scales, for instance, were in use in commercial transactions long before they became scientific instruments. With Galileo, the theory of scales became part of the theory of physics. The principles which explain why scales give invariant measures (and the limits within which they do so) are part of Galileo's theory. A homogeneity is thus established between the object of physics and its methods, which truly establishes physics as a science. Should the use of scales give inconsistent results, this would have direct theoretical implications; it would be neces-

sary to revise or transform *well-defined* aspects of the theory. Conversely, any development within the theory of physics can be 'translated' into its methods and instruments (including its mathematical instruments). This process is precisely what Pêcheux terms the 'methodical reproduction' of the object of a science, that is to say the process in which a science creates its own *Spielraum*, varies its questions and, in doing so, internally adjusts its theoretical discourse, making it both consistent and necessary. Established sciences, of course, develop their instruments themselves; the 'invention' of these instruments is internal to them and is a kind of 'realization' of theory. But, writes Pêcheux, whenever an instrument or experimental device is transferred from one branch of a science to another, or from one science to another (and this often occurs), it has to be somehow reinvented in order to become an instrument or device belonging to that science or branch of science. Pêcheux concludes by stating that sciences raise questions by using instruments, and that the internal adjustment of a scientific discourse to itself therefore consists in the theoretical appropriation of instruments. It is this that makes scientific activity a *practice* (Herbert 1966: 161-162).

We have now a fairly clear impression of what Pêcheux understands by the term 'scientific instrument', and of what he wants his system of automatic discourse analysis to be. In his view, such an instrument could not be conceived independently of a theory capable of encompassing or producing a theory of that instrument. Therefore, the elements needed to construct that instrument should in a sense be 're-invented' or 'appropriated' by the theory in question. This is particularly true of Pêcheux's borrowings from linguistics. Automatic discourse analysis cannot be simply an 'application' of linguistic analysis. At the beginning of his book he criticizes, therefore, the application of linguistics to 'textual analysis'. The same holds for any borrowings from logic, computer studies and so on. This implies that automatic discourse analysis cannot be simply one of the many instruments already used by social scientists, or an instrument which completes an extant set because of its ability to perform tasks other instruments cannot perform.

Pêcheux was a philosopher by training, but for reasons which were deeply rooted in his personal history and family background, he was a philosopher who was fascinated by machines, instruments and technologies. Nor was he just any philosopher. He was a philosopher who was convinced that the classical practice of philosophy was meaningless, or at best a failure, especially where the sciences were concerned. In referring to

the classical practice of philosophy with regard to the sciences, I refer to a practice of philosophy which attempts to legislate in matters of science, scientificity, epistemological legitimacy and the like. Pêcheux was convinced that a mere philosophical critique of the social sciences would not get us very far, but he was also convinced that the social sciences are not sciences at all, but merely ideologies. From his point of view, the only valid critique of such ideologies was that made by sciences whose field they currently occupied. This is precisely what he means when he says that a science is initially the science of the ideology with which it breaks. In this way, Pêcheux adopts a philosophical position (in the tradition of Bachelard, Canguilhem and Althusser). Whilst he was critical of the traditional philosophical approach to the sciences ("Let us leave Kant at his tribunal", Herbert 1966: 144), he was by no means prepared to accept that scientific practices could be independent of philosophical practice.

On the contrary, a new type of philosophical practice is absolutely necessary because the traditional practice of philosophy has, among other things, played a crucial part in the elaboration of the social sciences which he regarded as ideologies or pseudo-sciences. On the other hand, he was also convinced, as we have seen, that scientific practices need instruments ('material' or 'abstract'), even though the use of instruments is no guarantee of the scientificity of a practice. Defining a new instrument was, for Pêcheux, the best way of avoiding the traps of a traditional philosophical critique. It was also, in his view, the only possible way of gaining a hearing from social scientists, who have always (and sometimes with good reason) dismissed the philosophical criticisms addressed to them. Pêcheux was engaged in a debate with both philosophers and social scientists but, given the current state of research (which was largely an effect of academic divisions of intellectual labour), those two groups of interlocutors clearly required a different approach.

We can, then, understand why Pêcheux stressed the role of instruments when he addressed himself to social scientists. He felt that if he foregrounded the theoretical and philosophical aspects of his project, his debate with the social scientists would be focused on such issues, and his reference to instruments would seem to be no more than an illustration of his view. That would have been in complete contradiction with his conception of the scientific instrument. When, on the other hand, he addressed himself to philosophers, as he did in *Cahiers pour l'analyse*, he merely mentioned that, if an ideological mutation was to be provoked (that is, effects that could not be contained in the dominant ideology), it would be

necessary to construct an instrument precisely in the ideological space in question. This is what he does at the end of the second 'Herbert' paper. And he adds that one cannot initiate a speculative dialogue with just any interlocutor, and that one cannot experiment in just any condition (Herbert 1968: 92).

Pêcheux's criticisms of the use made of instruments in the social sciences make a crucial point. Whilst he did regard his automatic discourse analysis as an instrument, it was in no sense comparable to the extant instruments in the social sciences. He did not, however, simply denounce the way those instruments were used: he also tried to explain why it was possible for them to be used that way, and how one particular view of instruments had come to dominate the field. His criticism of the use made of instruments in the social sciences blends with his criticism of the social sciences themselves, and of their links with politics.

### **Instruments in the social sciences**

In his first paper, Pêcheux criticizes those conceptions of 'scientific practice' which see it merely as an extension of 'technical practices'. Such a traditional view of scientific practices fits in with the epistemology of empiricism, but fails to establish a distinction between scientific practices and other types of practice involving theorization, speculation and the use of instruments. It fails, for instance, to identify clearly the difference between alchemy and chemistry (a point which Pêcheux develops as an illustration). To return to the example of scales: quite aside from the commercial uses to which they were put, scales were used to weigh all sorts of things, such as blood, urine, wool and atmospheric air. Almost everything which could be weighed was weighed; comparisons were made, and whole theories were elaborated on the basis of these observations. But while scales were being used in this haphazard way, there was no 're-invention' of the instrument, no theoretical 'appropriation' of the instrument. Scales were regarded simply as instruments which gave 'objective' data concerning reality, and it was assumed that it was possible to speculate and to draw conclusions on the basis of that data. This 'objectivity' was in fact no more than a transposition of the adequacy of the instruments that was apparent in the commercial practices within which they had been developed and used.

In a sense, scales are a by-product of, among other things, com-

mercial practices but at the same time they open up the possibility of such practices. 'Technical practices' are, says Pêcheux, *determined* in the sense that they respond to an external demand, and *determinant* in the sense that it is the set of possibilities they open up which makes it possible for that demand to exist. From the point of view of 'technical practices', the only requirement to be fulfilled was that the repeated use of scales should give constant results; if two quantities of the same commodity were weighed separately and then together, the latter result should be equal to the sum of the first two results, and so on. The price of twice the amount of a commodity could, for instance, then be legitimately stated to be twice the price of the first amount; given that a certain weight of gold corresponded to a unit of weight of some commodity, a correspondence could be established between any given weight of that commodity and a weight of gold. A system for measuring the quantities of commodities involved in commercial transactions was thus created by reference to weight. Though a whole technology of scales was developed, it did not embody any real theory of scales or of additivity. The properties of scales were received facts verified empirically. The same was true of the established weights of blood or urine.

Because the measurements obtained were sufficiently reliable, people felt free to speculate on the basis of their findings. So, why should one not feel free, for instance, to weigh brains and state that the weight of the brain is a measure of intelligence? That assumption was in fact made on the basis of theories which saw the brain as the organ of thought and intelligence. Certain physical anthropologists began to determine the average weight of the brain in various human races, and to relate its weight to the assumed level of intellectual ability of those races as related to their relative distance from the animal kingdom. Other experiments were of course more sophisticated, and their findings cannot be dismissed so easily. But when Binet stated that his IQ-tests provided a measure of intelligence, was he so far removed from the former view? We are dealing here with what Canguilhem (1977) terms a (pre-)scientific ideology. Canguilhem, who refers specifically to the field of the life sciences, in which most of his work was done, describes such an ideology as a discourse which, given that it is not possible to obtain any real experimental verification of its claims within its own field, establishes its credibility on the basis of a maximal number of analogies with data established in other fields.

This illustration of a specific (but common) ideological use of

tools and instruments calls for two comments.

(1) Such a technical use of instruments is clearly an extension of the other uses to which they are put. This does not imply, however, that scientific practices can be placed in a continuum with technical practices. It is of course true that not everything about technical practices is false; the measurements obtained are not inaccurate, and comparison based upon them cannot be regarded as being entirely unfounded. It cannot be said that they represent an absence of knowledge. But this does not mean that scientific use of tools can be placed in a continuum with technical use, as in the rationalizations made by empiricist epistemologies and philosophies of knowledge.

(2) At first sight, the technical use of instruments appears to be quite divorced from common social demands relating to the sphere of production (such as those involved in commercial transactions). But from another point of view, it can be seen to be bound up with a different kind of social demand and command. This is perfectly clear in the case, though this example may seem rather schematic or simplistic, where experiments in weighing brains are used to legitimize blatantly racist positions. It is of course possible to find examples of anthropologists using the same instrument to opposite ends. The important point is that the use of instruments is directly related to the defence of ideological positions (or to attacks on ideological positions) and to interventions within the field of ideological struggle. This implies that (a) we cannot dismiss the use of an instrument on the grounds that it always has the same ideological or political orientation, and that (b) the social demand or command, which seemed to have disappeared to make way for more noble or disinterested activities (obtaining knowledge of the world), is in fact still present. The two 'Herbert' papers suggest that this whole process of divorcing the use of instruments from social demands (its conditions of possibility) has something to do with the division between workers and non-workers in a class society. In that sense, they begin to outline an account of the historical roots of empiricist epistemologies and philosophies of knowledge.

In the second 'Herbert' paper, Pêcheux analyses ideology as a dual form (Herbert 1968: 77):

(1) In terms of the process of production, ideology is, according to Pêcheux, a primary process in which technical and operational concepts originating in the labour process are divorced from their operational sequence and recombined within an original discourse;

(2) In terms of the social relations of production, ideology is a

process which produces and maintains differences essential to the workings of social relations of production within a class society (primarily the fundamental 'difference' between worker and non-worker). At this level, the function of ideology is to make the agents of production 'recognize' their place within the relations of production.

The 'social sciences' developed in societies in which the aim of political practice is primarily, according to Pêcheux, to transform social relations in such a way that the overall structure of social practice remains unchanged. The 'social sciences' are a direct extension of ideologies developed in relation to political practice (1966: 156). In their present state, he goes on, they consist in the application of a technique to an ideology of social relations and they are designed to adapt and re-adapt social relations to an overall social practice, regarded as the invariant of the system (1966: 156-157). Pêcheux's discussion of political practice brings us back to discourse analysis. He states that the instrument of political practice is *discourse* or, to be more accurate, that the function of political practice is to transform social relations by re-formulating social demands through discourse (1966: 150).<sup>1</sup>

### **Language, discourse and ideology**

We can see, then, that in Pêcheux's view the 'social sciences' are primarily techniques which are closely bound up with political practice and with the ideologies which have developed alongside that practice, and that their instrument is discourse. This idea reappears in the second 'Herbert' paper. If, writes Pêcheux (1968: 81), man is regarded as an animal who communicates with his 'likes', we can never understand why it is through the general form of discourse that dissymmetries and differences are established between the agents of the system of production. We can now understand why, in his attempt to intervene within the ideological field of the 'social sciences', Pêcheux chooses discourse and discourse analysis (the elaboration of a theory of discourse and the construction of the

---

<sup>1</sup> For Pêcheux this does not mean that politics is purely a matter of discourse. As an illustration, he analyses the case of Law and juridical practice which simultaneously try to rationalize 'customary law' and to realize 'the rational essence of Law'. Juridical practice attempts to transform something that exists 'in nature' into something that exists 'in reason' (Herbert 1966: 150). This transformation is a re-formulation, and it implies discourse.



instrumental device of automatic discourse analysis) as the specific site for both a theoretical and a practical intervention. There are two reasons for doing so:

- (1) the hidden link between the 'social sciences' and political practice (the term 'social sciences' refers primarily to social psychology and sociology but also to psychology itself, even though it is usually regarded as a 'human science' rather than a 'social science');
- (2) the link between political practice and discourse.

Pêcheux completely rejects the conception of language which reduces language to an exchange of 'information', or which sees it as an instrument for the inter-human communication of significations which exist in themselves and which can be defined independently of language. That theory of language is, in his view, an ideology whose function within the 'human and social sciences' (where it is dominant) is to mask the link between these sciences and political practice, to blur that link and, at the same time, to place those 'sciences' in a continuum with the natural sciences. Even though animals are not endowed with language in the strict sense of that term, they do communicate. If, therefore, language is seen as an instrument of communication (albeit an elaborate, complex and efficient instrument), men and human societies can be equated with animals and animal societies. Although dissymmetries and differences between the agents of the system of production are established through the general form of discourse, this is not openly effected by issuing the order: 'Stand there, this is your place in the system of production'. It is, that is, not a matter of 'communication' reinforced if necessary by some form of physical coercion or threat. Coercion may exist, and in a sense it does always exist. Someone may, for example, be ordered to a specific work-station, but a work-station is not a place within a system of production. That is not the point. The point is that we have to understand how the agents of the system spontaneously recognize their place without being formally ordered to adopt a place and even without 'knowing' that they have a place within the system of production. When someone is ordered to a specific work-station, he or she has already recognized his or her place in the system; the individual already knows that he or she is a worker, with its implications. The same applies when someone is appointed to the bench, crowned king, or even when someone decides to become a writer. The process whereby individuals are, so to speak, 'put in their place' remains invisible; we see only its external features and effects. In order to understand this process, which simultaneously 'places' individuals and

conceals the way in which they are 'placed', and the role language plays in that process, we have to abandon the communication view of language. This does not mean that language has no role to play in communication, but it does mean that its communicative role is merely the tip of an iceberg.

It is precisely in order to break with the traditional or instrumental concept of language that Pêcheux introduces the idea of discourse, and attempts to elaborate a conception of discourse which is original in theoretical, conceptual and empirical terms. In attempting to break with the instrumental concept of language, he is in fact following a trend which was already of considerable importance in France at the time he was writing. We have already had occasion to mention structuralism, and it is now necessary to make some comments on that topic.

### **Structuralism and language**

Pêcheux can no more be considered a structuralist than Lacan, Foucault or Althusser. Yet structuralism does place the emphasis on language, and this emphasis can also be found in Lacan and Foucault as well as in Pêcheux. Structuralism turned linguistics into the leading human science, and structuralists attempted to define their methods with reference to linguistics, and to transfer a set of linguistic concepts to almost every field within the 'human and social sciences'. Structuralism identified culture with language to such an extent that any account of a cultural phenomenon became a linguistic analysis, or something approximating to a linguistic analysis (semiology). This is not, however, true of, for instance, Lacan. Lacan does not attempt to reduce psychoanalysis to linguistic analysis. But his conception of psychoanalysis does stress that it is a 'talking cure', which operates exclusively through speech. In doing so he is arguing against psychological, biological and even sociological or anthropological trends within psychoanalysis. He refers to Saussure and Jakobson. He re-interprets Freud's *Verdichtung* and *Verschiebung* (condensation and displacement) in terms of metaphor and metonymy. And he elaborates a concept of the unconscious as being structured like a language, and of the subject as being a 'language-being' or a 'speech-being'. But it is to be noted that all Lacan's borrowings from linguistics (and from other scientific fields) are subject to a theoretical and operational reworking.

In structuralism, linguistic concepts and methods are simply

transferred to other fields without undergoing any fundamental reworking. To that extent, structuralists are rather like the scientist who weighed brains. They are therefore not – and this is a fundamental point – in a position to get away from the traditional idea that human nature (or the human mind) can be an explanatory principle.<sup>1</sup> This idea is inherited from Christian theology (which sees God as the ultimate explanatory principle behind human nature or the human mind, a principle with a privileged position) and from classical philosophy, which uses it to elaborate its conception of the human subject (which goes by various names, such as *Reason*). Structuralism cannot rid itself of the idea of the specificity of the human sciences which are defined (as opposed to other sciences) in terms of the specificity of their object, *man*. The notion of 'man' is, however, a *petitio principii* in that it presupposes that 'man', as an object of science, is something specific and well-defined. Structuralism therefore retains the idea that the human sciences provide the basis for a revival of humanism. In France, structuralism became the primary philosophy of the human sciences which asserted their specific difference from other sciences. As a result, it created an opening for all kinds of reductionism and for attempts to make human nature an explanatory principle in all respects, including the biological.

But at the very moment when this structuralist philosophy was being elaborated, theorists such as Lacan, Althusser, Foucault and Derrida were, in different ways, completely rejecting this conception of the subject and the associated conception of the 'sciences of man' (cf. Wahl 1968). Foucault (1970: 366-367) writes: "Western culture has constituted, under the name of man, a being who, by one and the same interplay of reasons, must be a positive domain of *knowledge* and cannot be an object of *science*".<sup>2</sup> Lacan (1966: 859) writes: "There is no science of man, because the man of science does not exist; only its subject exists". And Derrida (1970: 264-265) writes in a critique of Lévi-Strauss's structuralism:

---

<sup>1</sup> See, for example, Lévi-Strauss (1966), and the critique of his position formulated in Derrida and Mosconi (1975).

<sup>2</sup> There are – at the theoretical if not the practical level – many similarities between Foucault's work on discourse analysis and that of Pêcheux. The concept of a 'discursive formation' can, for instance, also be found in Foucault, even though he uses the term in a slightly different sense. Foucault never attempts to develop an instrumental system for automatic discourse analysis (cf. Foucault 1971 and 1972).

There are thus two interpretations of interpretation, of structure, of sign, of freeplay. The one seeks to decipher, dreams of deciphering, a truth or an origin which is free from freeplay and from the order of the sign, and lives like an exile the necessity of interpretation. The other, which is no longer turned toward the origin, affirms freeplay and tries to pass beyond man and humanism, the name man being the name of that being who, throughout the history of metaphysics or of onto-theology – in other words, through the history of all his history – has dreamed of full presence, the reassuring foundation, the origin and the end of the game. The second interpretation of interpretation, to which Nietzsche showed us the way, does not seek in ethnography, as Lévi-Strauss wished, the ‘inspiration of a new humanism’.<sup>1</sup>

A common strand runs through all these positions, which have been described as representing ‘a theoretical anti-humanism’: the need to rid theory of the transcendental subject in all its forms, including both the forms it takes in *theoretical* humanism and less overt forms (as in certain types of pseudo-materialist reductionist theories of the human mind). In this way they open up a field of questions and practices which are rendered impossible or inconceivable by the notion of the subject. Lacan, Foucault and Derrida all refer to language, the sign and discourse. The quotation from Derrida – and roughly equivalent statements could be found in the work of Lacan and Foucault – provides the key. Derrida criticizes attempts to decipher “a truth or an origin which is free from freeplay and from the order of the sign”. Language (or freeplay, or the order of the sign, or discourse) is not an origin or something harbouring a truth which exists independently of it. It is always-already there for any speaking being, and it is this which defines the position of the subject, the position of any possible subject. It defines the subject as a position, and not as a thing in itself or as a substance. No ‘positive’ definition of the subject as entity will be found in Lacan, Foucault or Derrida; they define only the position of the subject. It thus becomes possible to account for

---

<sup>1</sup> Pêcheux always regarded Nietzsche as a crucial figure in philosophy, arguing that his work is essential to any understanding of what is at stake in current philosophical debates and controversies. Shortly before he left us, he was planning to work specifically on Nietzsche.

the appearance of the transcendental subject and for the consequences of its appearance in terms of the 'always-already-there' quality of language rather than in terms of some reference to the inherent credibility of human nature. Language is not, then, turned into a substitute for 'human nature' or the 'structure of the human mind' as explanatory principle or as origin. And insofar as the subject is concerned, all attempts at reductionism are rendered non-pertinent.

### **Subject, discourse and ideology**

At the time when he wrote *Automatic Discourse Analysis* and the two 'Herbert' papers, Pêcheux was a follower of Althusser rather than of Lacan, Derrida or Foucault. We have seen that he was primarily concerned with the link between discourse and political practice and that in his view that link involved ideology. The second 'Herbert' paper is therefore devoted to outlining a general theory of ideologies. Following Althusser, it is with reference to ideology that Pêcheux introduces the subject and refers to it as the elementary ideological effect. It is insofar as the individual is a subject that he or she is 'interpellated' to take a given place in the system of production. In a paper to which Pêcheux often refers, Althusser writes: "Like all obviousnesses, including those that make a word 'name a thing' or 'have a meaning' (therefore including the obviousness of the 'transparency' of language), the 'obviousness' that you and I are subjects – and that that does not cause any problems – is an ideological effect, the elementary ideological effect" (1971a: 171-172). Why 'elementary'? What does that mean? It means precisely that this 'effect' is not the result of something. The subject that is interpellated is always-already a subject. More specifically, Althusser writes (p. 170): "there is no ideology except by the subject and for subjects", adding that an ideology always exists in a practice or that, in other words, no human – that is to say social – subject can be the agent of a social practice except insofar as he or she is a subject.

These formulations were put forward after the publication of *Automatic Discourse Analysis* and the two 'Herbert' papers. They are, however, a systematization of the theoretical theses underlying Althusser's work on Marx's *Capital* (cf. Althusser and Balibar 1970; published in French in 1965), and Pêcheux was familiar with that work. It comes, therefore, as no surprise to find that the 'Herbert' papers are consistent

with those theses. Althusser's work on *Capital* takes the form of a re-reading which attempts to break with the established dogmatic reading of Marx. This rereading of Marx was made in accordance with a method which is defined by Althusser as 'symptomatic reading': a reading which focuses upon the discontinuities, shifts, stumbling blocks and reformulations to be found in the text.<sup>1</sup> This method implies that Marx's texts have to be compared with one another before they are referred to anything outside them.<sup>2</sup> The idea is to open up Marx in order to promote further theoretical elaborations without losing sight of his contributions, rather than regarding his work as a kind of Bible. This method obviously influenced Pêcheux. It might be said that one of the objectives of his work on discourse analysis and the theory of discourse was the theorization and systematization of this method.

Althusser's rereading of Marx is not, however, based solely upon a method; it also involves the use of 'philosophical tools'. In his 'Elements of self-criticism' (1976), Althusser explains that, despite appearances, he never was a structuralist, but that he was guilty of a much more serious and more compromising passion: he had been a Spinozist. Taking the view that every philosophy has to make a detour through other philosophies in order to define itself and to understand its own specific difference<sup>3</sup>, he explains that, just as Marx had to make a detour via Hegel, he himself had to make a detour via Spinoza. In the course of this detour via Spinoza, Althusser discovers a concept which enables him to explain Hegel's lingering concept of the subject as origin (or end). It enables him, in other words, to explain the roots of Hegel's idealism, and therefore to understand what Marx meant by the enigmatic remark to the effect that he had 'inverted' Hegel. According to Althusser, Spinoza helps us to understand how and why such an 'inversion' was possible, and to disclose the 'dif-

---

<sup>1</sup> A parallel was drawn, at that time, between Althusser's rereading of Marx and Lacan's 'return to Freud'. It is to be noted that Althusser refers to Freud when he attempts to define his method. The question of the subject is, however, already present in the first 'Herbert' paper, at least to the extent that it criticizes the subject of traditional philosophy.

<sup>2</sup> It has sometimes been seen as a structuralist method because it resembles the structuralist methods applied by Vladimir Propp to folk tales (1968) or by Lévi-Strauss to myths (1963); both authors compare different versions of a single tale or myth.

<sup>3</sup> The link between this point and Althusser's method is to be noted: it is only via texts that a philosophy can relate to other philosophies.

ferences' between Marx and Hegel. In a sense, Althusser's thesis is that Spinoza's category of an 'effect without (external) cause (or end)', which supports the famous *verum index sui et falsi*, anticipates Marx on a *specific but crucial* point relating to the central category of idealism: the subject as origin, essence or cause. Althusser is thus attacking the very concept of the subject that Lacan, Derrida and Foucault are attacking. But he does so on a quite different basis and for quite different purposes (to understand the similarities and differences between Marx and Hegel). In his self-criticism, Althusser also explains that, whilst his detour via Spinoza had some positive results, it was not without negative effects. According to Althusser, Hegel supplies Marx with the notion of contradiction, but there is definitely no notion of contradiction in Spinoza. That absence leads Althusser to see ideology as a universal element in history. He therefore constructs a 'theory' of ideology which overlooks the differences between ideological regions, and the antagonistic class trends that run through them, divide them, and confront them to one another. This 'theory' of ideology is precisely the theory that is outlined in the paper on 'ideological state apparatuses' (Althusser 1971a). In other words, Althusser believed that he had been saved from falling into the trap of structuralism by Spinoza (and by the critique of the traditional subject of philosophy which he found in Spinoza), but he in fact fell into the trap laid by that powerful critique, which prevented him from seeing the role of contradiction and of class struggle in ideology.

And what do we find in the 'Herbert' papers? A reference to Spinoza in the first, and an attempt to outline a general theory of ideologies in the second. The papers clearly follow Althusser's line. This becomes all the more apparent if we compare Althusser's position with that of Lacan, Derrida or Foucault. As we have seen, Althusser shares their position regarding the status of the subject. It is with regard to the question of the subject that Althusser, like Derrida, Foucault and Lacan, spells out why he parts company with structuralism<sup>1</sup>, and dismisses the idea that the human sciences could be grounded on the specificity of human nature. But there is also a fundamental difference between Althusser and the others. The difference is that, whereas Lacan, Foucault and Derrida relate the subject to the impossibility of escaping 'freeplay and the order of the sign', as Derrida puts it, Althusser relates the subject to the

---

<sup>1</sup> In the case of Foucault, this is particularly clear in the conclusion to *The Archaeology of Knowledge* (Foucault 1972).

impossibility of escaping ideology. "Ideology *has no outside*", writes Althusser (1971a: 175). He does not say that ideologies have no outside. He does of course accept that there are different ideologies and different ideological positions. These different ideologies or ideological positions are antagonistic (but not contradictory). *An* ideology has an outside, but its outside is made up of other ideologies. If science exists, it must exist among ideologies. Yet, Althusser makes a distinction between science and scientific theories. Scientific theories are statements, and as such they embody ideologies. In a sense, all theories are ideological. All theories are provisional. A theory cannot be true; it can only be more true than another theory. In other words, Althusser's subject is the subject of ideology, and there is no subject other than the subject of ideology.

Althusser's subject is not the subject as seen by Lacan, Derrida or Foucault. Foucault (1972: 95-96) writes: "To describe a formulation *qua* statement does not consist in analysing the relations between the author and what he says (or wanted to say, or said without wanting to); but in determining what position can and must be occupied by any individual if he is to be the subject of it". There is no other way of being a subject. In other words, to be a subject is to occupy the position of the author of an utterance. Discourses are utterances. The elementary unit of discourse is the utterance. Foucault's definition of the subject is consistent with his conception of discourse. And we are now in a position to state that his subject is the subject of discourse. We must also bear in mind Foucault's aim: to arrive at a new definition of the field occupied by the history of ideas, a definition which could renew it by overcoming its shortcomings, namely its taking psychological subjectivity as an explanatory principle. Foucault's subject is the subject of 'the order of discourse' (Foucault 1971). Derrida's aim is to renew philosophy by sparing it the attempt to look for a truth or an origin outside of freeplay and the order of the sign. His subject *is* the subject of freeplay or of the order of the sign. Lacan's aim is to renew psychoanalysis, and his subject is the subject of 'the unconscious structured like a language'. Language is the condition of the unconscious, and it introduces a discrepancy between speaking beings and their reality. And Althusser's aim is, as we have seen, to renew Marxism and historical materialism.

We have, then, a number of attempts at renewal, and they all see the subject and the status of the subject as the crucial issue. But the intersections between the definitions of the subject given by Foucault, Derrida and Lacan are more obvious than those between their definitions



and Althusser's. The definitions of the subject offered by Lacan, Derrida and Foucault are bound up with language or with signs (in the strict sense, or as distinct from signals). An appeal to ideology does not have the same implications as an appeal to language. Althusser is not greatly concerned with language, and this brings us to Pêcheux's central concern: to link language to ideology. The only means he has at his disposal is Althusser's indication of a parallel between the obviousness of the 'transparency' of language and the obviousness that we are subjects (the elementary ideological effect). Althusser establishes a parallel, but not a link between the two. And it is in order to establish that link that Pêcheux introduces what he calls *discourse*, and attempts to elaborate both a theory of discourse and an operational method for discourse analysis.<sup>1</sup>

### Pêcheux's theory of discourse and his discourse analysis

Pêcheux works in the space between the 'subject of language' and 'the subject of ideology'. This has a bearing on all his work, and not only on his *Automatic Discourse Analysis*. In *Language, Semantics and Ideology* (Pêcheux 1982a) he attempts to clarify the links between the 'obviousness of meaning' and the 'obviousness of the subject', and locates discourse in a region intermediate between language and ideology (defining 'language' not as an empirical object, but as the object of linguistics, or what Saussure calls *langue*). In this later work he does not abide by the formulations advanced in the 'Herbert' papers or in *Automatic Discourse Analysis*. Like Althusser, he abandons the idea of developing a general theory of ideology (or of ideologies). He begins to work on a new problem, that of the connection between the object of linguistics and the object of the theory and analysis of discourse.<sup>2</sup> This is not merely a theoretical problem; it is also a problem inherent in the system of discourse analysis he has elaborated. It relates to the kind of linguistic analysis that is required if the analytic system is to become operational and capable of realizing its

---

<sup>1</sup> This implies that Pêcheux and Foucault have very different conceptions of discourse.

<sup>2</sup> Almost at the same time, Foucault moves from discourse, knowledge and the history of ideas to power and knowledge (cf. Foucault 1977 and 1980).

potential.<sup>1</sup> Ultimately, this problem is also a theoretical problem, namely that of the link between the subject of language and the subject of ideology. Pêcheux never abandons this problem, even though he does reformulate it completely. In his last book, written in collaboration with Françoise Gadet (Gadet and Pêcheux 1981), he again deals with linguistics and argues that the ambiguities which arise when it is confronted with the disjunction between what makes sense and what does not make sense, raise a theoretical and political question. Citing Kundera (1986: 172), he asserts that 'even a metaphor is worth fighting for'.

One last comment should be made concerning Pêcheux's strategy. We have already seen why he divorced his account of his system for automatic discourse analysis from any statement concerning the theoretical (and political) problems which led him to elaborate that system. Like any strategy, this strategy has its disadvantages, because it is possible to see the system as an instrument or tool to be used in empirical research without a theoretical effect. It has been used that way. Pêcheux foresaw the problem and stated that his instrument was not to be used in that way. Rather, he saw his system as a kind of Trojan Horse which could be infiltrated into the social sciences in order to bring about a revolution (it should be noted that Foucault saw his 'archaeology' in a similar light). This strategy has had some effect; using the instrument has led many researchers to raise questions they would not otherwise have encountered. This is exactly what scientific instruments do according to Pêcheux: they rather raise questions than provide answers. Pêcheux's attempt to specify the status of 'scientific instruments' is still worthy of attention. It is particularly relevant in judging the claim that the Turing machine, computers, connexionist networks and neuronal networks have laid the foundations for a 'new science of mind'. Unfortunately, Pêcheux is no longer with us in our struggle against the re-appearance of that monstrous claim.

---

<sup>1</sup> See Chapter 5 below for the way Pêcheux deals with what he calls the 'appropriation' of instruments, in this case linguistic and computing 'instruments'.

## Pêcheux's theory of ideology and the subject

Françoise Gadet, Jacqueline Léon, Denise Maldidier  
and Michel Plon

At the end of the prefatory note to his article 'Freud and Lacan', Louis Althusser (1971b: 193) writes:

An accurate assessment of these [psychoanalytic] concepts can only come from the serious study of Freud and Lacan which each one of us can undertake; the same is true for the definition of the still unsolved problem of this theoretical discipline already rich in results and promises.

Whatever the intrinsic merits of this article (and its author has never denied that it is somewhat approximate), it is historically important and an essential point of reference, due to its date of publication, its impact and the breach it opens up in the barriers which, since Politzer (1928), Marxist thought had erected between itself and psychoanalysis, especially in France. Pêcheux's relationship with psychoanalytic theory or, to be more accurate, the elements of that relationship that can be excavated, can only be understood within the theoretical conjuncture of which Althusser's article is an expression.

Jacques Lacan was born in 1901. He trained as a doctor and psychiatrist, studied under Clérambault, and in 1932 he published his thesis, which breaks with the constitutionalist current within psychiatry, adopts a dynamic clinical approach and turns to a concrete psychology which allows him to promote the concept of 'personality' which was revolutionary at that time. In this work we can already detect an interest in what lies outside psychiatry and forebodings of the path taken by the young doctor, friend of the surrealists, as he moves towards his discovery of Freud. From this point onwards, Lacan's public life and work gradually

became part of the history of psychoanalysis in France; they changed its course and eventually came to orchestrate it in its entirety (cf. Roudinesco 1990). In 1966, Lacan published a volume of almost one thousand pages, entitled *Écrits* (cf. Lacan 1977). It brings together most of the articles he had written since 1936, the year of his first public intervention as an analyst at the Fourteenth Congress of the International Psychoanalytic Association in Marienbad, where he read a paper on the mirror stage; some of the papers have been reworked, whilst others appear in their original form. The collection, which – more so than any other work of the period – was to become the bedside book of a whole generation of French intellectuals, closes with ‘La science et la vérité’. This text, which was the opening lecture of the Seminar given by Lacan at the *École Normale Supérieure* (ENS) in 1965-1966, was originally published in the first issue of *Cahiers pour l'analyse*, the journal of the *Cercle d'Épistémologie* at the ENS.

Because Pêcheux was a ‘normalien’, who studied philosophy under Louis Althusser, and was a member of the *Cercle d'Épistémologie*, it might be thought that he would inevitably have been familiar with Lacanian thought. But this is clearly not the case, as is obvious both from the first edition of ADA69 and from Pêcheux’s subsequent work. Although he always treats psychoanalytic theory with the utmost respect – a respect which is expressed either silently, or through references, or in the form of limited attempts to use Freudian and Lacanian concepts – his view of psychoanalysis remains somewhat ‘gauche’, and his borrowings relate less to the unconscious itself than to theorizations of its workings and effects. Neither Freud nor Lacan figures in the bibliography of ADA69, and psychoanalysis itself is mentioned only in veiled terms (cf. the note on p. 118 below). It may be possible to explain this silence in tactical terms relating to the academic strategy adopted by Pêcheux at the time, to the theoretical orientation of the series in which ADA69 appeared, to the Piagetian position of François Bresson (who was the series editor), or even to Pêcheux’s institutional position in the psychophysiology and psychology section of the *Centre National de la Recherche Scientifique*. That section was heavily influenced by positivist conceptions which privileged the development of a psychology as compatible as possible with the neurosciences, and the supporters of that psychology made no secret of their hostility to psychoanalysis.

All these reasons are valid, and as such perfectly acceptable, but they do not solve the problem completely. A reading of the two articles

Pêcheux published in *Cahiers pour l'analyse* under the pseudonym 'Thomas Herbert' (1966, 1968) suggests a number of hypotheses involving factors other than purely tactical considerations, and those factors may explain what can only be termed Pêcheux's elliptic use of Freudian and Lacanian theory.

In the 'Herbert' articles, psychoanalytic theory or 'psychoanalysis as science of the unconscious' is evoked, and the names of Freud and Lacan are mentioned in the second article, where Freud is quoted only once, and Lacan is merely mentioned in passing. The overall perspective and the modalities of these references give some indication as to the role Pêcheux ascribes to psychoanalytic theory in the conceptual apparatus he is elaborating and as to the extent of his knowledge of developments within Lacanian theory. A first point concerns the central role given in his conceptual apparatus to historical materialism which is, to use one of Pêcheux's favourite expressions, 'put in command'. This is eminently clear in the 'Herbert' articles, but it is also apparent, albeit in rather more veiled terms, from the reference in ADA69 to the concept of 'conditions of production' (p. 79 below). In Althusserian circles, the concept of production, taken from the sphere of economic activity and of material production, was systematically imported in the sphere of intellectual activity; the theorist, the philosopher, the writer, the painter and the musician were regarded as workers like any other worker (cf., for example, Macherey 1966). The role of historical materialism is also obvious in some of the examples, e.g. when Pêcheux describes positions A and B as designating "determinate positions within the structure of a social formation", and when he adds that "sociology can describe the bundle of objective features that characterize them; within the sphere of economic production, for example, the positions of the 'boss' (director, factory manager), the executive, the foreman and the labourer are marked by identifiable differential properties" (p. 85 below). It should be noted in passing that the reference to sociology, which reappears at the end of ADA69 when Pêcheux discusses the possible applications of his theory, is to something very different to official sociology. Pêcheux is referring to the new sociology which Louis Althusser and his pupils thought they could develop on the basis of historical materialism or the science of social formations. This is the real principle behind ADA69, and it largely determines the role given, at least momentarily, to psychoanalytic theory and the perspective within which it is inscribed. To go straight to the point, even though all these questions should be described in much greater detail than is possible

here, we might say that the role of psychoanalysis is 'regional'; that is the term Pêcheux uses in ADA69, where he speaks in rather ambiguous terms of, for instance, a 'regional theory of the signifier' (see p. 72 below).

The role given to psychoanalysis in the work of Thomas Herbert might be described as subordinate as well; psychoanalysis intervenes in a schema designed to account for the transformation of ideological elements into a conceptual system. In the second 'Herbert' article, the ambiguous modality of the recourse to psychoanalysis is even more pronounced; certain psychoanalytic concepts are used as 'instruments', and it is pointed out in rather curious terms that they "were originally constituted for psychoanalysis" (1968: 83). This comment provides the author with an opportunity to raise a problem, but he formulates it in such a way as to introduce a further ambiguity as he refers to "the relationship between the *analytic unconscious* and the social unconscious of ideological repression". A rereading of this article (which must address the reasons why there is a rather fundamental – and not a merely tactical – difference between the 'Herbert' articles and ADA69) provides many other indications which reveal that, at this time, Pêcheux's project was still oriented towards a theoretical horizon implicitly dominated by the fantasy of articulating historical materialism (as the bedrock) with the theory of the unconscious (as a regional contribution). The only explicit reference Thomas Herbert makes to Freud is a telling indication of the significance of this fantasy and of the meaning of its organization; a passage from the *Introductory Lectures on Psychoanalysis* (Freud 1971: 311-312) is used to demonstrate that Freud 'does not ignore' the question of 'the reproduction of man as labour power', suggesting that he is therefore suitable company for Marxist theorists (Herbert 1968: 90). These lines, which reveal, it has to be said, a prodigious level of curiosity and understanding, implicitly show how Althusser's appeal for a 'serious study' of Freud and Lacan met with the lingering effects of Marxism's excommunication of psychoanalysis as a 'petty-bourgeois' ideology. The fantasy functions according to an ambivalent mode of repulsion and fascination, and it finds a further expression in the persistent presence in ADA69 of a veritable social psychology, even though Pêcheux's stated aim is to dissociate himself from it.

One central concept allows us to locate the point around which most of the difficulties we have so far evoked crystallize: the concept of the subject. The concept of the subject or the subject of the unconscious, as elaborated by Lacan throughout his work, undergoes a series of transformations which, in order to avoid contradiction, rarefy the concept,

removing it definitively from any, even a metaphorical, form of localization and substantialization in an early stage, inscribing it exclusively within the register of structure (cf. Ogilvie 1987).

In Lacan's 'mirror stage' period the category of the imaginary was firmly anchored in the body and opened up only onto the alienating perspective of identification. Lacan ran his Seminar year after year, and in 1969 it migrated to the Law Faculty near the Pantheon after having been driven out of the ENS during the roll-back and restoration which characterized the months that followed May 68. By this time, Lacan had already elaborated his topography of the real, the symbolic and the imaginary; he was now working on topology, organizing his reflections around the theme of Borromean knots, and had set off in search of the mathemes of psychoanalysis and to protect it from the imaginary deformations inherent in the literality of discourse. The subject of the unconscious constantly appears and is then effaced only to reappear as a residue; it arises from within the site of the symbolic, the place of the *Other*, as distinct from the imaginary *other* that is related to the place of the ego, the subject of psychology and of social psychology. In other words, for Lacan the subject is not a basic *datum*, a site within the psyche from which the operations of recognition and misrecognition – which constitute the internal mechanisms of ideological reproduction as conceived by Thomas Herbert – are controlled. With his celebrated axiom 'the signifier represents the subject for another signifier' (Lacan 1966: 840), Lacan indicates the absolute incompatibility of his 'subject of the unconscious' with any form of localization in which a subject can be identified as a support for operations that are definitely psychical.

A sentence from the second 'Herbert' article indicates the distance separating the Lacanian concept of the subject from his. Arguing against those theories which see society as a system in which each element is a reflection of the whole, he writes (1968: 82):

If, on the contrary, we apply to the question which concerns us the utterance which Lacan formulates to a (partly) different end – namely 'the signifier represents the subject for another signifier' –, we find that the syntactic chain of signifiers assigns the subject its place by identifying it with a certain point in the chain (the signifier which represents it), and that this mechanism of differential identification is none other than the 'society-effect' of

which the dissymmetries can be explained by this mechanism.

This is a major misunderstanding and we can, in general terms, characterize it in two ways. On the one hand, explicit use is made of a problematic of application; Pêcheux's use of the term 'partly' presupposes that Lacan's utterance can be used to ends other than those for which it was forged. On the other hand, the concept of 'subject' developed here implies the idea of a subject that exists prior to the operation of identification by the signifier. This operation is said to assign a place to the subject, which would be articulated with social processes (designated by the expression 'society-effect').

The whole passage on the conditions of production of discourse (pp. 82-94 below) is of vital importance in that it is one of the few 'non-technical' passages to describe certain of ADA69's theoretical premises. Initially, it is a matter of choosing between two families of schemata which can be used to describe linguistic *behaviour*. But the very notion of a choice leads to the immediate ratification of the notion of behaviour, which is central to psychology. The 'reaction' schema of stimulus and response is obviously rejected in favour of the 'information' schema given in the article by Moscovici and Plon (1966). This reference is not unproblematical; although the behaviourist schema is rejected by Moscovici and Plon, their work is firmly grounded in psychology, and both their article and the experiment it describes ignore Freud and Lacan, as well as contemporary developments in linguistics.

The schema which is introduced designates elements A and B as 'addresser' and 'addressee'. A and B are 'points' or 'positions' which "are represented within the discursive process within which they are brought into play". But a few lines later Pêcheux refers to "the positions which A and B ascribe to themselves and to one another; the image they have of their own position and of the position of the other" (p. 85). There is something distinctly odd about positions which have images of their own positions and of the positions of other ... positions!

We then find a summarizing and explanatory schema in which one can read that  $I_A(A)$  signifies 'the image of A for the subject placed at A', a notion which can be expressed as: 'Who am I to talk to him like that?'. Such a question could only be asked by the 'subject' of social psychology, and that 'subject' belongs to the realm of the imaginary. The schema unwittingly confirms this by its failure to make any distinction between 'I'



and 'me' (the ego).

It might be said that this passage brings us to the heart of the problem and to the source of the misunderstandings that will arise as a result of ADA69. Compared with the article by Moscovici and Plon and, more generally, with the social psychology of 'communication', Pêcheux's approach does seem to represent a major advance in that it breaks with its formalism if only because it makes a decisive point which social psychology overlooks for structural reasons. Pêcheux takes into account one essential parameter: the topic under discussion or the referent of discourse, and this implies a willingness to admit that the way in which something is talked about changes everything. But at the same time that breakthrough – and anyone who doubts that it is a breakthrough should look at the reactions of the psychologists – leads to the suggestion that it is possible to develop a field of knowledge and research without taking psychoanalytic theory into account. This is an illusion in two senses. It is impossible to ignore psychoanalytic theory without lapsing back into psychology. And it is impossible to ignore psychoanalysis without subverting the whole project, without undermining its very foundations. Pêcheux constantly returns to all these points. All these questions – and others raised by this passage and by the references to the work of Ducrot – are dealt with in admirable fashion by Paul Henry (1977).

Only a rudimentary account of these difficulties has been given here, and they are not specific to Pêcheux alone; they have to be related to the conjuncture within which ADA69 was conceived, and they testify to the extent of the breach opened up by Lacan's contribution, a breach whose importance is often underestimated, not in the least because it is papered over by psychoanalysts themselves. In Pêcheux's later work, these difficulties are attenuated insofar as the fantasy of the articulation between Marxism and psychoanalysis begins to lose its hold over him; the publication of Althusser's article on the ideological state apparatuses (in 1970) will provide an unexpected support for that fantasy, but it is also marked with the seal of the impasse to come.



## Pêcheux's linguistic references

Françoise Gadet, Jacqueline Léon, Denise Maldidier  
and Michel Plon

We will attempt to define Pêcheux's concept of the language [*langue*] by providing a few brief monographs on six names and themes relating to the linguistic conjuncture that prevailed in France in the early 1960s: Saussure and structuralism; the reception given to Chomsky and generative grammar; Harris; Jakobson; Benveniste and enunciation; Culioli. It is necessary to introduce these six themes because it is in fact apparent that the concept of the language from which Pêcheux derives his concept of discourse is shaped by a number of contributions, which he soon discovers to be heterogeneous, and by the philosophical positions he adopts within the theoretical landscape of the 1960s. Broad as his linguistic horizons may be, his actual grammatical practice in fact often reproduces a traditional grammar which is basic to the French educational system: parsing, clause analysis, and the principles of rhetoric. This provides a basis, and more recent contributions introduce only minor modifications. In his later writings, Pêcheux rapidly becomes critical of the concept of the language that is at work in ADA69, and then attempts to modify it. Yet the format of the elementary utterance remains constant until the early 1980s, when the ADA69 programme is abandoned, and when the use of DÉREDEC means that comparisons no longer have to be restricted to sequences of equal length (see Chapter 8).

### Saussure and linguistic structuralism

It is obvious that, even at the ADA69 stage, Pêcheux had read Saussure very carefully, and that he continued to do so throughout his career (see, for example, Haroche, Henry and Pêcheux 1971 and Gadet and Pêcheux

1981). This is worthy of note in that the period was, in general, characterized by a rather vague interest in Saussure, who was a reference rather than an object to be worked upon. In the 1960s, Saussure was in fact read in a variety of ways.

- In France, the structuralist reading was influenced by Martinet (whose *Éléments de linguistique générale* appeared in 1962) and by Mounin (whose *Saussure ou le structuraliste sans le savoir* appeared in 1968). This reading had yet to become what Lepschy (1966) calls 'the vulgate', but it did have its reductive side since it made little use of the *Sources manuscrites du Cours de linguistique générale*; these had been studied by Robert Godel, but they were not well known even though they had been available since 1957.

- The sociolinguistic reading was essentially a militant and critical reading; it was primarily designed to show that the distinction between the language and speech is an adequate instrument for dealing with the problems of discourse and of the use of the language in a social context.

- Benveniste's article 'Saussure après un demi-siècle' (1966: 32-45) aside, 'philological' readings of Saussure's text had little impact outside specialist circles. Godel's *Sources* appeared in 1957, but it was published by a small Swiss house (Droz) with only a small distribution network. Engler had begun his monumental undertaking (five volumes published between 1967 and 1974), and he had already published related articles in the *Cahiers Ferdinand de Saussure*, but that journal had only a limited circulation. It was only from 1970 onwards that Normand's articles began to appear in journals with a specifically Marxist theoretical orientation (*La Pensée*, *Dialectiques*) and in *Langages*, France's most important linguistic journal (Normand 1970, 1973, 1978). We can see, then, that it took a long time for linguists to take an interest in the sources.

- In terms of literary readings, in 1961, Starobinski began to publish the articles which were later published in book form as *Words upon Words* (1979). These dealt with an aspect of Saussure's work which was still largely unknown: the anagrams. In France, knowledge of the anagrams was popularized by people such as Jakobson, Kristeva, Todorov and Barthes, and then by Calvet (1975). In the mid-1970s, the Americans took over where these authors had left off, the major event being the publication of the 'two Saussure's' issue of *Semiotexte*.

Although he did not in fact adopt any of these categories of readings, even in 1969 Pêcheux proved himself to be very familiar with Saussure's text; his reading was well-informed, intelligent and original,

and relied upon a real ability to work with Saussurean notions. Paul Henry recalls that at this time Pêcheux had already worked on the *Cours de linguistique générale*, and had read the *Sources manuscrites* (Godel 1957), Starobinski, and even Saussure's work on the *Nibelungen*. The effects of this detailed knowledge can be seen in ADA69:

- In terms of his general conception of the language, Pêcheux's interest in the functioning of languages rather than their function shows that he had already grasped the basis of the displacement brought about by Saussure and the basic postulate on which modern linguistics has been founded since Saussure: the language is a system.

- Whilst he did, like the sociolinguists, conclude that the opposition between the language and speech cannot address the problematic of discourse, he did not try to resolve the problem by diluting that opposition; on the contrary, he concentrates on the axis of speech, which is not fully developed by Saussure.

- As to the role given to 'metaphoric effect' (see pp. 96-100 below), Pêcheux was definitely influenced by Jakobson's metaphor/metonymy pair as described in 'Linguistics and poetics' (Jakobson 1981), but the main influence may have been his own understanding of Saussure's position on the language, an understanding which seems to owe something to his study of both the concept of value and the anagrams.

This Saussurean stance does not, however, preclude the adoption of certain banal 'structuralist' formulations ('structuralist' in the vulgarate sense).

### **Chomsky and generative grammar**

In France, generative grammar began to be diffused in 1965. The fourth issue of *Langages*, which was entitled *La grammaire générative*, appeared in 1966, and Ruwet's introductory study was published in 1968. The first translation of Chomsky appeared in 1969 (*Syntactic Structures*). In ADA69, generative grammar is a stimulating theoretical horizon rather than a source for formal, conceptual or methodological borrowings. On reading Pêcheux, one has the impression that Chomsky – by integrating the sentence and a non-subjective linguistic creativity into his theory – has taken over the lead from Jakobson in the triumphal march of linguistics. ADA69 expresses something of the enthusiasm that the 'Chomskyan Revolution' inspired amongst certain linguists. But for Pêcheux, the

Chomskyan revolution provides food for thought rather than a solution. He reacts against it by resisting the dizzying temptation to use the model of generative grammar to construct a mechanism for the production of discourse. But he also endorses it by borrowing, albeit in metaphorical terms, the Chomskyan opposition between deep structure and surface structure. As we shall see, this opposition allows him to postulate a relationship between discursive structures which can be analyzed in terms of sites of surface effects and an 'invisible structure' which determines them.

It should be noted that the opposition between Saussure and Chomsky is not always clearly evaluated: linguistics is said to be part of a 'science dealing with the sign' (p. 71 below), but the assumption that the theory of the sign is synonymous with linguistics is quite alien to the Chomskyan conception of language (on this point, see Gadet 1989, especially Chapter 3: 'The Sign').

## **Harris**

Harris provides the main inspiration for Pêcheux's method of analysis, even if it is somewhat heteroclite. The first French translation of a paper by Harris (1952) appeared in the journal *Langages* in 1969 (issue 13). The name of the American linguist figures in the bibliography of ADA69, but there is only one reference in the text (Pêcheux 1969a: 49). We are not in a position to provide an explanation for this absence, and we can only outline a number of possibilities.

Pêcheux made explicit his debt to Harris in later work (cf. Chapter 6 and 7 below; see also Woetzel and Geier 1982). It seems to us that Harris does more than supply certain analytic procedures; he provides the inspiration for the development of the ADA69 method itself. The similarity between Pêcheux and Harris is particularly clear in the recording of the discursive surface, which is in fact the 'linguistic analysis' phase of ADA69. The text is reduced to elementary utterances which recall Harris's 'kernel sentence'. Pêcheux uses transformations (a grammatical technique which is an essential feature of Harris's method), and attempts to use those operations to arrive at an optimal regularization of the discourse under analysis in order to constitute semantic domains. Although it is a 'method' that is being borrowed rather than a 'theory of language', Pêcheux is well aware of the fact that it is based upon 'theoretical presuppositions which

the linguist must make explicit and criticize' (Pêcheux 1969a: 85).

In ADA69, the implicit reliance upon Harris leads to the emergence of questions pertaining to synonymy and substitutability, and to semantic variance and invariance; these will later centre on the question of paraphrase (see Chapter 6), which Harris introduces in connection with the distinction between incremental and paraphrastic transformations (Harris 1957). The question of paraphrase, which is the motor behind the whole method, will continue to be of central importance up to the appearance of *Matérialités discursives* (Conein et al. 1981).

### Jakobson

In 1963, a collection of articles of Jakobson was published in French, with a preface by Ruwet. These articles provided Pêcheux with both food for theoretical thought and instruments for linguistic analysis. At the level of theoretical propositions, Jakobson is used negatively in an attempt to find in linguistics anti-subjective positions: Pêcheux rejects the view that as we move from phonemes to discourse we move *gradatim* from the necessity of the system to the contingency of freedom, and are dealing with increasingly effective combinatory rules (see p. 74 below). But, on the positive side, Pêcheux finds in Jakobson propositions which make it possible to open up linguistics, to expand its limits whilst still remaining within the framework of structuralism. This explains his adoption and reformulation of the famous schema of communication (p. 84 below), and his reference to the passage in which Jakobson argues that the unity of language should be seen as 'a system of interconnected subcodes' (p. 76 below). Pêcheux's uncritical citing of Jakobson indicates that he is looking for a theoretical support which can account for discursive variation within an invariant language (Jakobson 1981). Pêcheux's other references to Jakobson concern the framework of linguistic analysis. They occur in the chapter devoted to the analysis of the discursive surface (Pêcheux 1969a: 73-77) and concern, albeit in rather confused fashion, the notation used to record the elements within the verbal syntagma which refer to the relationship between utterance and enunciation (tense, voice, mode, person). Jakobson's suggestions are adopted in a strikingly utilitarian manner. They are seen as purely technical references which make it possible to give a certain consistency to the 'form of the utterance' (cf. pp. 112-113 below).

### Benveniste and enunciation

It is clear that in 1969 Pêcheux did not address the issue of enunciation. This is confirmed by the minor role given to Benveniste. None of the three references to Benveniste shows any real understanding of the gap opened up in structuralism by his recognition of the role of enunciation. Pêcheux soon recognizes this; in 1975 he states that ADA69 remained 'blind' to the problem of enunciation (p. 148 below). Pêcheux relies upon Benveniste when, as opposed to Jakobson, he refers to the sentence as the unit of discourse and as the frontier of a domain which is irreducible to the order of grammar. But he draws no theoretical conclusions. And whilst he describes Benveniste's position as implying that speech is infinitely creative, he clearly has ideological doubts. Benveniste appears to Pêcheux to be the linguist of subjectivity. It is not clear just how Pêcheux came across Benveniste's work<sup>1</sup>, but it seems that he initially saw him as having taken a step backwards, as having effected a return to a psychological subject which had been victoriously driven from the theoretical scene by Saussure and structuralism. If, like Normand (1985), we accept that linguistics discovered the theory of enunciation by reading the work of Jakobson rather than that of Benveniste, it would appear that Pêcheux shared the dominant view. This temporary failure to recognize Benveniste's importance can be explained in terms of strictly theoretical reasons. Being primarily concerned with the question of the subject, Pêcheux inevitably soon came to attach a new importance to the problem of enunciation. This is clear in his reflections on the necessary illusion which constitutes the subject and in his theory of the two 'forgettings' (pp. 129, 137-140 below).

### Culioli

We know that Pêcheux took an interest in the work done in Culioli's seminar in Nancy in 1967, and that he worked at the *Centre d'Études de Traduction Automatique* (CETA) in Grenoble, where Culioli worked as well. He worked with Culioli on the question of determiners (cf. Culioli, Fuchs and Pêcheux 1970). This convergence of interest is signalled by two

---

<sup>1</sup> Did he read 'De la subjectivité dans le langage', which appeared as early as 1958 in the *Journal de psychologie*, before its republication in *Problèmes de linguistique général* (1966: 267-276)?



important points in ADA69:

- Pêcheux uses Culioli's term *lexis* (see pp. 112-113 below). This is, however, a purely terminological borrowing: "The term *lexis* was merely mentioned in ADA69" (Pêcheux and Fuchs 1975: 52). From this, Pêcheux derives the idea of an eight-position schema resulting from the application of a 'form' (which introduces determinations and modal values) to a morpho-syntactic structure.
- As Pêcheux stresses in a note, his analysis of the determinations of the noun and verb to be found in an utterance is directly inspired by Culioli (Pêcheux 1969a: 70).

On the whole, ADA69 attracted little attention from linguists when it was published. There was, however, one exception, and it opened up a whole new future. In the review section of issue 151 of *La Pensée*, the 'journal of modern rationalism', which carried Althusser's famous article on 'Ideology and ideological state apparatuses', the linguist Geneviève Provost devoted an important note to ADA69 (Provost-Chauveau 1970). Whilst her article did raise some strictly linguistic issues (the problem of synonymy) and did query the method of analysis, it concentrated primarily on Pêcheux's contribution to discourse analysis. She called it "an interesting approach to current trends relating to discourse analysis" (p. 135). It is an indication of the convergence between Pêcheux's preoccupations and those of a group of researchers working, like Provost herself, with the linguist Jean Dubois on a new academic discipline: discourse analysis. The recognition that they had a common object went hand in hand with their search for methods for linguistic analysis (Harris was an implicit reference for Pêcheux, and a central point of reference for Dubois's group; note that it was Dubois who had translated Harris's 'Discourse Analysis' in 1969). Their work was inscribed within a broadly Marxist view of relations between language and social class. Provost has the merit of having immediately recognized that the publication of ADA69 was a major event. The history of discourse analysis in France – insofar as it can be reconstructed afterwards – was heavily influenced by the conceptual positions of ADA69.



# Appendix

## Pêcheux's approach to formalization

Françoise Gadet, Jacqueline Léon, Denise Mالدیدیر  
and Michel Plon

The considerable role given to formalization in ADA69 is, for Pêcheux, inscribed within a double perspective. On the one hand, it is inscribed within an epistemological perspective intended to define procedures which can be repeated and compared, and to define, so to speak, a heuristics for discourse analysis. On the other hand, it is inscribed within an operational perspective which allows empirical findings to be obtained and which therefore offers a theoretical and methodological alternative to content analysis. For Pêcheux, it is not, then, merely a matter of developing the ADA69 apparatus internally by formalizing it. It is also a matter of computerizing it, of developing an instrument capable of meeting both the above demands (cf. Chapter 2).

Pêcheux's approach to formalization applies an algebraic framework (set theory, Boolean algebra, theory of graphs) rather than a logical framework. Certain elements are also borrowed from the domain of formal grammar (finite state automata, stacks and lists). Paul Henry recalls having worked at this time with Pêcheux on Berge's theory of graphs (Berge 1958) and on Chomsky and Miller's (1963) studies on the formal analysis of natural languages. Finally, the whole apparatus can be expressed in the form of algorithms which can be used directly for computer programming.

The attempt to automatize discourse analysis also coincided with contemporary French work on automatic translation, notably that carried out by the *Centre d'Études de Traduction Automatique* (CETA) in Grenoble. The strategy adopted for automatic translation at this time consisted of elaborating recognition grammars for source and target languages as a preliminary to translation. Pêcheux held that the elaboration of a French recognition grammar would make it possible to automatize the construction of elementary utterances, which is the first — manual — stage

in the ADA69 procedure, and to improve upon the rudimentary nature of the representation of the linguistic sequence, which he had already described as provisional. Similar concerns inspired the research carried out under the leadership of Pêcheux independently of CETA, which was designed to construct a morpho-syntactic analyzer for French. A partial analyzer dealing with functional forms was developed in 1971-72 (see Fuchs et al. 1972 for a description).

The late 1960s saw the introduction of computers in the human sciences in France (the *Centre de Calcul pour les Sciences Humaines* of CNRS was established in 1969), and Pêcheux became interested in computerized methods of analysis because he was critical of lexicometric and documentary analysis programmes like Gardin's SYNTOL (see Cros, Gardin and Lévy 1964) and *The General Inquirer* (Stone et al. 1966). The ADA69 programme was, together with the lexicometric programme designed by a team in St. Cloud, one of the first operational programmes in the field of computer-assisted textual analysis.

Three versions of the programme were developed. First, the ADA69 programme written in FORTRAN IV by Pêcheux and P. Duval was used at the CNRS's *Centre de Calcul pour les Sciences Humaines* from 1972 onwards. This version was also used at the *Université de Québec à Montréal* and, thanks to N. Pizarro, in Madrid. Pêcheux's continued and privileged links with the research being carried out in Grenoble led to the development of two versions of the ADA69 apparatus, both at the *Université de Grenoble II*. The first version was written in ALGOL W by P. Bizard and M. Dupraz in 1972. This version is identical to the FORTRAN IV version at the algorithmic level, and it gave rise to several applications. The second version was written in ALGOL W by C. Del Vigna in 1975. It made it possible to test certain reworkings of the algorithmic propositions, thanks mainly to its interactive configuration.

Almost twenty researchers in the human sciences who problematized their disciplines (sociology, psychology, social psychology, psycholinguistics and linguistics) within the framework of discourse theory, used ADA69 between 1971 and 1981 despite the labouriousness of the preliminary coding. This indicates the interest aroused by the novelty of the methodological and theoretical approach offered by ADA69. By 1980, the general inadequacy of the programme, which was only partly revised, to developments in discourse theory, together with its linguistic and even technical inadequacies, led Pêcheux and his team to look for an alternative to the ADA69 programme. A methodological breakthrough was made

possible by DÉREDEC, the software developed by P. Plante at the *Université de Québec à Montréal*, which included a syntactic analyzer for French. Unlike ADA69, which could only carry out well-defined tasks, DÉREDEC provided a programming environment which made it possible to carry out modular procedures, and it opened up new prospects in automatic discourse analysis (see Chapter 8).



## **PART II**

### **AUTOMATIC DISCOURSE ANALYSIS**





## Automatic discourse analysis

Michel Pêcheux

### 5.1. Linguistics and textual analysis: relations of proximity

Prior to recent developments in linguistic science, whose origins can be found in the *Course in General Linguistics* (Saussure 1983), studying a language usually meant studying *texts* and asking various kinds of questions about them. These questions were normative or descriptive, and they derived both from the classroom exercise which is still known as 'text interpretation'<sup>1</sup> and from the activities of grammarians. One asked 'What is this text about?' and 'What are the main "ideas" to be found in this text?'. At the same time, one also asked 'Does this text conform to the norms of the language in which it is written?' or 'What are the norms specific to this text?'. All these questions were asked simultaneously because they all relate to one another; to be more specific, questions concerning the semantic or syntactic usages revealed by the text helped to answer questions concerning the meaning of the text (what the author 'meant'). In other words, the classical science of language purported to be both a *science of expression* and a *science of the means of expression*; the study of grammar and semantics was a means to an end (understanding the text) just as, at the level of the text itself, means of expression were pressed into service to further the aims of the producer of the text (making oneself understood).

In these conditions, it is because all human beings are, to a certain extent, grammarians, that they can understand their fellow human beings, and it is only because the language specialist is, first and foremost, a human being capable of self-expression that he can carry out scientific

---

<sup>1</sup> According to Saussure, this is a form of philology in that its primary aim is to "establish, interpret and comment upon texts" (Saussure 1983: 1).

work.

The point of the conceptual displacement introduced by Saussure is that it destroys this collusive homogeneity between the practice and theory of language: once the language [*langue*] has to be thought of as a *system*, it ceases to be understood as having the *function* of expressing a meaning, and it becomes an object whose *functioning* can be described by a science. (To take up the chess metaphor Saussure uses to conceptualize the object of linguistics, we might say that we are no longer concerned with *the meaning of a single game*, but with *the rules that make the game as such possible* – regardless of whether or not a particular game has been played.)

The effects of this displacement are as follows: a 'text' can in no sense be a pertinent object for linguistic science, as it does not function; what does function is the language, that is, a body of systems permitting rule-governed combinations and substitutions of specific elements. The mechanisms involved operate in a dimension smaller than the dimensions of the text; the language [*langue*], defined as the object of science, can be contrasted with speech [*parole*], defined as the non-scientific residue of the analysis: "By distinguishing between the language itself and speech, we distinguish at the same time: (1) what is social from what is individual, and (2) what is essential from what is ancillary and more or less accidental" (Saussure 1983: 13-14). The study of language, which once purported to have the status of *a science of expression and of the means of expression* and which tried to handle large-scale phenomena, thus takes the position which is that of linguistics ever since.

But, as is the rule in the history of science, the retreat by which linguistics establishes its scientificity leaves uncovered the territory it has abandoned, and the questions which linguistics has to refuse to answer continue to be asked for both theoretical and practical reasons:

- 'What does this text mean?'
- 'What meaning does this text contain?'
- 'How does the meaning of this text differ from that of another text?'

These are all different forms of a single question, and various answers to them have been provided by what is termed *content analysis*, sometimes also known as *textual analysis*.

I will examine the various types of answer to be found in contemporary analytic practices, which I will classify according to the manner in which the terrain abandoned by linguistics is approached.

## Non-linguistic methods

There are, first of all, methods of analysis which are apparently unrelated to linguistics. These were the first to appear, and they were developed at almost the same time as the displacement described above was taking place, unnoticed because of the lack of historical distance. These methods therefore set themselves the task of answering the question in what might be termed a 'pre-Saussurean' manner. They stand outside contemporary linguistics, which does not mean that they do not rely upon concepts originating in linguistics – it simply means that those concepts are outdated in terms of contemporary linguistic theory.

### *The frequency-count method*

The term 'frequency-count method' designates the process of counting the number of times a linguistic sign (usually a word or lexical element) occurs within a sequence of fixed dimensions, and of defining a frequency which can be combined with other frequencies; this provides a comparability test that can be applied to several items in the same sequence, or to a single item found in several parallel sequences. The major advantage of this method is that it facilitates the development of adequate statistical methods for processing data<sup>1</sup>.

Here, the relationship with the domain of linguistics is reduced to a minimum: it might be said that the only concept of linguistic origin is that of the biunique nature of the relationship between signifier and signified, which allows the same thought-content to be noted whenever the same sign appears. But this concept belongs to a pre-Saussurean theoretical field, and contemporary linguistics relies to a great extent upon the idea that it is only because a given term has *several meanings* that it has *a* meaning within a language, and that is tantamount to denying that the relationship between signifier and signified is biunique.

An alternative way to formulate this criticism in definitive terms is to note that even if we increase the number of frequency-counts, we will still be unable to account for the organization of the text, or for the networks of relationships that exist between its elements; it is as though the

---

<sup>1</sup> The rank/frequency ratio (the Loup-Zipf-Mandelbrot law) is the most important result obtained with this method.

surface of the text were a population and as though different censuses were being carried out: the description of the population may well be as detailed as one could desire, but it tells us nothing about the *meaning-effects* which constitute the content of the text. The data that is collected is objective, but a price has to be paid for its objectivity, and it is difficult to use it for the purposes for which it was intended.<sup>1</sup>

### *Analysis by thematic categories*

The method we have just described is *infra-linguistic*: insofar as its object is a sort of textual *demography*, it is not concerned with the functioning of a system of elements, but simply with the existence of certain types of linguistic material. Its contribution to linguistic theory cannot be denied, but it provides no answer to the question of the meaning contained in the text, or to that of the difference between the meaning of one text and another.

Classic content analysis<sup>2</sup> — as described, for instance, by Cartwright (1953) — does attempt to provide an answer to this question. What it looks for in a text is a series of *meanings*, which the coder detects by identifying the indicators that are bound up with them; in other words, the functional relationship between signification and means of expression takes on its true importance here. The analysis takes place at a *supra-linguistic*

---

<sup>1</sup> It should, however, be noted that the analysis of co-occurrence (contingency analysis) does allow us to identify a particular type of relationship between elements (namely their simultaneous presence within a single unit of a text) (cf. De Sola-Poni 1959: 62f).

[Pêcheux refers here to the statistically-based methods which were increasingly used in France in the 1960s. In 1982, he will inscribe these criticisms within the context of the struggle against the 'quantitative objectivism' of the period. The reference to linguistics in ADA69 is, that is, part of an attempt to 'displace the question from the domain of the quantitative to that of the qualitative' (p. 190 below). But when he discusses the constitution of the corpus in 1975, Pêcheux does admit that 'a preliminary statistical analysis may be helpful to an initial formulation of hypotheses' (p. 146 below).]

<sup>2</sup> [For Pêcheux, the project of an automatic discourse analysis is to a large extent inscribed within the framework of a critique of and an alternative to techniques for content analysis, which were increasingly popular in the human sciences at that time and which had already been automatized in the form of methods for documentary analysis (SYNTOL, *The General Inquirer*; see p. 58 above). He published a number of articles on this subject (Pêcheux 1967, 1968 and 1973).]

level, as the aim is to arrive at the *meaning* of a segment of the text by traversing its linguistic structure; coding or characterizing a segment means placing it in an equivalence class defined by an analytic table of meanings. The criterion for classification is the coder's judgement as to the absence or presence of the predicate under consideration or as to the intensity of its presentation.

Judgement is therefore based upon *indicators* whose linguistic pertinence has not been established (a word, a sentence, a 'theme'), and it requires complementary psychological qualities such as sensitivity or adaptability, and an ability "to detect subtle differences of meaning but also to neglect differences that do not make a difference for a specific purpose" (Cartwright 1953: 461). In other words, this method presupposes that the coders have *acquired a certain culture*, that they have *learned to read*.<sup>1</sup> Leaving aside the well-known problem of inter-coder reliability, let me make what seems to me to be the essential point: in this perspective, analysis cannot be a sequence of objective operations producing a univocal result (and a coder who tried to simulate such objectivity would produce a routine, mechanical piece of work with no analytic validity), yet "[i]f the coding is to be carried out by a team of coders, it is necessary that they all come to apply the same definitions and frame of reference to the coding" (p. 462). We must assume the existence of an explicit or implicit consensus<sup>2</sup> between the coders as to the modalities of their reading; in other words, the text can only be analyzed within *a shared system of values which has a meaning for the coders and which constitutes their mode of reading*. This method imposes the relationship between expression and means of expression on the text, and it therefore also imposes its effect, namely an overlap between the theoretical function of the analyst and the practical function of the speaker (cf. pp. 63-64 above). The ultimate danger is therefore that the results of the analysis will reproduce the reading-grid that made it possible (whatever the degree of the coder's

---

<sup>1</sup> [This is the first reference to the concept of 'reading'. It is only in the last section of the text (p. 118 below) that the ADA69 project is problematized by the use of the term reading. The notion of reading gradually takes on a central importance in Pêcheux's work (e.g. pp. 144, 186 below), and comes to play a vital role in the research group ADELA (*Analyse de Discours et Lectures d'Archives*; Discourse Analysis and Readings of Archives).]

<sup>2</sup> Such a consensus can be reached through group discussion or by using a device such as a round robin.

probitly, sensitivity or accuracy) because of the phenomenon of participatory reflection that occurs between the *object* and the *method* designed to apprehend that object.<sup>1</sup>

### Para-linguistic methods

The methods described above are non-linguistic in that they do not deal with the specific level of the sign and derive from psychological or sociological methodologies, but there are also more recent methods which refer *overtly* to modern linguistics<sup>2</sup> and which provide a different answer to the question of the *meaning* contained in the text. We have here a paradox which has to be resolved: why is it that disciplines such as ethnography, literary criticism and mass communication studies appeal to linguistics in order to answer a question which arises on the very terrain linguistics abandoned when it constituted itself?

I would suggest that we can resolve the paradox as follows: the various disciplines mentioned recognize the basic theoretical fact which marks the beginning of linguistic science, namely the *transition from function to functioning*; moreover, they interpret that event not as a closure which makes it impossible to ask certain questions, but as a sign that a new possibility is being opened up, namely the possibility of making the displacement from function to functioning for a second time – this time at the level of the text. In other words, given that syntactic systems exist, they posit the hypothesis that mythical or literary systems also exist, that *texts, like language, function*: the epistemological homogeneity that is assumed to exist between the facts of language and phenomena with the dimensions of a text therefore guarantees that it is possible to use the same conceptual instruments. The relationship between paradigm and syntagma, for example, is extended to different functional, and therefore analytic,

---

<sup>1</sup> Automatization does not alter the question to any great extent. The *General Inquirer* method (Stone et al. 1966) consists of noting the occurrence of words and sentences corresponding to pre-existing categories which have been incorporated into the recognition programme. There are, of course, other programmes, and the analyst *chooses* one to suit his needs. In other words, he usually chooses it in accordance with the theoretical presuppositions governing his reading.

<sup>2</sup> Or, to be more accurate, to its concepts (the opposition between paradigm and syntagma, for example) or its instruments (generative grammars, transformational systems).

levels: these disciplines hope to achieve the ideal of linguistic scientificity by transposing the instruments of linguistics. But does this necessarily mean that they achieve their ideal? At this point we encounter the resistance specific to the level and the dimensions of the object: the disjunction between the theory of the language and the practice of the speaker may appear to have been established, but the disjunction between the theory of myth and the practice of myth<sup>1</sup> remains problematic. The work of one distinguished specialist leads one to wonder whether or not that disjunction can be conceptualized:

There is no real end to mythological analysis, no hidden unity to be grasped once the breaking-down process has been completed. Themes can be split up *ad infinitum* [...]. Consequently the unity of the myth is never more than tendential and projective and cannot reflect a state or a particular moment of the myth [...]. Myths, like rites, are 'in-terminable'. And in seeking to imitate the spontaneous movement of mythological thought, this essay, which is also both too brief and too long, has had to conform to the requirements of that thought and to respect its rhythm. It follows that this book on myths is itself a kind of myth. (Lévi-Strauss 1969: 5-6)

It seems that we find here a 'pre-established' harmony between the *producer of the myth* and the *analyst*, just as we previously found a pre-established harmony between the *speaking individual* and the *grammarian* (cf. p. 63 above); in other words, the 'functioning' of the text is still closely related to its function, and the displacement has not therefore been fully accomplished.

We must bring out all the implications of the fact that the object of an analysis is not normally brought into existence *by the desire of the analyst*, and the elucidation of this point appears to be one of the con-

---

<sup>1</sup> [The reference is obviously to semiological and semiotic approaches. Note that these were commonly used in France at that time. Although Pêcheux quotes Lévi-Strauss, he is obviously thinking of Barthes whose *Elements of Semiology* appeared in *Communications* 4 (1964), and perhaps of Todorov, whose anthology of texts by the Russian formalists appeared in 1965. This anthology included an important text by Propp (1968). See also the comments made on this topic in Chapter 7, p. 190.]

ditions of existence of a scientific semiological practice.<sup>1</sup> We find here the origin of the methodological difficulties relating to the constitution and closure of a 'corpus'; if the object of the analysis is not *conceptually* defined, as an element in a process whose structure has to be constructed, that object in fact remains an *object of desire*, and this has two implications. On the one hand, the constitution of the object depends upon whatever it is in the mind of the analyst that leads him to posit its existence. On the other, the analyst feigns to encounter it as a natural given, and that frees him from all responsibility.

The problem therefore concerns our mode of access to the object, and the conceptual guide-lines described below will be articulated around this point.

We have seen that, in analyzing myths, one has no norm to define what does and what does not belong to the corpus. When, to take another example, one is faced with a juridical or scientific text, this problem appears not to arise. In such cases, there is a juridical or scientific institution to which the text can be referred. We must therefore differentiate between a *documentary analysis* carried out within an institutional reference for purposes which correspond, in general, to the ends of that institution, and an analysis which could be described as 'non-institutional' – the analysis of a myth being one such example<sup>2</sup>. (The methodological convergence whereby methods for automatic documentation are applied to a non-institutional analysis should therefore occasion some surprise.) Documentary analysis in fact basically presupposes that *equivalence classes* are defined *a priori* by the institutional norm itself; speaking of the modalities of the storing of the data required for the analysis of a document, Gardin writes:

Whatever stance we adopt, the fact remains that we must first establish the norms we are dealing with; we must, that is, in some way or another constitute a 'classification'

---

<sup>1</sup> The psychoanalytic relationship is a special case in that the object of the analysis is brought into existence by and for the desire of the analyst *too*.

<sup>2</sup> [This can be seen as the first formulation of an opposition which will in later texts (Pécheux 1982b) take the form of a distinction between 'logically stabilized discursive universes' (e.g. scientific discourse) and 'non-logically-stabilized discursive universes' (specific to a socio-historical space).]



in which the position of each key word reflects the semantic relations in which it stands to other terms (for example, 'temporal lobe' is part of 'telencephalon') or groups of terms (for example, 'ataxia' is part of 'motor-control disorder'). (Cros, Gardin and Lévy 1964: 42)

We can now understand the importance of the essential preliminaries to any analysis, which Mounin states clearly:

[The analyst] constitutes, for every type of object, the symbol-code which registers the presence or absence of all the *distinctive features* of each type of object that is to be described and classified. Coding is therefore preceded by a technological analysis designed to establish an inventory of all the distinctive features required for the description of objects of that type, or in other words for the establishment of an exhaustive framework containing a definition of each object. (Mounin 1963: 114-115)

It is, then, because there is already an *institutionally guaranteed* discourse on the object that the analyst can rationalize the system of semantic features characterizing the object: the system of analysis has the same level of theoretical development as the institution which defines its norms, and it therefore allows the position of a particular content to be defined in relation to those norms. The work of Ackermann and Zygouris (1966) shows, for example, that it is possible to measure the progressive conformation of a group of subjects to the scientific norms imposed upon them by an educational institution.

Several questions arise as a result of this analysis. They can be formulated as follows:

- If we accept it as proven that no science dealing with the sign can be constituted unless it abandons the terrain of the *function* of expression and meaning and situates itself on that of *functioning*, *what type of functioning* is to be assigned to the object in question here?
- In what sense is the concept of 'institution' pertinent to the construction of the concept of that object?
- If we take 'text' to mean any organized linguistic object submitted for analysis, can we retain that concept to designate the object of an analytic practice which takes into account the answers to the previous two

questions?

## 5.2. Conceptual guidelines for a theory of discourse

### Theoretical consequences of certain Saussurean concepts

In Chapter 3 of Saussure's *Course* we find two *definitions* of the concept of the language. The first states the properties of the object to be defined: The language [*langue*] "is the social part of language [*langage*], external to the individual, who by himself is powerless either to create it or to modify it" (1983: 14). This definition contrasts the language, a concept that can be used scientifically, with speech, or that part of language which belongs to the individual to the extent that he can create and modify it. The second definition relates the object to other objects situated on the same plane:

A language [*langue*] [...] is a social institution. But it is in various respects distinct from political, juridical and other institutions. Its special nature emerges when we bring into consideration a different order of facts. A language is a system of signs expressing ideas, and hence comparable to writing, the deaf-and-dumb alphabet, symbolic rites, forms of politeness, military signals, and so on. It is simply the most important of such systems. It is therefore possible to conceive of a science *which studies the role of signs as part of social life*. It would form part of social psychology and hence of general psychology. We shall call it *semiology*. (1983: 15)

With this definition, Saussure effects a double division: he contrasts *one* semiological system ('the most important': the language) with the *set* made up of all semiological systems, which are regarded as having a potentially equivalent scientific status and as coming within the field of the regional theory of the signifier. But Saussure also introduces a further opposition by using the term institution: this allows him to divorce juridical, political and other institutions from the series of semiological institutional systems, and simply to exclude them from the field of the regional theory in question.

Saussure thus regards the language as a homogeneous scientific object (belonging to the region of 'the semiological') whose specificity is supported by two theoretical exclusions:

- the exclusion of speech, which is inaccessible to linguistic science;
- the exclusion of '*non-semiological*' institutions from the zone of pertinence of linguistic science.

Let us now elucidate the consequences of the two definitions we have described.

### *Implications of the Saussurean opposition between the language and speech*

The opposition between the language and speech belongs to the post-Saussurean linguistic tradition:

The antinomy between the language and speech is total. Speech is an *act*, and therefore an *actualized* manifestation of the aptitude for language. It presupposes a context, a concrete and determinate situation. The language, in contrast, is a virtual system which can only be actualized in and by speech. But it is still true to say that the two principles are interdependent: the language is merely the residue of countless speech acts, whilst the latter are merely the application, the utilization of the *means of expression* [emphasis added] supplied by the language. It follows that speech is an act or an *individual* activity which contrasts sharply with the *social* character of the language. (Ullmann 1952: 16, cf. 1951: 27-29)

This text brings out the consequences of the operation of exclusion effected by Saussure: even if he did not explicitly intend it to do so, it is a fact that this opposition authorizes the triumphal reappearance of the speaking subject as *subjectivity in action*, as an active unit of subjectivity with intentions that can be realized thanks to the means it has at its disposal; it is, in other words, as though linguistic science (which takes the language as its object) had set free a residue: the philosophical concept of a free subject who can be regarded as the unthinkable obverse of the system, as its necessary correlate. Speech, defined as the use of the

language, appears to be *a road to human freedom*: following this strange road, which leads from phonemes to discourse, means moving *gradatim* from the necessity of the system to the contingency of freedom, as Jakobson's text (1971: 242-243) suggests, despite the many qualifications he introduces:

Thus, in the combination of linguistic units there is an ascending scale of freedom. In the combination of distinctive features into phonemes, the freedom of the individual speaking is zero: the code has already established all the possibilities which may be utilized in the given language. Freedom to combine phonemes into words is circumscribed; it is limited to the marginal situation of word coinage. In forming sentences with words the speaker is less constrained. And finally, in the combination of sentences into utterances, the action of compulsory syntactical rules ceases, and the freedom of any individual speaker to create novel contexts increases substantially, although again the numerous stereotyped utterances are not to be overlooked.

Insofar as the language is defined by a body of *rules which are universally present* within the linguistic 'community', we can therefore see why the mechanisms which characterize it were initially sought at the level of the elementary combinations and substitutions without which no speech is possible, because they are its indispensable means. They were sought *at the bottom of the scale*, or at a hypothetical level *lower than that of the sentence*. Recent developments in certain areas of linguistic research (and above all the appearance of generative grammars) seem, however, to have pushed back these limits, and tend to constitute *a linguistic theory of the sentence*, without necessarily leaving the system of the language. Whereas Saussure thought that the language did not create anything, the workings of generative grammar reveal that *a form of non-subjective creativity is at work within the language itself*.

Are we to assume that linguistic science will therefore gradually extend its empire and that it will succeed in accounting for the whole of the 'scale' by using increasingly powerful combinatory instruments?

It appears that there is a fundamental difficulty here, and that it relates to the theoretical horizon of linguistics, even in its current form.

We can spell out this difficulty by saying that it is not certain that the theoretical object that allows us to conceptualize language [*langage*] is a homogeneous whole; it is possible that the conceptualization of phenomena at the top of the scale requires a shift of theoretical perspective, or a change of ground, which introduces concepts external to the region of contemporary linguistics. In my view, the now classical problem of the 'normality of the utterance' is an exemplary index of this difficulty: under current conditions, the functioning of a generative grammar presupposes a type of speaker who might be described as having been *neutralized*, as being bound up, that is, with the universal normality of 'canonic utterances' in which the position of equivalence classes (for example: animate subject + inanimate object) is defined *a priori* as being a property of the language. The 'anomalous utterance' is therefore defined in relation to what is assumed to be normal in the language. In many respects, this thesis appears to be extremely fragile, as the following example demonstrates. In an attempt to determine whether the sentence belongs to the order of the language or to that of speech, Saussure writes (1983: 123):

To the language, and not to speech, must be attributed all types of syntagmas constructed on regular patterns. [...] Exactly the same holds for sentences and groups of words based upon regular models. Combinations like *la terre tourne* ('the earth rotates'), *que vous dit-il?* ('what does he say to you?'), etc. correspond to general combinatory types, which in turn are based in the language on specific examples heard and remembered.

Take the sentence *la terre tourne*. A pre-Copernican linguist who was, by some miracle, familiar with generative grammars and the work of contemporary semanticists would certainly have posited an incompatibility between the constituent parts of the sentence, and would have declared the utterance anomalous.<sup>1</sup>

---

<sup>1</sup> [This passage obviously refers to the concept of grammaticality introduced by generative grammar. Pêcheux does not use the term itself, and speaks of the 'normality of the utterance'. In doing so, he is referring to the issue of 'acceptability to a speaker' rather than raising the question of the correct grammatical formation of utterances. The return to Saussure and the discussion of the example given do nothing to clarify the problem; the issue of combinatory regularity is replaced by the question of what can or cannot be said on the basis of the

In other words, a sentence cannot always be said to be normal or anomalous with respect to a universal *norm* inscribed in *the* language: the sentence must be referred to the specific *discursive mechanism*<sup>1</sup> which made it possible and necessary in a given scientific context. It seems, that is, essential to question the identity Saussure implicitly establishes between *universal* and *extra-individual* by demonstrating that it is possible to define a level which is intermediate between individual singularity and universality, namely the level of the *particularity* which defines linguistic 'contracts' specific to one or another region within the system, or bundles of 'norms' which are, to a greater or lesser extent, locally defined, and which are therefore unlikely to spread from one region to another. As Jakobson writes (1981: 21):

No doubt, for any speech community, for any speaker, there exists a unity of language, but this over-all code represents a system of interconnected subcodes; every language encompasses several concurrent patterns, each

---

scientific knowledge of a given period.

This discussion is not dissimilar to the debate in the 1970s between semanticists and generative grammarians over the acceptability of sentences such as 'John called Mary a Republican, and then she insulted him' (Lakoff 1971: 33), in which the use of 'insulted' would imply that it is an insult to call someone a Republican. (The debate in question arose in the 1970s, when Lakoff's work was published in French.)

<sup>1</sup> [By referring to a 'discursive mechanism', Pécheux begins to effect the displacement mentioned earlier by proposing an extralinguistic explanatory principle. The distinction between 'linguistic' and 'extralinguistic' will be reformulated in Pécheux (1982a) and in Chapter 6 (pp. 140-141) below, where a distinction is made between linguistic base and discursive process. But that formulation implies ideology. In 1969, Pécheux introduces the question of ideology by using philosophical terms (universal/universality, extra-individual, individual singularity) without referring to ideology as such. We are not in a position to say whether he is avoiding the issue, or whether this is an indication that his problematic is still incomplete. In the article on the Saussurean break (Haroche, Henry and Pécheux 1971), Pécheux introduces the concept of 'discursive formation', and relates it to that of 'ideological formation'; at that time, the question of the role of semantics within linguistics, which is a minor issue in ADA69, becomes central. As a result, Pécheux's uncritical acceptance of Jakobson's proposal that the 'unity of language' should be seen as 'a system of interconnected subcodes' (quoted here) makes way for a theoretical decision which precludes the possibility of resolving the question of 'discourses pronounced on the basis of different positions' by relating these discourses 'to different sub-systems within the language' (see Haroche, Henry and Pécheux 1971: 97).]

characterized by different functions.

The concept of a 'semantic field'<sup>1</sup> certainly represents a step in this direction, as it relates to semantic constraints at the level of morphemes, or to *in praesentia* and *in absentia* relations between them within a given area of meaning. It does not, however, account for the sequential effects that are bound up with discursivity. In other words, the concept of a semantic field can account for the first of the two meanings of the word 'rhetoric' (rhetoric as knowledge pertaining to the choice of words and 'mutual alliances' between words), but not for the second (rhetoric as knowledge pertaining to disposition, the order and sequences of ideas, etcetera). To adopt the terminology of logic, we might say that the local normality which controls the production of a given type of discourse concerns not only *the nature of the predicates attributed to a subject, but also the transformations those predicates undergo in discourse, and which lead discourse to its ends (and to its end).*

I propose to use the term *process of production* to designate all those formal mechanisms which produce a discourse of a given type in given 'circumstances'. It follows from the above that the study of discursive processes implies two kinds of research:

(1) The study of the specific variations (semantic, rhetorical and pragmatic) that are bound up with the specific *process of production* under consideration. These are to be studied against the 'invariant background' of the language (which means primarily syntax as a source of universal constraints). The concepts and the methodology used are further specified below (cf. pp. 95ff.).

---

<sup>1</sup> [Pêcheux gives no references here. The question of 'semantic fields' is raised by Mounin (1963: 71-74), who uses the term to cover the *sprachliche Feld* of Jost Trier and the German linguists, the 'area of meaning' of the Anglo-American school, Matore's *notional field* and Guiraud's *lexicological fields*. The notion of a 'semantic field' arises in contemporary discussions and serves to illustrate in contradictory fashion both the potential and the limitations of structural analysis (cf. Todorov 1966). The notion of a semantic field opens up a breach in structural homogeneity.]

Major studies in lexicology have been undertaken in an attempt to specify relations between the language and society in particular domains, Jean Dubois's thesis on the vocabulary of the Commune (Dubois 1962) being the archetypal example. The wealth of material produced in this area (cf. the *Cahiers de lexicologie* and *Langue française* 2 (1969)) forms the background for the emergence of discourse analysis, in both its ADA69 version and in Dubois's 'Harrisiatt' version.]

(2) The study of the connection between the 'circumstances' of a discourse – which we will in future refer to as its *conditions of production* (cf. pp. 82ff.) – and its process of production. In current linguistic theory, this perspective is represented by the role assigned to *context* or *situation*, to the specific background which makes it possible to formulate and understand discourses. We will now attempt to elucidate this aspect of the question by making a critical examination of the Saussurean concept of *institution*.

### *Implications of the Saussurean concept of institution*

According to Saussure, the language is one social institution<sup>1</sup> among others; it is therefore possible to stipulate the specific difference which situates it within the series of institutions, as though it were a species within a genus. Everything seems to be clear once we specify that this specific difference is called *the semiological*. In the *Course in General Linguistics* we find, however, another type of difference which once more involves 'other' institutions. It is essential to evaluate it critically. Saussure writes that:

Other human institutions – customs, laws, etc. – are all based in varying degrees on natural connexions between things. They exhibit a necessary conformity between ends and means. [...] A language, on the contrary, is in no way limited in its choice of means. (1983: 76)

We find here an indication of the reversal we described earlier. Saussure shows that the language cannot be defined in terms of a 'necessary conformity between ends and means' (teleological harmony); and in order to make it perfectly clear that what he is saying is new, he refers to the functional properties of other institutions as though they were *obvious*. In other words, it is because Saussure still thinks of institutions in general as means *adapted* to ends that he can show that the language is a special case in that it has no naturally determined means. Criticising Saussure for his ignorance of what the sociologists of his day were beginning to perceive is

---

<sup>1</sup> [See Chapter 6.1., pp. 124-132 below for a critique of the term institution, and its difference with 'ideological state apparatus'.]



of course out of the question; let us simply note that in the *Grande Encyclopédie Française*, Fauconnet and Mauss (1901: 168) define sociology as the science of institutions and add that "an institution is a fully established ensemble of acts and ideas which individuals find before them and which, to a greater or lesser extent, imposes itself upon them". Saussure could have accepted this definition as a description of the language [*langue*], that is of the 'social part of language [*langage*]'.<sup>1</sup>

It cannot in fact be denied that one of the most decisive results of contemporary sociology is the ability to make a distinction between the *apparent function* of an institution and its *implicit functioning*: the norms of social behaviour are no more transparent for the actors than the norms of the language are transparent for the speaker. "The objective meaning of their behaviour [...] possesses them because they are dispossessed of it" (Bourdieu 1990: 4). In retrospect, it appears that Saussure was influenced here by the necessary illusion of the non-sociologist, who regards institutions in general as functions with an explicit finality.<sup>1</sup>

This is not without its relevance to the theory of discursive processes. Take the example of a deputy speaking in the Chamber. From a strictly Saussurean point of view, his discourse belongs, as such, to the order of speech and exhibits the 'freedom of the speaker', even though it does of course belong to the order of the language insofar as it is a syntactically correct sequence. But the sociologist perceives the same discourse as *part* of a functioning mechanism, as, that is, belonging to a system of norms which are neither purely individual nor fully universal, but which derive from the structure of a political ideology and which therefore correspond to a certain *position* within a given social formation.

In other words, a discourse is always pronounced on the basis of given *conditions of production*. The deputy, for example, either belongs to a political party supporting the government, or to an opposition party; he is either a spokesman for one or another group representing one or another interest, or he is 'isolated'. He is, whether he knows it or not, situated within the *balance of power* that exists between the antagonistic elements of a given political field: the status of what he announces, promises or denounces varies in accordance with the position he occupies. The same declaration may be a fearful weapon or a ridiculous farce, depending upon the relationship between the orator's words, his position and what he

---

<sup>1</sup> We also find traces of the distinction between 'apparent function' and 'implicit functioning' in Merton (manifest function and latent function) and in Durkheim.

represents. A discourse can be a direct political act or an empty gesture designed to allay suspicion, which is another form of political action.<sup>1</sup> We might evoke the concept of the 'performative' introduced by Austin to underline the necessary relationship between a discourse and its position within an extra-linguistic institutional mechanism.<sup>2</sup>

If we pursue our analysis of political discourse — which is used here merely as an exemplary representative of various types of discursive processes — we find that it must also be related to the *relations of meaning* within which it is produced. Any given discourse refers to another discourse; it is a direct or indirect answer to it, and it 'orchestrates' its major themes or demolishes its arguments. In other words, the discursive process has no actual beginning: a discourse is also supported by pre-existing discursive material, which it uses as raw material, and the orator knows that, when he *evokes* an event that is already the object of discourse, he is reviving in the minds of his listeners the discourse in which that discourse was mentioned, together with the 'distortions' introduced by the current situation. He is also able to exploit those 'distortions'.<sup>3</sup>

This implies that the orator in some way senses where his listener stands as a result of his own position; his ability to imagine or anticipate

---

<sup>1</sup> [Note that the term 'discourse' has yet to acquire a clear theoretical status. Pécheux moves here from a theoretical definition (discourse as 'part of a functioning mechanism' deriving from the 'structure of a political ideology') to descriptions in which the specification of a discourse as 'political' has a purely empirical value. Even so, it has to be recalled that, given the conjuncture of the time, it required considerable courage for someone working in an academic institution to make use of concepts derived from Marxism and to refer to the need to situate discourse in terms of 'conditions of production' and 'balance of power'.]

<sup>2</sup> [This is Pécheux's only reference to Austin and the problematic of the 'performative'. It is clear that he had not worked on this issue; Austin does not appear in the original bibliography. A French translation of Austin (1962) was published in 1970. It was Benveniste (1966: 267-276) who introduced linguists to themes of analytic philosophy. That Slakta's essay on Austin did not appear until 1974 shows how long it took the theme of speech acts to penetrate French linguistic circles. Ducrot's first writings, which helped to popularize these themes, date from the period 1965-70.]

<sup>3</sup> [This is the first descriptive reference to a discourse being supported by an earlier discourse.]

the listener's expectations can be decisive.<sup>1</sup> The ability to predict *what the other will think* appears to be a constituent element of all discourse, with variations that depend as much on possible states of mental pathology pertinent to verbal behaviour<sup>2</sup> as on the reactions that are institutionally allowed to the listener: a rambling conversation and a sermon 'function' differently in this respect. In certain cases, the listener or audience can be expected either to *block* or to *support* a discourse by making either direct or indirect interventions, which may be either verbal or non-verbal.

A deputy speaking in the Chamber may, for example, be interrupted by an adversary who, because he occupies a different position (in other words, because his discourse corresponds to different conditions of production), may try to draw the orator on to his own ground, or may force him to answer a question on a subject he finds offensive. There is also a system of non-linguistic signs such as, in the case of parliamentary discourse, applause, laughter, uproar, booing and heckling, which allow the audience to make indirect interventions. These signs are usually gestures (acts at the level of the symbolic), but they can also spill over into direct physical interventions. Given the current state of the theory of the signifier, we unfortunately do not yet have a theory of *gesture as symbolic act*, and this leaves many problems unresolved.<sup>3</sup> What is significant when, for instance, 'anarchists' hurl bombs into the parliament: is it *a symbolic gesture* signifying an extremely brutal intervention or *an attempt at physical destruction* directed against a political figure who was judged harmful to their cause?

Many of the questions we have just evoked will remain unanswered here. It is not my intention to outline a *sociology* of the conditions of production of discourse, but to define theoretical elements which will allow us to conceptualize discursive processes in their generality; as a general proposition, we can say that *linguistic phenomena of dimensions larger than the sentence can indeed be considered as a 'functioning'*,

---

<sup>1</sup> Pagès (1955) notes that the sender follows and 'attunes' his discourse in accordance with his assumptions about 'a relatively determinate audience'. In certain cases, he adds, the sender is informed about the 'echo' of his previous messages on the receiver, and modifies his assumptions accordingly.

<sup>2</sup> On this point, see especially Irigaray (1985).

<sup>3</sup> It is, however, to be noted that Greimas (1968) does provide elements of such a theory.

provided that we immediately add that *their functioning is not entirely linguistic in the current sense of that term*, and that it can only be defined if we refer to the *mechanism* which gives the protagonists and the object of discourse their respective *positions*. We have termed this mechanism the 'conditions of production' of discourse.

We are advancing the hypothesis that a given state of conditions of production corresponds to a definite structure of the process of production of discourse within the language. This means that, under given conditions, the set of discourses that can be generated in those conditions will manifest semantic-rhetorical invariants that are stable within the set under consideration and characteristic of the process of production in question. As will be demonstrated below, this implies that *it is impossible to analyze a discourse as though it were a text*, as though, that is, it were a self-contained linguistic sequence; it is necessary to relate it to *all the discourses that can possibly be generated* on the basis of a given state of conditions of production.

We will therefore begin by proposing an operational definition of the *state of conditions of production*, and we will then describe the theoretical and methodological prerequisites for a representation of the *process of production* corresponding to any given state.

## Conditions of production of discourse

### *Structural elements of conditions of production*

An *extrinsic* description of linguistic behaviour in general (as opposed to an *intrinsic* analysis of the spoken chain) can be based on one of two rival families of schemata:

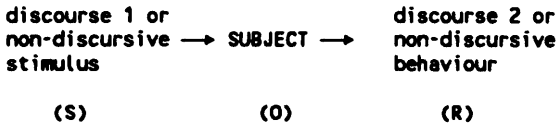
- (1) a 'reaction' schema deriving from psycho-physiological and psychological theories of behaviour (the 'stimulus-response' or 'stimulus-organism-response' schema);
- (2) an 'information' schema deriving from sociological and psycho-sociological theories of communication (the 'sender-message-receiver' schema).

Current thinking still appears to be dominated mainly by the former schema. As Moscovici and Plon (1966: 720) write:

The majority takes the view that the basis of language is

to be found in the organization of the nervous system, which is its material matrix, and not in communication, which is said to be its function. We can therefore say that theoretical progress at the psycho-sociological level is not enough and that current options must change if social psychology is to be brought into line with other psychological disciplines and if we are to promote an understanding of language.

The application of the stimulus-organism-response (or *S-O-R*) schema to verbal behaviour produces the following representation:



This representation has the disadvantage of masking both the position of the producer of (S) and that of the addressee of (R); this is perfectly legitimate when the stimulus is *physical* (for example, a change in the intensity of light) and when the response is *organic* (for example, a variation in the response recorded by the E.E.G.). In that case, the experimenter simply constructs an apparatus which, experimental artifacts aside, functions independently. In an experiment on 'verbal behaviour', on the other hand, the experimenter is part of the *apparatus*, whatever the modality of his *presence* — physical or otherwise — within the conditions of production of the response discourse; in other words, the stimulus is a stimulus only insofar as it refers to the 'verbal communication' situation in which a provisional pact is established between experimenter and subject. Moscovici and Plon note in this connection that:

The Skinnerian attitude implies the exclusion of the effect of the rules and norms which individuals establish between one another from the examination of human behaviour in general and of linguistic behaviour in particular. It therefore also minimizes the symbolic dimension language acquires through its association with those rules, and the role, which is far from negligible, it plays

in their constitution. (1966: 718)

In other words, the *S-O-R* schema implies too many theoretical 'forgettings' for it to be retained in its present form.

The information schema, on the other hand, has the advantage of introducing both the protagonists and the 'referent' of discourse. In his concise survey of 'the constitutive factors in any speech event, in any act of verbal communication', Jakobson (1981: 21-22) writes:

The *addresser* sends a *message* to the *addressee*. To be operative the message requires a *context* referred to (the 'referent' in another, somewhat ambiguous, nomenclature), graspable by the addressee, and either verbal or capable of being verbalized; a *code* fully, or at least partially, common to the addresser and addressee (or in other words, to the encoder and decoder of the message); and, finally, a *contact*, a physical channel and psychological connection between the addresser and the addressee, enabling both of them to enter and stay in communication.

The schema thus becomes:



where

A = 'addresser'

B = 'addressee'

R = 'referent'

(L) = linguistic code common to A and B

→ = contact established between A and B

D = verbal sequence sent by A to B

With respect to the verbal sequence D, it will be noted that information theory, which provides the basis for this schema, leads Jakobson to describe the information that is transmitted as a *message*; in view of what has been said above, I prefer to use the term *discourse*, which implies that we are not *necessarily* dealing with information transmitted from A to B but, more generally, with a 'meaning-effect' that occurs between points A

and B.

We can now specify the various structural elements of the conditions of production of discourse.

It is, first of all, quite clear that the elements A and B designate something other than the physical presence of individual human organisms. If what has been said above means anything, it implies that A and B designate determinate positions within the structure of a social formation, and that sociology can describe the bundle of objective features that characterize them; within the sphere of economic production, for example, the positions of the 'boss' (director, factory manager), the executive, the foreman and the labourer are marked by identifiable differential properties.

My hypothesis is that these positions are *represented* within the discursive processes in which they are brought into play. It would be naive to imagine that a *position* (as a bundle of distinctive features) functions as such in the discursive process. It is represented within it, that is: it is *present, but transformed*; in other words, what functions in the discursive process is a series of imaginary formations designating the positions which A and B ascribe to *themselves* and to *one another*; the image they have of their own position and of the position of the other. If this is the case, rules for projection must exist within the mechanisms of any social formation, and they must establish both *situations* (which can be defined objectively) and *positions* (representations of situations). It should be added that it is highly probable that this correspondence is not biunique, that different situations can therefore correspond to a single position, and that a single situation can be represented by several positions. This is not a random process, and it is governed by laws that only a sociological investigation can reveal.<sup>1</sup> All that we can say is that any discursive process presupposes the existence of imaginary formations which can be designated as in figure 1 on the next page, which outlines the manner in which the position of the protagonists of a discourse intervenes as a condition of production of discourse. We have to add that the 'referent' (the 'context' or 'situation' in which the discourse appears) is also a condition of production. This is expressed by R in figure 2. It should again be stressed that the referent is an *imaginary object* (a subject's viewpoint), and not a physical reality.

---

<sup>1</sup> [Pêcheux criticizes the sociologism inherent to this passage in Chapter 6, p. 125 below.]

Figure 1

	Expression designating imaginary formations	Signification of expression	Implicit question (the answer to which subtends the corresponding imaginary formation)
A	$I_A(A)$	Image of position A for the subject placed at A	'Who am I to talk to him like that?'
	$I_A(B)$	Image of position B for the subject placed at A	'Who is he for me to talk to him like that?'
B	$I_B(B)$	Image of position B for the subject placed at B	'Who am I for him to talk to me like that?'
	$I_B(A)$	Image of position A for the subject placed at B	'Who is he to talk to me like that?'

Figure 2

	Expression designating imaginary formations	Signification of expression	Implicit question (the answer to which subtends the corresponding imaginary formation)
A	$I_A(R)$	A's 'view' of R	'What am I talking to him about like that?'
B	$I_B(R)$	B's view of R	'What is he talking to me about like that?'



Finally, we have already noted (pp. 80-81 above) that any discursive process presupposes that the sender *anticipates the receiver's representation*, and that the strategy of his discourse is based upon his expectations. We will therefore form the expressions:

A	$I_A(I_B(A))$	B	$I_B(I_A(B))$
	$I_A(I_B(B))$		$I_B(I_A(A))$
	$I_A(I_B(R))$		$I_B(I_A(R))$

They express the manner in which A represents B's representation at any given moment of discourse, and vice versa.

It should be noted that, given that we are, hypothetically, talking about *expectations*, these values exist *prior to* B's possible replies, which sanction A's anticipatory decisions. A's expectations of B must, for example, therefore be conceptualized as derivatives of  $I_A(A)$ ,  $I_A(B)$  and  $I_A(R)$ . We can symbolize this derivation by means of the following expressions, which, for the moment, serve simply to clarify our hypotheses as to the specific nature of the derivation in each case:

$$\begin{aligned} I_A(I_B(A)) &= f(I_A(B)).(I_A(A)) \\ I_A(I_B(B)) &= g(I_A(A)).(I_A(B)) \\ I_A(I_B(R)) &= h(I_A(R)).(I_A(B)) \end{aligned}$$

We can see that in each case A's expectations of what B will say depend upon the 'distance' A imagines there to be between A and B; we can make a formal distinction between discourses in which the orator tries to *transform the listener* (attempts to persuade, etc.) and discourses in which *orator and listener identify with one another* (the phenomenon of cultural complicity, 'winks' signifying agreement, etc.).

It follows from the above that state  $n$  of the conditions of production of a specific discourse  $D_x$  addressed by A to B on the referent R — which we will write as  $\Gamma_x^n(A, B)$  — can be represented by the following

schema<sup>1</sup>:

$$\begin{array}{ll}
 \Gamma_x^n(A, B) & I_A^n(A) \\
 & I_A^n(B) \\
 & I_A^n(R) \\
 & I_A^n(I_B^n(A)) \\
 & I_A^n(I_B^n(B)) \\
 & I_A^n(I_B^n(R))
 \end{array}$$

This calls for a number of comments. The first concerns the *nature* of the elements of the above schema. It has already been pointed out that they are *imaginary representations* of the various instances of the discursive process; we can now further specify our hypothesis as to their nature by adding that these various formations are themselves the result of earlier discursive formations (deriving from other conditions of production) which no longer function but which have resulted in the adoption of implicit positions guaranteeing the possibility of the discursive process under consideration. Whereas the 'phenomenological' thesis posits the view that the *perceptive apprehension* of the referent, the other and the self are *prediscursive conditions of discourse*, we are assuming that perception is always traversed by an 'already-heard' and an 'already-said' which gives the imaginary formations that have been uttered their substance. The concepts of *presupposition* and *implication* described and used by Ducrot (1966: 20-21) bring into play a similar hypothesis.<sup>2</sup> Describing a situation

<sup>1</sup> It will be noted that some rhetorical figures (both syntactic and semantic) can refer explicitly to elements of  $\Gamma_x^n$ . For example:

$I_A^n(I_B^n(A))$ : *You will think I am being indiscrete*

$I_A^n(I_B^n(R))$ : *"How strange", you will say*

This does not necessarily mean that any fragment of discourse can be referred in univocal fashion to any one instance. The question of whether or not these higher level expressions are pertinent to the problems under consideration has been ignored here.

<sup>2</sup> [The reference to Ducrot (1966), whose work on presupposition was beginning to be discovered by linguists at this time, provided a linguistic basis for the hypothesis that a discourse is supported by a previous discourse, a hypothesis which is formulated here in terms of the 'already-said'. This has to be seen as the starting point for a whole new

which 'can no longer be conceived in purely chronological or geographical terms as a spatio-temporal localization', he writes: "The 'knowledges' the speaking subject ascribes to his listener are an integral part of the 'discourse situation' to which the presuppositions refer. It therefore concerns the image the participants in the dialogue have of one another."

It is also clear that, in a given state of conditions of production of discourse, the elements constituting that state are not simply juxtaposed; relations between them may vary, depending on the nature of the elements that are brought into play. It seems possible to advance the view that the elements of  $\Gamma_x^n$  do not all have the same efficacy, and that one element may become *dominant* within a given state of conditions. The rules governing this have yet to be defined.  $\Gamma_x^n$  thus proves to be an ordered sequence, perhaps of a vectorial type, in which certain terms have the property of determining the nature, value and position of other terms.

Take, for example, a series of discourses characterized only by the fact that they all deal with the question of 'freedom'. Depending on whether we have a therapist addressing a patient, a philosophy teacher talking to students, or a prison governor explaining regulations to detainees, we find a *displacement of the dominant element* in the conditions of production of discourse. In therapeutic discourse, as conceived by classical psychiatry, it is the patient's self-image,  $I_B(B)$ , that is at stake in discourse. In the pedagogic relationship, it is the students' image of what their teacher is telling them,  $I_A(I_A(R))$ , that dominates discourse (in relation to  $I_A(R)$ ). Finally, in the discourse of the prison governor, everything is conditioned by the image the prisoners construct of the representative of the rules,  $I_A(A)$ , on the basis of his discourse. For the prisoners, it is a matter of knowing 'what they can get away with'; for the governor, it is a matter of letting them know 'how far they can go'.

Within this perspective, the object of a sociology of discourse is to identify the links between *power relations* (which are external to the discourse situation) and the *relations of meaning* manifested within it by

---

problematic. The 'preconstruct' (elaborated by Pêcheux and Paul Henry and described by the latter in Henry 1975 and 1977) provides an alternative to 'presupposition'; when it is separated out from Ducrot's problematic of communication, it can be seen as a trace of utterances pronounced elsewhere. It is part of a conceptual set made up of 'discursive formation', 'ideological formation', 'intradiscourse' and 'interdiscourse'. It is also to be related to the theorization of the 'two forgettings' (pp. 138-140 below) and to 'Three stages of discourse analysis' (Chapter 9 below).]

systematically bringing out the *variations in dominance* which we have just mentioned.

*Outline of a formal representation of discursive processes*

As we stated earlier (p. 82 above), we are advancing the hypothesis that, given the conditions of production of discourse  $D_x$  in state  $n$ , i.e.  $\Gamma_x^n$ , it is possible to make them correspond to the process of production of  $D_x$  in state  $n$ . That process will be designated  $\Delta_x^n$ . But we have also seen (pp. 88-89 above) that a given state of conditions of production must be understood as resulting from sedimented discursive processes; we can therefore see that it is impossible to define the origin of conditions of production. An *origin* is literally *unthinkable*, as it would imply *infinite regress*. It is, on the other hand, possible to investigate transformations of conditions of production on the basis of a given state of conditions. We will deal in turn with two questions:

- (1) the question of the correspondence between  $\Gamma_x^n$  and  $\Delta_x^n$
- (2) the question of the transformation  $\Gamma_x^n \rightarrow \Gamma_x^{n+1}$

The abstract operations we are about to apply to the elements defined above make it possible, I believe, to outline a formal description of discursive processes. The formulations given below are incomplete and provisional; my present aim is simply to demonstrate the possibility of such a theory in general terms, and to situate the particular case represented by my current work in relation to more complex phenomena, which will be left on one side for the moment.

RULE 1

*The process of production of discourse  $D_x$  (in state  $n$ ) results from the compounding of the conditions of production of  $D_x$  (in state  $n$ ) with a given linguistic system  $L$ .*

We will adopt the symbol  $\circ$  to designate this operation of compounding, and we will write:

$$\Gamma_x^n \circ L \rightarrow \Delta_x^n$$

This rule can be interpreted as follows:  $\Gamma_x^n$  operates as a principle of selection and combination on elements of the language  $L$ , and it uses them

to constitute a system of semantic links representing the matrix of discourse  $D_x$  in state  $n$ , or in other words, to constitute *semantic domains* and *dependencies between domains*. It should be added that this operation in fact reveals a number of hierarchical levels. As will be demonstrated below (p. 102), the constitution of an utterance – an elementary sentence – does not obey the same semantic, rhetorical and pragmatic laws as the disposition of utterances in a discursive sequence.

Although he operates on the basis of different theoretical premises, an interesting parallel can be found in the work of Doležel (1964: 51), who writes:

By using elementary units of the code and the rules of the code, the source of linguistic information – the encoder – produces concrete messages – discourses – which are a representation of sets of extra-linguistic events and which transmit information about those events.

Our earlier theoretical remarks should have alerted the reader to the divergence of opinion that can be noted here: the concepts of *information*, *message* and *extra-linguistic event* in particular derive from empiricist presuppositions, and I have already pointed out the difficulties they pose. Doležel's project is, however, in many respects illuminating for our purposes, as he goes on:

The fundamental unit obtained as a result of the process of encoding is the sentence; the sentence or a sequence of sentences constitutes the linguistic message, the discourse. [...] We must establish and specify the set of rules whose application allows words to be organized into sentences, and sentences to be organized into messages, during the process of encoding. (Doležel 1964: 52)

It should be added that the author explicitly advances the hypothesis of the stationary character of the 'parameters of the language', which will also be adopted here.

## RULE 2

*The compounding of any process of production  $\Delta_j^i$  with a determinate state  $n$  of the conditions of production of*

*discourse  $D_x$  results in a transformation of that state.*

We will adopt the symbol  $*$  to designate this operation of compounding, and we will write:

$$\Gamma_x^n * \Delta_y^j \rightarrow \Gamma_x^{n+1}$$

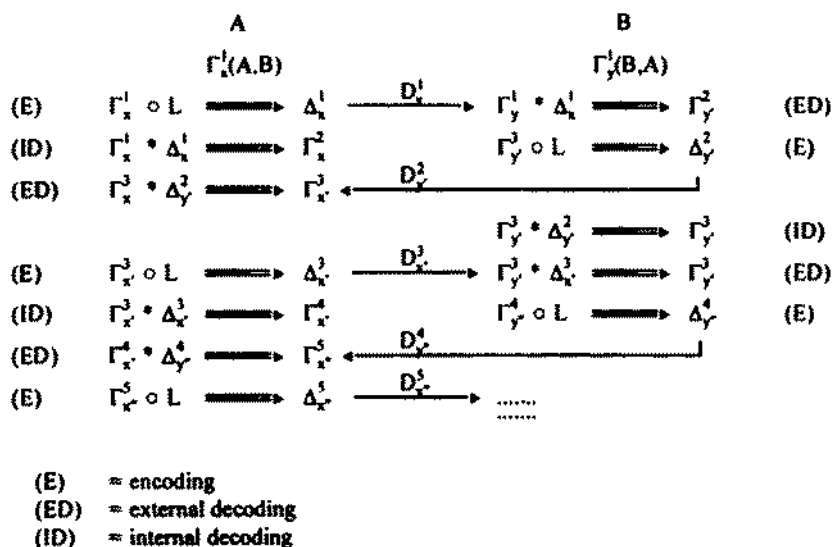
This rule brings out the transformational effect on the state of conditions of production that results from the presence of a particular process within the discursive field; it is clear, first of all, that a discourse addressed by A to B modifies the state of B insofar as B can compare with his 'anticipations'.

But any orator is potentially his own listener, and that implies that what is said by A also transforms the conditions of production specific to A by allowing him to pursue his discourse; 'disorders of narrative behaviour' (such as losing the thread of the narrative, or constantly going back to the beginning) can be interpreted as a disturbance of this mechanism.

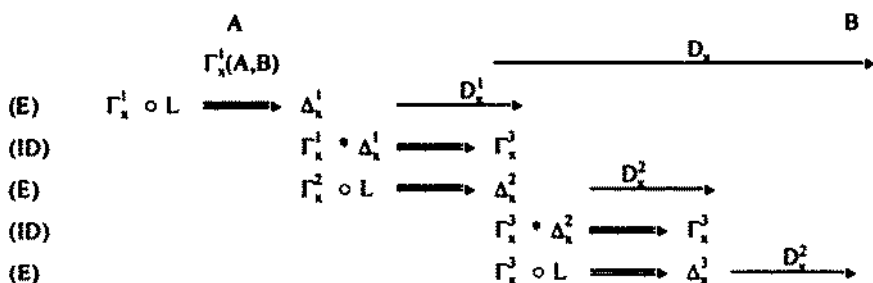
Both rules call for certain comments. In the first place, we can see that Rule 1 corresponds to the *sending* of the discursive sequence, whereas Rule 2 concerns its *reception*; they play, that is, a role comparable to that of the operations which are often described as *encoding* and *decoding* respectively. It is, however, to be noted that the opposition between language and reality, which often provides the basis for these two concepts, does not function here, and that the symmetry between encoding and decoding, which is often evoked as a necessity, also disappears.

In the second place, Rule 2 (the 'decoding rule') can take one of two functional modalities, which I propose to call external decoding and internal decoding. Any discourse situation involves internal decoding, but the existence of external decoding is bound up with the 'response' addressed by the addressee B to the initial addresser A. In certain discourse situations (writing a letter, a broadcast speech) that response may well be absent. We will now present the *general case* and then contrast it with this *particular case*.

A discourse situation in which A and B 'reply' to each other can be represented as follows. We can see that, at each 'step', the discourse of one of the protagonists is modified by that of the other:



Let us now, in contrast, consider a discourse in which the addresser A receives no response (no discourse and no symbolic gesture) from the addressee B.



We have here a particularly simple case because, as can be seen from the above, the series of states  $\Gamma_x^i$  can be deduced from  $\Gamma_x^1$  and because discourse  $D_x$  can be assimilated to the sequence

$$D_x^1 + D_x^2 + \dots + D_x^n$$

Under these conditions we can say that  $\Gamma_x$  (which integrates  $\Gamma_x^1, \Gamma_x^2, \dots, \Gamma_x^n$ ) is the condition of production of discourse  $D_x$  (which integrates  $D_x^1, D_x^2, \dots, D_x^n$ ) and that it corresponds to the process of production  $\Delta_x$  (which integrates  $\Delta_x^1, \Delta_x^2, \dots, \Delta_x^n$ ).

We will deal here only with the particular case of *monologic discourse*, which we can identify in such activities as narrative, prayer, testimony and demonstration, with the case, that is, where the addressee is present in the situation only insofar as the addresser has an image of him. The analysis of dialogic situations and of situations in which a third party may be present would require us to take into account more complex relations (several interacting conditions of production), which would imply further research. For the moment, our problem is the following:

Given a defined state of conditions of production ( $\Gamma_x$ ) of a monologic discourse ( $D_x$ ), and given a finite set of empirical discursive realizations  $D_{x1}, D_{x2}, \dots, D_{xn}$ <sup>1</sup> which represent  $\Gamma_x$ , how to determine the structure of the process of production ( $\Delta_x$ ) corresponding to  $\Gamma_x$ , that is, the set of *semantic domains* in play in  $D_x$  and their relations of dependence.

We will assume that it is possible to determine empirically a set of senders identifiable with respect to the state of the conditions of production of  $D_x$  (but not, of course, with respect to all discourse in general). As we have already pointed out (p. 85), the constitution of this set implies both the monitoring of the *objective sociological variables* characterizing the position of the addresser and the monitoring of the *imaginary formations* specific to the situation of  $D_x$ .<sup>2</sup>

---

<sup>1</sup> It should be noted that the designation of a particular realization of  $D_x$  (such as  $D_{x1}, D_{x2}$ , etc.) is not to be confused with that of a sub-sequence (such as  $D_i^j$ ) corresponding to a  $\Gamma_i^j$  state of conditions of production.

<sup>2</sup> The latter consists of verifying their content by asking implicit questions such as 'Who am I to talk to him like that?' (see p. 86).



## Towards an analysis of the process of production of discourse

Classical structural linguistics described its approach as follows: given a corpus of facts of the language, we must find notions and relations which allow us to make a simple, exhaustive and non-contradictory description of those facts. The theory of generative grammar inverts this relationship; it asks: what linguistic rules are consciously or unconsciously applied to produce correct sentences in a given language? Analysis gives way to synthesis; we are therefore dealing with a system of rules rather than a system of elements. (Todorov 1966: 5)<sup>1</sup>

Let us assume that the results of this *Copernican revolution*, which organizes the language around a 'speaking subject', are directly applicable to discourse theory; if that were the case, our primary objective would be to obtain a set of rules capable of generating *discourse*, and we could quite unproblematically forego any analysis of *the surface effects of the discursive sequence*, as such an analysis would be an outdated *Ptolemaic* preoccupation. But according to our hypothesis it is, as we have seen, illegitimate to transpose results from a 'speaking subject' (who has been *neutralized* with regard to the conditions of production of discourse) to a hypothetical 'subject of discourse'; what has been said above in fact presupposes that there is no *universal psychological subject* capable of supporting the production of all possible discourses in the sense that the subject represented by a generative grammar is capable of generating every grammatically correct sentence in a language. In other words, I regard the *methodological continuity* that is sometimes assumed to exist here as suspect in that the transition from the subject of language to the subject of discourse implies the existence of *selection* rules functioning at the level of 'terminal vocabulary', and in that those rules in fact involve the *analysis* of morphemes in terms of *semantic features*, even though it is generally recognized that the nature of such elements is highly problematic. This means that an *analytic digression* is necessary, but the analysis in question

---

<sup>1</sup> [The quotation from Todorov is taken from the opening article of the first issue of *Langages* (*Recherches Sémantiques*, 1966). It is part of a discussion of the 'impasse facing structural semantics', and is representative of the hopes raised in some quarters by generative theory in the linguistic conjuncture of the period 1965-70.]

usually remains implicit and non-systematic: it normally relies upon an *atomistic conception of signification*, and lexemes or morphemes are therefore analyzed arbitrarily, as though they were units that could be broken down into 'semes' which exist in themselves ('componential analysis'). Combinatory properties are deduced from rules of intersemic compatibility which are posited in an equally arbitrary manner.<sup>1</sup> It also seems that, in this domain, the principle of 'no elements; only relations and rules' has been forgotten.

Under these conditions, and given that an analytic digression appears at the moment to be inevitable, I think it preferable to state its principles explicitly. The series of discursive surfaces  $D_{x1}, \dots, D_{xn}$  constitutes a trace of the process of production  $\Delta_x$  of discourse  $D_x$ . It constitutes, that is, a trace of the 'deep structure' common to  $D_{x1}, \dots, D_{xn}$ . Our task is therefore to work backwards from these 'surface effects' to the invisible structure which determines them; only when we have done that will it be possible to realize a general theory of processes of discursive production as *theory of the rule-governed variation of 'deep structures'*.

### *The metaphoric effect*

Our question here is how to find at least one discourse in which two given terms (say  $x$  and  $y$ ) can be substituted for one another without altering the interpretation of that discourse. Let  $x$  and  $y$  be terms belonging to the same grammatical category in a given language  $L$ , let  $D_n$  be series of terms generated by  $\Delta_n$  in language  $L$ . Let  $D_n$  correspond to state  $\Gamma_n$ , which

---

<sup>1</sup> It is to be noted that Chomsky himself is more circumspect and cautious about this point than many of the theorists who have been inspired by his thought. It is also possible to envisage the constitution of a non-taxonomic semantics. [Though Pécheux adopts the designation 'Copernican Revolution' used in France to salute Chomsky's model, he indicates why his project and that of generative grammar are incompatible. Pécheux rejects the postulate of a universal psychological subject supporting the process of production of all possible discourses. Whilst Pécheux was tempted by the fantasy of a mechanism that could generate discourse, he argues in terms of the *analysis* of the processes of production of discourse. Another major difficulty of Chomsky's model is formulated here for the first time via the critique of atomistic conceptions of meaning. The underlying question is that of a 'universal semantics', of the role of semantics in linguistics. The same question arises in a pronounced form in Ilaroché, Ilenry and Pécheux (1971) and again on pp. 134-136, 158, and 176-186 below, as well as in Pécheux (1982a).]

is one of a set of possible states. Logically, three cases are possible:

- (1) x and y can never be substituted for one another,
- (2) x and y can sometimes be substituted for one another,
- (3) x and y can always be substituted for one another.

Let us consider cases (2) and (3), in which substitution is possible. In case (2), x and y are substitutable in a given context. For example:

x = *brillant* [brilliant]  
y = *remarquable* [remarkable]

In certain contexts, x and y are substitutable. For example:

*ce mathématicien est (x/y)*  
[this mathematician is (x/y)]

or:

*la démonstration de ce mathématicien est (x/y)*  
[this mathematician's proof is (x/y)]

But there are other contexts in which x and y are not substitutable. For example:

*la lumière brillante du phare l'aveugla*  
[the *bright* light from the lighthouse blinded him/her]

or:

*cette courbe comporte un point remarquable*  
[this curve reveals a *remarkable* point]

In case (3), in contrast, x and y are substitutable, *whatever the context*. For example:

x = *refrénér* [to suppress]  
y = *réprimer* [to suppress]

It is problematic to imagine a context in which substitution of these terms is impossible. Before a final decision is made to classify the *refrénér/réprimer* pair as an instance of case (3), however, all the discursive contexts possible in the given language must be examined. In other words, it is possible to discover whether or not an x/y pair belongs to case (2) *in a finite period of time*, but this is obviously not so with case (3). We will

use the term *local* or *contextual* synonymy to designate possibility (2), as opposed to possibility (3), which we will term *non-contextual synonymy*.

We can see that, if we have a finite set of discourses corresponding to a single state  $\Gamma_n$ , we must take the precaution of regarding all synonymies as contextual until it has been proven that certain of them remain constant throughout all the variations on  $\Gamma_n$  we have *studied*. It would thus seem that non-contextual synonymy is a borderline case, and that a contextual synonymy that has been verified against an increasing number of conditions of production tends to move towards that borderline. This relates to the question of *non-empty semantic intersections*. Personally, I would take the view that contextual synonymies are the rule and non-contextual synonymies the exception, which can be argued for on the basis of Saussure's theory of value:

In a given language, all the words which express neighbouring ideas help define one another's meaning. Each of a set of synonyms like *redouter* ('to dread'), *craindre* ('to fear'), *avoir peur* ('to be afraid') has its particular value only because they stand in contrast with one another. If *redouter* ('to dread') did not exist, its content would be shared out among its competitors. (Saussure 1983: 114)

It should be noted that it is in fact possible to envisage contextual synonymies between two groups of terms or two expressions which produce the same meaning-effect in relation to a given context.<sup>1</sup>

The 'sliding of meaning' between *x* and *y* is constitutive of the 'meaning' designated by *x* and *y*. Therefore we will call the semantic phenomenon produced by a contextual substitution a *metaphoric effect*; this effect is characteristic of 'natural' linguistic systems, as opposed to codes and 'artificial languages' in which a meaning is established in terms of a 'natural' metalanguage. In other words, a 'natural' system has no metalanguage which can be used to define its own terms: it is its own metalanguage.

We can thus see that it is absolutely necessary to have a series of

---

<sup>1</sup> [Although Pêcheux does refer to Saussure, the problem of contextual synonymy relates to the problematic of Harris, who is not mentioned here. Elsewhere Pêcheux does refer to Harris (see p. 175 below, where he refers to taking the possibilities opened up by Harris 'to their logical conclusion').]

sequences representative of any given set of conditions of production  $\Gamma_x$  if we are to discover the *semantic anchoring points* defined by intersections between metaphors. Let me explain this point by using an example whose *empirically improbable* character does not detract from its theoretical significance. Take state  $\Gamma_x$  and corpus  $C_x$ , a corpus of discourses that are truly representative of that state:  $C_x = D_{x1}, D_{x2}, \dots, D_{xn}$ . Let us use letters to designate the words making up the discourse under consideration; a different letter corresponds to a different word, and vice versa. Let the sequences of the  $n$  discourses be:

$D_{x1}$	<i>a</i>	<i>g</i>	<i>d</i>	<i>b</i>	<i>h</i>	...	<i>y</i>
$D_{x2}$	<i>j</i>	<i>g</i>	<i>d</i>	<i>b</i>	<i>h</i>	...	<i>y</i>
$D_{x3}$	<i>j</i>	<i>k</i>	<i>d</i>	<i>b</i>	<i>h</i>	...	<i>y</i>
$D_{x4}$	<i>j</i>	<i>k</i>	<i>m</i>	<i>b</i>	<i>h</i>	...	<i>y</i>
$D_{x5}$	<i>j</i>	<i>k</i>	<i>m</i>	<i>x</i>	<i>h</i>	...	<i>y</i>
$D_{x6}$	<i>j</i>	<i>k</i>	<i>m</i>	<i>x</i>	<i>w</i>	...	<i>y</i>
.							
.							
.							
$D_{xn-1}$	<i>j</i>	<i>k</i>	<i>m</i>	<i>x</i>	<i>w</i>	...	<i>y</i>
$D_{xn}$	<i>j</i>	<i>k</i>	<i>m</i>	<i>x</i>	<i>w</i>	...	<i>z</i>

It can be seen that it is assumed that the sole difference between each discourse  $D_{xi}$  and the preceding discourse  $D_{x(i-1)}$  is a *single substitution*, and that the whole context remains constant. We have therefore a series of metaphoric *effects* (*a/j*, *g/k*, *d/m*, etc.) which have the effect of preserving a semantic anchoring point on the surface of the text as the borderline case  $D_{xn}$  no longer contains *any* of the terms belonging to  $D_{x1}$ , though it is, by definition, its semantic equivalent.

The only function of this perfectly fictitious – and perfectly impossible – example is to bring out what is meant by the preservation of an invariant across morphemic variations: the same system of representations is reinscribed across *variants* which repeat it step by step. In my view, it is the repetition of the identical across necessarily diverse forms that characterizes the mechanisms of a process of production: the ‘deep structure’ thus proves to be a tissue of interdependent elements supported and guaranteed by the metaphoric effects which permit the generation of an almost infinite series of ‘surfaces’, but which at the same time impose

functional limits beyond which the 'deep structure' would explode.<sup>1</sup>

Under these conditions, the reciprocal comparison of various forms of the surface – by repeatedly bringing the discourse into its own presence – allows us to reveal the invariant structure of the process of production corresponding to a given state: the variations are the symptom of this structure.

We will now explain how that comparison can actually be made.

*From discursive surface to the structure of the process of production*

Let us consider the theoretical example given above. We used it simply to represent the *metaphoric effect*, but indicated that it was impossible to realize such an example. We now have to specify the determinations which were provisionally left aside in our abstract representation. We will demonstrate in turn:

- (1) the concrete impossibility of the extreme hypothesis as to the existence of *two discourses belonging to the same structure of production but having no terms in common*;
- (2) the consequences of this point for the notion of *context*, and the theoretical work that this notion implies;
- (3) the existence of a *dominance-effect* within the production of a given discursive sequence (see pp. 105–106 below); this has the effect of delineating *zones of pertinence* within the sequence in accordance with a given process of production  $\Delta_x$ .

---

<sup>1</sup> The term 'surface', which was originally introduced by Chomsky (surface structure/deep structure), should be seen here in its geometric context: a surface is a juxtaposition of discursive lines  $D_{x1} \dots D_{xn}$ . Rather than the linear sequence being related to the underlying operations of which it is a trace, each discursive line should be related to all the lines that are parallel to it in a given state of conditions of production. Deep structures do not, therefore, lie beneath the surface, but within the relationships between each surface (in Chomsky's sense) and its variations on the surface (in my 'geometric' sense of the term).

[It should be noted that the expression 'discursive surface' is widely used in French discourse analysis, and that it is not specific to Pécheux and the group around him. The opposition between deep structure and surface structure is seen here as analogous with the opposition between invariant and variation. Pécheux stresses this analogy on p. 177 below, but he also criticizes the opposition between invariant and variation on the grounds that it can be reinscribed within traditional dichotomies such as denotation-connotation, or norm-deviation and that it represents a challenge to the thesis that 'metaphor is primary and constitutive' (p. 178).]

(1) So far as the first point is concerned, it is clear that the suggested hypothesis is almost untenable with regard to *any* two discourses, given the existence in language of a small number of very frequent operator-words whose use is not semantically bound up with any given context. Moreover, and this is more basic for our purposes, it seems that the semantic-rhetorical rules governing the sliding of meaning in any process of production  $\Delta_x$  establish certain blocks at various points. This means that within that discourse certain metaphors exist only in a 'sleeping state'. In such cases, the substitution which gives the term employed its meaning does not function within that discourse ('*lever du soleil*' [sunrise], for example, represents a sleeping metaphor to the extent that the current state of the laws of substitution does not authorize any form that can be commuted with *lever* [rise]).<sup>1</sup> To that extent, we can therefore assume that there are, in the series of discursive sequences, certain abutments whose existence is revealed by the repetition of certain terms *around which metaphoric displacements occur*.

(2) This is tantamount to saying that we do not necessarily move from one discursive sequence to another by means of *substitution alone*, and that two sequences are usually linked by *a series of metaphoric effects*. But if we also accept that several metaphoric effects can function between a given discourse and the rest of the corpus, this also means that the context for substitution is not necessarily the discourse as a whole, and that in turn leads us to raise the question of the problem of the segmentation of contexts within the discursive sequence. In the paper cited earlier, Jakobson (1971: 243) writes:

Any sign is made up of constituent signs and/or occurs only in combination with other signs. This means that any linguistic unit at one and the same time serves as a context for simpler units and/or finds its own context in a more complex linguistic unit. Hence any actual grouping of linguistic units binds them into a superior unit.

He adds that "The addressee perceives that the given utterance (message)

---

<sup>1</sup> As we have seen (pp. 88-89 above), the 'implicit discourse' that the orator 'demands' of his listener is not *a priori present in the terms* of the orator's discourse; for the orator, this opens up the possibility of generating *stylistic figures* by playing on the listener's expectations.

is a combination of constituent parts (sentences, words, phonemes, etc.) selected from the repository of constituent parts (the code)". If we took this text literally, we might suppose that, as we move from *phoneme* to *discourse*, we are in the presence of linguistic signs whose dimensions increase but which are still bound by the same rule of combination. If that were the case, it would be impossible to define the *context of substitution*, as we would not know the dimensions of the sign to which it is restricted.<sup>1</sup>

We can overcome this difficulty if we recognize that the *sentence* has a very specific status, and that it is the *boundary* separating *linguistics* from *discourse theory*. Benveniste (1966: 128) has some important remarks to make on this point:

With the sentence, we cross a boundary and enter a new domain [...]. We can segment the sentence, but we cannot use it for purposes of integration [...]. Because the sentence does not constitute a class of distinctive units which can become members of higher units, it is fundamentally different from other linguistic entities. The basis for that difference is that the sentence contains signs, but is not itself a sign.

We will for our part use the term *utterance* [*énoncé*] to describe the elementary sentence insofar as it is the unique object upon which the mechanisms of discourse operate. It follows from the above that discourse is not a higher unit produced by relations of combination and substitution between utterances, as an utterance is already *of the order of discourse*. "The sentence belongs to discourse", writes Benveniste (p. 130), "and indeed that is how we can define it: the sentence is the unit of discourse."

In other words, a substitution always has as its context an utterance, regarded as a combination and substitution of lexemes. It cannot, however, be said that an utterance has a context in the same sense, as

---

<sup>1</sup> [It has to be admitted that this discussion of the context for substitution is confused. It is claimed that Jakobson adopts a linear position and that he argues that the same rules of combination apply at the level of the phoneme and at that of discourse, yet the quotation used on p. 74 above – taken from the same passage – implies the opposite. It seems that Jakobson is cited here solely in order to highlight the value of Benveniste's comments on the sentence (see below), even though they are not really exploited.]



utterances may be linked by *relations of formal dependency*, which is to say that *the syntagmatic contiguity between elements* – the fundamental principle of the linguistic analysis of the sign at its various levels – gives way to a *logical-rhetorical connection* which is no longer confined to connexity: two utterances may be functionally related across a discursive space which is neutral with regard to the relation between them.

We can see, then, that our problem is one of relating the internal properties of utterances (as combinations of signs) to their external properties (as functional elements in discourse) so as to determine in which cases two given utterances can be given the same *semantic interpretation*, in the sense in which logic understands that term. We can posit the hypothesis that, if there is to be a metaphoric effect between terms  $x$  and  $y$  belonging to utterances  $U_a$  and  $U_b$ , which are situated respectively in discourses  $D_{xi}$  and  $D_{xj}$  (representing the same process of production  $\Delta_x$ ), then  $U_a$  and  $U_b$  must be given the same semantic interpretation. We will write this as

$$I(U_a) = I(U_b)$$

This implies that

- (a) the lexemes of  $U_a$  and  $U_b$  provide a *common context* for the substitution of  $x$  for  $y$ , to which we will refer as a *condition of paradigmatic proximity* between  $U_a$  and  $U_b$ , and that
- (b) there is an identical *functional position* of  $a$  and  $b$  in relation to two other utterances  $c$  and  $d$ , which belong to the discourses  $D_{xi}$  and  $D_{xj}$  respectively, and which must themselves have a same semantic interpretation:

$$I(U_c) = I(U_d)$$

Let us illustrate the above with an example. Take the following utterances:

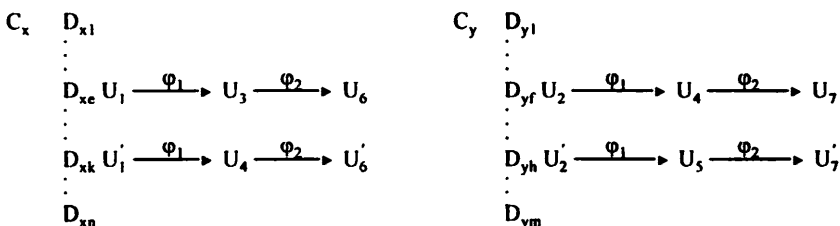
- $U_1 = U'_1 =$  *le shérif avançait prudemment en direction du salon*  
[the sheriff advanced cautiously towards the saloon]
- $U_2 = U'_2 =$  *l'orage grondait*  
[the storm was rumbling]
- $U_3 =$  *un coup de feu traversa la nuit*  
[a shot rang out in the night]
- $U_4 =$  *un éclair traversa la nuit*  
[a bolt of lightning shot through the night]

- $U_5 =$  *la foudre traversa la nuit*  
 [thunder rumbled through the night]  
 $U_6 = U'_6 =$  *la balle le frôla*  
 [the bullet grazed him]  
 $U_7 = U'_7 =$  *la grange était en flammes*  
 [the barn was in flames]

Let there be the following inter-utterance operators:

- $\xrightarrow{\varphi_1} =$  '*soudain*' [suddenly]  
 (temporal relation between state-utterance and event-utterance)  
 $\xrightarrow{\varphi_2} =$  ':'  
 (explanatory relation)

Let there be two corpora of discourses  $C_x$  and  $C_y$  such that



Finally, we will add the following *equivalent semantic interpretations*, which we will assume to have been obtained at an earlier stage in the analysis:

$$\begin{aligned}
 I(U_1) &= I(U'_1) \\
 I(U_6) &= I(U'_6) \\
 I(U_2) &= I(U'_2) \\
 I(U_7) &= I(U'_7)
 \end{aligned}$$

We see at once that  $U_3$ ,  $U_4$ , and  $U_5$  meet the 'condition of paradigmatic proximity', as the terms

*un coup de feu* [a shot]  
*un éclair* [a bolt of lightning]  
*la foudre* [thunder]

are substitutable in the context

*... traversa la nuit.*

$U_3$  and  $U_4$  also receive the same semantic interpretation in  $C_x$  because  $I(U_1) = I(U'_1)$ . This produces metaphoric effect  $M_1$  in  $C_x$ :

$M_1 =$  *un coup de feu*  
*un éclair*

Similarly,  $U_4$  and  $U_5$  receive the same semantic interpretation in  $C_y$  because  $I(U_2) = I(U'_2)$ . This produces metaphoric effect  $M_2$  in  $C_y$ :

$M_2 =$  *un éclair*  
*la foudre*

It is to be noted that the relations of semantic interpretation are not transitive, as

$I(U_3) = I(U_4)$

and

$I(U_4) = I(U_5)$

do not imply that

$I(U_3) = I(U_5)$

because, whilst the 'condition of paradigmatic proximity' between  $U_3$  and  $U_5$  has been met, the condition of 'identity of functional positions' has not. In fact,  $I(U_3) \neq I(U_5)$ , for  $I(U_1) \neq I(U'_2)$  and  $I(U_6) \neq I(U'_7)$ .

(3) It remains, finally, to explain what is meant by a 'dominance-effect' within the production of a given discursive sequence. We have until now been arguing in the following terms: 'Given state  $\Gamma_x$ , how can we analyze the process of production  $\Delta_x$  by analyzing a set of discourses representing that state?' This presupposes that every element in the discursive surface necessarily (and even with equal necessity) refers to  $\Delta_x$ ,

and that all discourses corresponding to the same state of production are therefore strictly parallel, or in other words absolutely isomorphic (due allowance being made for the metaphoric effects which differentiate them).

This is, it will be noted, a highly improbable eventuality: the parallelism is broken at certain points by 'individual' distortions of discourse. Discourse therefore appears to 'escape' the process of production thanks to the 'infinite creativity' and the 'unlimited variety' characteristic of speech.<sup>1</sup>

I think it is possible to account for this phenomenon without abandoning our earlier theoretical presuppositions, which are based upon the notion that the discursive process is *determined* by its conditions of production, and upon a rejection of the ideological notion of 'infinite creativity'. At this point, we will introduce the concept of *dominance* by stipulating that any situation in which discourse is produced can be characterized in terms of the *dominant process of production* ( $\Delta_x$ ) induced by this situation, but that the *concrete discursive* sequences which manifest  $\Delta_x$  necessarily result from the interaction between the dominant process and *secondary processes*. The latter are so entangled as to have the appearance of being aleatory, because at present we know nothing about the mechanisms of their interaction with the dominant process.

We are now in a position to formulate our present objective more accurately. Given a dominant state of conditions of production of discourse, there is a corresponding dominant process of production which we can reveal by comparing different empirical discursive surfaces resulting from that dominant state. The intersection-points defined by metaphoric effects allow us to extract the *semantic domains* determined by the dominant process and the *relations of logical-rhetorical dependency* it induces between those domains. The remainder of the discursive material we encounter at the empirical level falls outside the zone of pertinence of the dominant process.

---

<sup>1</sup> These terms are borrowed from Benveniste, who obviously relates discourse to speech in this way.

[Pécheux criticizes Benveniste's subjective position on 'infinite creativity'. He returns to Benveniste's conception in Chapter 6, where he refers to the "ideological duality of system and creativity (a system of signs as opposed to individual creativity)", and concludes that "'discourse' is no more than a new avatar of 'speech'" (p. 185 below). But there the discussion includes a consideration of processes of enunciation, which will become more important in Pécheux's subsequent work.]

This presupposes, let me repeat, that, in its textual materiality, *a discourse does not display an organic unity at a single level*, and that no such unity can be revealed on the basis of the discourse itself. This further presupposes that any particular discursive form necessarily refers back to the series of its possible forms, and that the references back from the surface of each discourse to the possible surfaces which are (in part) juxtaposed against it during the process of analysis, are themselves *pertinent symptoms* of the dominant process of production governing the discourse that is being analyzed.

### 5.3. Rules for a coded recording of the discursive surface

We will assume that the conditions defined above have been met, that is, that the series of discourses  $D$  submitted for recording and analysis corresponds to a single dominant state of conditions of production inducing a process of production  $\Delta_x$ .  $D_{x1}$ ,  $D_{x2}$ , ...,  $D_{xn}$  designate the  $n$  discourses that have been empirically recorded under the above conditions. They are considered to be representative of those conditions, and we will demonstrate below that there are formal means to determine, whether or not the 'corpus' that is so constituted is sufficiently *systematic* to be representative of the structure of the process of production  $\Delta_x$ . Our problem consists of the analysis of any given surface  $D_{xi}$  in terms of *minimal elements* linked by laws specific to the process of production  $\Delta_x$ .

The views expounded in the previous section presuppose that we define two levels of analysis:

- (1) analysis of *utterances* in  $D_x$  which are linked by functional relations characteristic of the process of production  $\Delta_x$  under consideration;
- (2) analysis of *lexemes* in utterances, which are linked by the laws of combination and commutation characteristic of the process of production  $\Delta_x$  under consideration.

This double analysis can only function on the basis of a double hypothesis as to the *linguistic object in general*, whatever the process of production  $\Delta_x$  under consideration:

- (1) Any linguistic sequence consists of a structured set of related *utterances* which are discernable on the basis of general linguistic laws.
- (2) Any linguistic utterance is composed of *lexemes* bound together by universally necessary morpho-syntactic relations deriving from a grammatical theory of the utterance.

It is, then, as though we had first to posit *properties* which remain *invariant* throughout a variety of processes of production, and as though those properties served as a referential context for the variations we wish to bring out. We will therefore begin by discussing the effects of this invariant on *the recording of the discursive surface*, which can be regarded as a preliminary stage essential to discourse analysis.

I do not claim to be providing anything more than an *outline* of the recording process, and I am well aware that I am leaving the linguist to take a large number of decisions which we cannot take for him. My only ambition here is to outline and justify the *general direction* we have to take, as it is a precondition for the next phase of the analysis. From the viewpoint of a study of the discursive process, it is that next phase, of discourse analysis proper, which is the real issue.<sup>1</sup>

### Relations between utterances

The following list summarizes our hypotheses as to a discourse's system of dependencies in the form of postulates:

- (1) A *discourse* is a linguistic sequence between two *semantic blanks* and corresponding to given discursive conditions of production.
- (2) A *sentence* is a part of a discourse between two *consecutive stop-markers* — or, in the case of an initial sentence, by a *blank* and a *stop-marker*.
- (3) A *proposition* is part of a sentence comprising only one verb taking a finite form.
- (4) A *reduced proposition* is a proposition that cannot be broken down into two or more *propositions* by means of a transformation 'liberating' latent propositions by completing them, such as in

*He asked him to answer immediately upon receipt of the letter*  
 → *He asked him [THAT] he would answer [AS SOON AS]*

---

<sup>1</sup> [Because of the provisional nature of the following description of a method for the automatic analysis of the recorded material, parts that are not essential for a proper understanding of the ADA69 project have been left out here, in particular a large part of the section on relations between utterances (pp. 40-66 in Pêcheux 1969a), of which only Pêcheux's summary is presented, and detailed discussions of the recording of the elementary utterance (Pêcheux 1969a: 77-85).]

*he received the letter*

(5) A *elementary utterance* is a proposition in which no latent propositions are left that can be 'liberated' by a transformation of an adjunct (an adjective, an attributive adjunct, or a relative clause) into a separate proposition, such as in

*The little dog sleeps by the fire*

→ *The dog sleeps by the fire [AND] the dog is little*

(6) Any elementary utterance can be recorded in the form of *an ordered set of fixed dimensions*; its elements are linguistic signs belonging to defined morpho-syntactic categories (see pp. 112-115 below).

(7) For any elementary utterance  $U_j$  there is at least one utterance  $U_i$  such that

$$U_i \rightarrow U_j$$

is a binary relation in which  $U_j$  is dominated directly by  $U_i$ . Conversely, any utterance  $U_j$  can dominate  $m$  utterances where  $m \geq 0$ .

(8) The beginning of the discourse is represented by the binary relation

$$U_0 \rightarrow U_1$$

in which  $U_0$  represents a semantic blank and  $U_1$  the first utterance of the discourse.

(9) The dependency relation between elementary utterances  $U_i$  and  $U_j$  may mark either

- *adjunction* of  $U_j$  to  $U_i$ ; or
- *coordination* as indicated by a *stop-marker* (there may or may not be a *syntagma qualifying* the marker, and there may or may not be an *anaphoric effect*); or
- *subordination*, as indicated by or a *subordinating term* or *syntagma*.

(10) The structure of a discourse can be represented by a list of binary relations.

### The recording of the elementary utterance

We have been dealing with the recording of the *functional dependencies* specific to a given discursive structure by showing that they can be represented by a set of binary relations taking the form

$$U_i \rightarrow U_j$$

where  $\rightarrow$  is a dependency operator, and where  $U_i$  and  $U_j$  are elementary utterances with a definite content. We will now specify the modalities of the recording of their content.

We depart from the following structure of each elementary utterance:

$$NP_1 + VP [+ P] [+ NP_2]$$

in which  $NP_1$  and  $NP_2$  are noun phrases, VP is a verbal phrase and P is a preposition.

The 'active' protagonist of the utterance is represented in the schema by  $NP_1$ : *subject noun phrase*. We will argue that  $NP_1$  is an indispensable component of any utterance – the case of 'meteorological' utterances ('it is raining' or 'it is snowing' are the usual examples given) is too marginal to represent a challenge to this rule.

The 'passive' protagonist of the utterance is represented by  $NP_2$ : *object noun phrase*, which can take several forms including that of absence. We will also argue that an *adjective* which is an *attribute of the subject* can also take the position of  $NP_2$ ; this implies that any form that can be inscribed in the position of  $NP_1$  can also be inscribed in the position of  $NP_2$ . The converse is not true.

VP, finally, represents the 'process of the utterance', and we will argue that it can always be re-established if it is not explicitly present. Apposition, in particular, can be systematically transformed into predication, and can be recorded as such:

*Waterloo, dismal plain, ...*  
 $\rightarrow$  *Waterloo is a plain*  
*The plain is dismal*

It goes without saying that we are not claiming that these various components are linguistic 'universals'; our aim here is simply *to demonstrate*



*the possibility of recording.*

### *The noun phrase*

The subject noun phrase necessarily includes a noun or a term representing a noun. In the latter case, the term standing in for the noun is related to it by an anaphoric phenomenon, which may be *external* or *internal*, e.g.:

*I declare the session open*

The active protagonist is the chairman of the session, *I* = the chairman X, or:

*The train arrived at the station. It (= the train) was on time*

We can thus see that pronouns and other stand-ins refer either to the *enunciation* or to an *earlier utterance*, and that the enunciation of that earlier utterance sanctions the introduction of what Benveniste (1966: 254) calls 'empty signs' into a subsequent utterance:<sup>1</sup>

['Empty' signs] do not refer to 'reality', are always available, and become 'full' when a speaker assumes them in every instance of his discourse. Having no material reference, they cannot be used wrongly; and as they assert nothing, they are not subject to the condition of truth and cannot be denied. Their role is to provide the instrument for a conversion, which one might term the conversion of language into discourse.

The noun itself is usually accompanied by a *mark of determination*. In the case of proper nouns and most anaphora, this mark is absent. We will term these marks DET<sub>1</sub> and DET<sub>2</sub>, depending on whether their association with

---

<sup>1</sup> [The terminology used here by Pêcheux may give rise to some confusion in that he contracts internal and external anaphoric phenomena (the phenomenon of substitution within the linear dimension of the text and a reference back to the protagonists of the enunciation respectively). The quotation from Benveniste obviously serves to clarify the status of elements referring to the enunciation (an external anaphora, to use Pêcheux's terminology).]

the noun forms a subject phrase or an object phrase:

$$\begin{aligned} NP_1 &\rightarrow DET_1 + N_1 \\ NP_2 &\rightarrow DET_2 + N_2 \end{aligned}$$

### *The verb phrase*

The verb phrase may be regarded as a *verbal lexeme* plus a *series of combined qualifications*. Take, for example, the sentence:

*Might he not have been poisoned deliberately?*

We have:

(to) poison: verbal lexeme  
deliberately: *adverbial* modulation of lexeme  
might he: *question* indicated by inversion  
have been -ed: mark of *past passive*  
might: mark of *supposition*  
not: mark of *negation*

The *adverb* is a *relatively free* element as compared to the rest of the phrase. On the one hand the position of the adverb may remain empty. On the other it may be displaced across the surface to a point at which it can, in certain cases, govern the *whole of the utterance* and link it to an *earlier utterance*. The problem is then to determine whether a given adverb functions as a *modulation of the verb* (as a kind of adjective applying to the verb) or as a *qualification of a dependency operator*, in which the adverb marks the *order of dependency* between two narrated events.

Now our final task is the recording of the other qualifications of the utterance, such as status (affirmative, negative, interrogative), tense, voice, mode, modality and emphasis. Each of these qualifications can be represented by a value. We will use the term '*form of utterance*' to refer to the vector constituted by the ordered set of values applying to each mark in the utterance under consideration. The form of utterance (F) can exercise its determination over every element in the utterance, and not merely over its verbal phrase. If we agree to term a canonically ordered set of such elements a *lexis*, we can say that *an utterance results from the application*

of a form of utterance to a given lexis:

$$U_n = F (DET_1, N_1, V, ADV, P, DET_2, N_2)$$

An elementary utterance can, thus, be recorded as an ordered set of *eight terms*, each of which corresponds to a determinate *morpho-syntactic category*.

A great deal of linguistic work remains to be done before the recording of utterances can be made automatic. Many of the points made here will probably be challenged as a result. The main point is to specify the essential linguistic pre-requisites for any analysis. In other words, what is proposed here is a *method* and not a *theory of language*. It goes without saying that the method is based upon theoretical presuppositions which the linguist must make explicit and criticize. The central issue is whether or not it is legitimate to represent a sequence by means of a set of binary relations linking elementary utterances represented as an ordered set of eight morpho-syntactic categories.

#### 5.4. The automatic analysis of the recorded material

Let  $C_x$  be a corpus of discourses  $D_{x1} \dots D_{xn}$  corresponding to stable conditions  $\Gamma_x$  defined by external criteria of a sociological nature (situation and position of sender, roles involved, etc.). In accordance with our earlier hypotheses (pp. 102-103), we posit the existence of two mechanisms:

- (a) a process of selection and combination applied to the language producing the set of utterances under examination;
- (b) the application of the set of utterances to itself.

With respect to the mechanism of selection and combination, it should be noted that certain constraints are imposed by the language itself – the need, for example, to select a complement to the object if it is to be combined with a transitive verb. Such constraints are external to the process of selection and combination discussed here, which represents constraints which are not imposed by the language. For example, the process of selection and combination carried out on the words ‘property’ and ‘theft’ is not the same in the following two sequences:

*Theft is an attack on individual property*  
*Property is theft*

We have taken the view that syntactic constraints are *stable within the language*<sup>1</sup> (effects of order, emphasis, etc., not being regarded as syntactic), and that both the specificity of combinatory constraints on the components within the utterance and the constraints on the relations between utterances represent the distance between the language and discourse. They represent, that is, the effect of the conditions of production of discourse on the process of production of discourse:

$$\Gamma_x^n \circ L \rightarrow \Delta_x^n$$

In order to reveal  $\Delta_x$  we will attempt to define both the points at which it is 'anchored' in the corpus — these points will be termed 'semantic domains' — and the relations of dependency between those domains. As said earlier (see p. 103 above), utterances  $U_i$  and  $U_j$  can be given the same semantic interpretation if

- (a)  $U_i$  and  $U_j$  are paradigmatically close to one another; and
- (b)  $U_i$  and  $U_j$  are linked by identical functional dependencies to utterances  $U_k$  and  $U_m$ , which are themselves paradigmatically close to one another.

This presupposes that we can define:

- a programme for the paradigmatic comparison of utterances;
- a programme for the formation of semantic domains by relating utterances on the basis of their relations of dependency.

## **Paradigmatic comparison of utterances**

### *Comparison of the contents of morpho-syntactic categories*

Take utterances  $U_i$  and  $U_j$ , both consisting of a sequence of terms ordered in accordance with morpho-syntactic classes. For example:

---

<sup>1</sup> The notion that the language forms an 'invariant background' (which means primarily that syntax is a source of universal constraints) is introduced on p. 77 above. See Chapter 6, pp. 138-141, for a further discussion of discourse as a system of selection and combination of lexical units which does not relate to the system of the language.

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
U <sub>i</sub> =	a	b	c	d	e	f	g	h
U <sub>j</sub> =	a	b	c	d	j	f	k	m

We now associate a number with the (U<sub>i</sub>, U<sub>j</sub>) pair in each morpho-syntactic category, adopting the convention that two identical terms within the same morpho-syntactic class are represented by the number 1, and different terms by the number 0. Here, for example, we obtain

$$(U_i, U_j) = 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0$$

### *Value of paradigmatic proximity*

In the above example, one could say that, on a scale from 0 to 8, the proximity (U<sub>i</sub>, U<sub>j</sub>) has the value 5. It is important, however, to consider the possibility of giving different values ('weights') to different morpho-syntactic categories. For example, it is reasonable to assign a higher value to lexical identity in the categories N<sub>1</sub> and V than in the categories DET<sub>1</sub> and DET<sub>2</sub>. We could ask a linguist to fix these values for once and for all, but it is also possible to keep them *variable*, depending on the aims of the analysis. Provisionally, I would like to propose the following weights:<sup>1</sup>

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
weight	3	2	5	5	3	3	2	5

### **The constitution of semantic domains and the analysis of their dependency**

Take two binary relations

$$\begin{aligned} U_i &\rightarrow U_m \\ U_j &\rightarrow U_n \end{aligned}$$

with an identical connector K<sub>1</sub>. If the value of the paradigmatic proximity

---

<sup>1</sup> See Chapter 7, p. 198, for another example of weighting.

between  $U_i$  and  $U_j$  and  $U_m$  and  $U_n$  respectively is greater than a certain threshold value, we can say that

$$\begin{array}{ccc} U_i & & U_m \\ & K_1 & \\ U_j & & U_n \end{array}$$

constitutes a zone of similitude, which may be amenable to expansion by the concatenation

$$\begin{array}{ccccc} U_i & & U_m & & U_q \\ & K_1 & & K_2 & \\ U_j & & U_n & & U_t \end{array}$$

if

$$\begin{array}{ccc} U_m & & U_q \\ & K_2 & \\ U_n & & U_t \end{array}$$

meets the same requirements. A concatenation of zones of similitude will be termed a *chain* of similitude. If two chains contain sequences that have enough paradigmatic proximity, if, that is, a binary relation in one chain can form a zone of similitude with a binary relation of another chain, we form a chain by vertical transitivity. Such a vertical chain is called a *semantic domain*.

### **Analysis of dependency between semantic domains**

Utterance  $U_n$  will be said to depend upon utterance  $U_k$  if there is a concatenation of direct dependencies between  $U_k$  and  $U_n$ :

$$U_k \rightarrow U_i \rightarrow \dots \rightarrow U_n$$

An utterance to the left of a sequence will be termed the *source* of that sequence. If two sequences  $S_x$  and  $S_y$  have *different sources* and if the source of  $S_y$  depends upon the source of  $S_x$ , we can say that  $S_y$  depends on  $S_x$  (or that  $S_x$  governs  $S_y$ ).

Two domains will be said to have completely the *same source* if, for any sequence in one domain, there is a sequence in the other domain which has the same source, and vice versa. A domain will be said to *include* another domain if the set of the sources of the sequences in the latter is part of the set of the sources of the first. Two domains will be said to *intersect* if the intersection between the sets of the sources of their sequences is not empty, even though they do not include each other. A domain will be said to *depend* on another if the intersection between the sets of the sources of their sequences is empty and if certain sequences in the domain depend upon certain sequences in the other, *whilst the converse is not true*.

In the light of the above, we can say that the process of production  $\Delta_x$  of a discourse is represented by the network of relations affecting the semantic domains that have been revealed.

## 5.5. Potential applications of automatic discourse analysis

The project I have described is in many respects incomplete.

On the one hand, the sociologist has been left with the task of providing a detailed definition of the specific features which characterize conditions of discursive production in terms of the *situation* and *position* of the protagonists of discourse in a given social structure. Incidentally, I have also provisionally ignored the question of *discourses which are not monologues* on the grounds that the solution to that problem depends upon the resolution of the particular case to which the present analysis is restricted. Further work must also be done on linguistics before rules can be established for the *recording of the discursive surface*. Finally, it has been left to the mathematician to define *a minimal sequence of algorithms* capable of carrying out the analysis.

On the other hand — and this relates directly to my own future activities — I am aware of the existence of a certain number of difficulties that have still to be overcome. The present method, for example, can account for *term for term equivalences* between two discursive surfaces to the extent that it compares parallel concatenations of utterances, but the possibility of *overall semantic equivalences* corresponding to different dependency structures remains an unresolved problem. Let me simply say that it seems possible to envisage the 're-injection' of the results of the present analysis into the initial discursive surfaces, and a further com-

parison of the concatenations that can be generated on the basis of the new state of the surface. This suggests the idea of a recurrent process consisting of *cycles of analysis*; the results obtained at the end of cycle  $n$  would thus provide an input for cycle  $n + 1$ .

If we look at the results obtained with the current analysis programme, we find that the dominant process of production  $\Delta_x$  is represented by a set of 'domains' which are linked to one another in various ways. I think that it is therefore possible to represent each domain by one or more *propositions* of the  $g(x)$  or  $m(x,y)$  type, as appropriate, and to define by deduction the transformations that affect the predicates of the propositional variables brought into play in the process of production. We can thus arrive at the logical rules defining semantic coherence and the transformation of that coherence. In other words, we arrive at the rules governing the semantic effect produced by  $\Delta_x$ . If this is the case, we have, then, an instrument that will allow us to identify the type of process in play (the structure of narrative, as distinct from, for example, the structure of a logical argument) and to provide *theoretical facts* that can be integrated into a theory of discourse which is a general theory of the production of meaning-effects.<sup>1</sup>

The ultimate purpose of this undertaking is, then, to provide the preconditions for a practice of *reading*, defined as the systematic detection of symptoms representing meaning-effects within the discursive surface. Before I briefly outline the uses to which such a practice could be put, it is important to clarify one last point of vital importance: I refer to the *principle* of this reading, which might be described as the 'principle of double difference'. I have shown that a controlled comparison of discursive surfaces deriving from the same conditions of production  $\Gamma_x$  allows us to reveal the *internal differences* through which the *invariant in discourse*

---

<sup>1</sup> I again stress that a discourse theory can no more replace a theory of ideology than it can replace a theory of the unconscious. It can, however, intervene in the field of such theories.

[At this stage, the theory of ideology is no more than a project elaborated by one 'Thomas Herbert', who is haughtily ignored by the author of ADA69. Herbert is also the author of a 'theory of the unconscious', but the reference to it here might suggest that it too remains to be elaborated. This veiled allusion in a discussion of potential applications seems to be indicative both of Pêcheux's tactical prudence and of his continued belief in the advent of an eldorado which will fulfil what we have termed the fantasy of articulation (see Chapter 3, p. 44).]



$x$  is manifested<sup>1</sup>; I have defined this invariant as process of production  $\Delta_x$ . We thus arrive at a representation of the semantic effects *present* in  $\Delta_x$ . Our earlier discussion of the 'implicit discourses' to which a given discursive surface refers (see pp. 88-89 above) suggests, however, that the *external differences* between  $\Delta_x$  and other processes  $\Delta_y$ ,  $\Delta_z$ , ... that make up its specific exterior must also be taken into consideration. I believe, in other words, that process  $\Delta_x$  is characterized, not only by the semantic effects realized within it or by *what is said* in discourse  $x$ , but also by *the absence of a certain number of effects* which are present 'elsewhere', 'elsewhere' being precisely what we will call the *specific exterior* of  $\Delta_x$ . This implies that we can only define the absence of a meaning-effect as *the specific absence of something that is present elsewhere*: the implicit 'unsaid' which characterizes a process of production  $\Delta_x$  is therefore represented by the distortion induced in  $\Delta_x$  when it is compared with  $\Delta_y$ ,  $\Delta_z$ , ..., which thus become *the real cause of the absences characteristic of  $\Delta_x$* . The 'errors' and 'omissions' specific to the discourse of a science in a given state are, for example, visible only if that discourse is compared with the discourse that corrects it. Similarly, a stylistic device exists only in relation to a process which is implicitly assumed to involve the addressee, and upon which the addresser relies.

The modes of the insertion of the practice of discourse analysis into various sectors of research pose a great number of problems which will not be dealt with here. Once again, I will simply indicate certain lines of research by way of example.

### *The field of sociological investigation*

Insofar as sociology sets itself the task of investigating the *relationship between the power relations and the meaning relations specific to a given social structure*, it regards the discourse of the social subject as representing the relationship between his socio-economic *situation* and his ideological *position* within that structure. Anything the subject says must always be related to the conditions under which he says it: the content of an interview a factory manager grants a sociologist is less pertinent than the

---

<sup>1</sup> [The term 'discourse' can refer to either an oral or a written sequence of variable dimensions (as, e.g.  $D_x$  in 5.2. above, p. 87), or to a 'discursive process' (as in 'the discourse of science'). Here it is used in the second sense (cf. p. 141 below).]

confrontation of his discourse with *what he says and does in other contexts* – in other words, in relation to other discursive roles whose effects can be found *elsewhere* – and to the subject's practice which can be described sociologically as placed within a field of practices.

In other words, the introduction of the 'principle of double difference' must allow us to define both  $\Delta_x$  – the dominant discursive process – and the specific differences it displays with regard to other processes corresponding to other discursive conditions of production.

This may allow us to shed light on problems such as those of implicit cultural values, implicit and explicit forms of consensus and differentiation, and the implication of the expected response in any question that is asked.

### *The field of the history of science*

The identification of the 'epistemological break' between a science and the terrain it leaves in order to constitute itself has been demonstrated to be one of the crucial problems to be solved by the history of science. The analysis of the conditions under which a new scientific discourse is established by using means borrowed from existing sciences or from 'non-scientific' representations can be described as the interrelating of several processes of production whose interaction generates, in certain conditions, a new process which disrupts the coherent rules governing the earlier discourse. If it is true to say that reading a scientific text means relating it to the discourse from which it has separated, we can see that the practice of analysis requires us to reveal that element in a text  $D_y$  which produces a discrepancy – a difference ascribable to the nature of its predicates and the transformations they undergo – between it and the dominant process of production  $\Delta_x$ . It is precisely  $D_y$  which prevents  $\Delta_x$  from repeating itself indefinitely.

The study of the processes from which a science borrows the elements it uses as metaphors in order to understand itself and to make itself understood,<sup>1</sup> that of the 'context' of a scientific work – the constellation of discursive processes with which it argues and struggles – and

---

<sup>1</sup> As Canguilhem notes, Aristotle and Harvey used different metaphors. Aristotle believed that blood irrigates the body *just as* water irrigates the land. Harvey, on the other hand, saw the circulation of blood *as* a hydraulic system complete with pumps and sluices.

that of the process of the 'diffusion' of knowledges within a system of pre-scientific representations, all raise problems which the type of analysis outlined here may help us to resolve. I repeat that an immense body of work remains to be undertaken before these various possibilities can be realized in concrete terms. Given the *volume* of material to be processed, the implementation of these analyses is in fact dependent upon the *automation of the recording of the discursive surface*. In my view, there is no alternative, and any preliminary or arbitrary *reduction* of surface  $D_{xn}$  by means of techniques of the 'coded résumé' type is to be avoided because it *presupposes a knowledge of the very result we are trying to obtain*, namely a representation of the  $\Delta_x$  corresponding to the class of discourse from which  $D_{xn}$  has been extracted.



## Overview and prospects

Michel Pêcheux

ADA69 has, over a period of years, given rise to a number of publications which deal with it both at the level of its method (Haroche and Pêcheux 1972a) and at the level of its experimental applications (cf. Gayot and Pêcheux 1971, Gayot 1973, Pêcheux and Wesseliuss 1973). The remarks, interpretations, criticisms and even distortions it has provoked at both these levels (cf. Dubois 1970, Provost-Chauveau 1970, Cipolli 1972, Dupraz 1972, Trognon 1972, Borillo and Virbel 1973, Fisher and Veron 1973, Robin 1973: 103-107, Del Vigna and Dupraz 1974) make it necessary to provide a general overview. I will attempt to dispel certain ambiguities, to rectify certain errors and to admit the existence of certain unresolved difficulties, and at the same time to outline the basis for a new formulation of the question in the light of more recent work, much of which remains unpublished, on the relationship between linguistics and discourse theory.

In order to remove any ambiguities which might lead to the necessary critical work specific to a theoretical field being confused with an attempt to retreat from or to abandon that field, the general epistemological framework for this undertaking will be described in 6.1. That framework is provided by the articulation of three regions of scientific knowledge:

- (1) historical materialism as a theory of social formations and of their transformations, and as a theory of ideologies;
- (2) linguistics as a theory of both syntactic mechanisms and processes of enunciation;
- (3) discourse theory as a theory of the historical determination of semantic processes.

It should be added that these three regions are in a sense traversed and articulated by a reference, which has to be made explicit, to a theory of subjectivity, and that theory is of a psychoanalytic nature.

As a result, one of our central questions will be reformulated as the question of reading, as the question of the reader-effect which constitutes subjectivity and which is characterized by the fact that the effect's conditions of existence must be concealed from the subject if it is to be realized. We will discuss what in this 'forgetting' pertains typically to the domain of linguistics in relation to what pertains to non- or pre-linguistic domains.

The second part of this chapter (6.2.) will be devoted to a detailed discussion of the various aspects of ADA69 that have given rise to criticism – a discussion which cannot be included in the general theoretical framework of 6.1. Whenever possible, we will indicate how outdated aspects can be reformulated (locally, and within the framework of the original problematic), and we will attempt, insofar as it is possible to do so, to establish the conditions for a radical transformation of the very terms of the problematic. We will attempt to establish, that is, the conditions for a revolution which everyone agrees is necessary but whose form cannot be predicted at the moment. If it is true that one only destroys what one replaces (and from this point of view, ADA69 is designed to destroy 'content analysis'), we have a theoretical responsibility for preparing the ground on which the displacement and replacement invoked by the metaphor of 'revolution' can take place. More specifically, our project presupposes that we overcome the discrepancy between practical procedures for processing texts and the level reached in discussions as to relations between the three regions mentioned above, and, above all, that we overcome the discrepancy between *discourse analysis* and *discourse theory*.

## 6.1. Social formation, the language and discourse

### **Social formation, ideology and discourse**

The organizational focus of this part (6.1.) is provided by the relationship between the three regions mentioned. It must immediately be pointed out that, given present working conditions in the university, everything concurs to make it very difficult to articulate these three regions theoretically. Quite apart from the fact that some may find this attempt to articulate them of dubious theoretical taste, it is true to say that, even with the best theoretical and political will in the world, it is difficult to overcome

the organizational and epistemological obstacles bound up with the balkanization of knowledge and, above all, with the university's repression and distortion of historical materialism. Experience tells us that it is very difficult to avoid spontaneous translations which turn historical materialism into a 'sociology' and which restrict discourse theory to 'the social aspect of language' and so on. Even Marxist researchers who are quite capable of making a lucid critique of their original disciplines often remain blind to certain academic and idealist aspects of neighbouring disciplines, so much so that they believe that they can find in them 'instruments' which are of immediate use in their own practice, and even in their critical practice.

The manner in which we propose to formulate this articulation is obviously not immune to that thread, as it is coextensive with the current conditions of academic practice. Adopting the most recent state of this formulation (Haroche, Henry and Pêcheux 1971), we will begin by positing the view that the region of historical materialism which concerns us here is that of the ideological superstructure in its relations with the mode of production dominating a given social formation. Recent Marxist work (especially Althusser 1971a) shows that it is not enough to regard the ideological superstructure as the expression of an 'economic base', as though ideology consisted of a 'sphere of ideas' hovering above a world of things and economic facts. In other words, the region of ideology must be characterized by a specific materiality which is articulated with economic materiality. More specifically, the workings of the ideological instance must be regarded as being 'determined in the last instance' by the economic instance insofar as the ideological instance can be seen to be one of the non-economic conditions for the reproduction of the economic base or, more accurately, for the reproduction of the relations of production inherent in that economic base.<sup>1</sup> The particular modality of the workings of the ideological instance in relation to the reproduction of relations of production takes the form of what has been termed the *interpellation* or subjection of the subject as ideological subject in such a way that everyone is, without realizing it and whilst under the illusion that one is acting of one's own free will, *obliged to take their place* in one or another of the

---

<sup>1</sup> Relations of production are not, as functionalist sociology claims, caught up in an eternal cycle of repetition. In reality, and insofar as relations of production correspond to class relations, we have to speak of their *reproduction/transformation*. This is not the place to develop this point, which is of crucial importance to historical materialism [cf. Pêcheux 1982a: 97ff.].

antagonistic social classes which exist within the mode of production (or in a category, stratum or class fraction connected with one of those classes).<sup>1</sup> The continued reproduction of class relations (and this, as we have seen, is both an economic and a non-economic process) is guaranteed by the existence of those complex realities which Althusser designates 'ideological state apparatuses', and which are characterized by the fact that they involve practices associated with places or relations between places which relate to class relations but which do not reproduce them exactly. At any given historical moment, class relations (the class struggle) are characterized by a conflict between political and ideological positions within those apparatuses. Such positions are not

created by individuals, but are organized into *formations* displaying relations of antagonism, alliance or domination. We will use the term *ideological formation* to characterize an element capable of intervening as a force confronting other forces within the ideological conjuncture that characterizes a social formation at a given moment; every ideological formation thus represents a complex set of attitudes and representations<sup>2</sup> which are neither 'individual' nor 'universal' but relate more or less directly to mutually antagonistic '*class positions*'. (Haroche, Henry and Pêcheux 1971: 102)

We now have to raise the question of the relationship between ideology and discourse. In the light of the above, it is clear that it is impossible to

---

<sup>1</sup> Being the most fully developed form of ideology, bourgeois ideology tells us a great deal about earlier historical forms, as well as about the workings of the ideological instance in general. Bourgeois forms of interpellation must not, however, be projected on to those earlier forms. It is not, for example, self-evident that interpellation always refers the subject to its determination. Insofar as it is 'a representation of an imaginary relationship', the autonomy of the subject is in fact closely bound up with the appearance and extension of bourgeois juridico-political ideology. In social formations dominated by other modes of production, the subject may represent its determination as being imposed by certain constraints or by the will of others; the relationship that is so represented is no less imaginary.

<sup>2</sup> We make no secret of the fact that by using terms such as 'attitudes' and 'representations', which are borrowed from the vocabulary of sociology, we are introducing an element of ambiguity. *Practices*, in the Marxist sense of that term, are neither 'forms of social behaviour' nor 'social representations'.



identify ideology with discourse (to do so would imply the idealist conception that ideology is a sphere of ideas and discourses), and that the discursive must be regarded as one of the material elements in what we have termed ideological materiality. In other words, it seems to us that the discursive is a *species* belonging to the *genus* of the 'ideological'; this implies that the ideological formations we have just mentioned

necessarily<sup>1</sup> have as one of their components one or more interconnected *discursive formations* which determine what *can or should be said* (or articulated in the form of a harangue, a sermon, a pamphlet, an exposé or a programme) on the basis of a given position within a conjuncture (Haroche, Henry and Pêcheux 1971: 102),

or in other words within a relationship between positions inside an ideological apparatus and inscribed in class relations. We can therefore state that any discursive formation derives from specific *conditions of production*<sup>2</sup> which can be identified on the basis of the above considerations.

"Ideology interpellates individuals as subjects" (Althusser 1971a: 170); this constitutive law of ideology is never realized 'in general'. It is always realized through a complex set of determinate *ideological formations* which play a necessarily uneven role in the reproduction and transformation of relations of production at any given historical stage of the

---

<sup>1</sup> This necessity relates to the specificity of language, which is an inherent characteristic of human beings in that they are ideological animals.

<sup>2</sup> The term *production* may result in a certain ambiguity here. In order to dispel any possible ambiguity, a distinction will be made between the economic meaning of the term, its epistemological meaning (production of knowledge), its psycholinguistic meaning (production of a message) and, finally, the meaning it acquires in the expression 'production of an effect'. The term is to be understood primarily in the latter sense. We will, however, see later that the mechanisms of the realization of the discourse produced by the subject are also involved. The use of the term is also, in our view, a polemical device to combat the frequent use of 'circulation' or even 'creation' to describe processes of signification.

It should be added that verbal (phonic or graphic) materiality is one of the prerequisites for economic production, in that it is both an infrastructural precondition for commerce (and, more generally, contracts) and a precondition for the social mobilization of the productive forces (transmission of 'instructions' for using instruments of labour) and for the 'education' of the labour force. The meaning of the expression 'conditions of production' will be clarified below (pp. 132, 141, 143-144).

class struggle because they have both 'regional' characteristics (of Law, Ethics, Knowledge, God, etc.) and class characteristics. Discursive formations therefore intervene as component elements in ideological formations on the basis of both their regional and their class characteristics. To take an example: in the feudal mode of production the *dominant ideology* takes the form of the *religious ideological formation*. It realizes the 'interpellation of individuals as subjects' through the religious ideological state apparatus, which 'specializes' in relations between God and men, God's subjects, and in specific forms of ceremonies (services, baptisms, marriages and funerals, etc.) which, although they take a religious form, in fact intervene in juridical relations and in economic production, and, therefore, in feudal relations of production themselves. Various discursive formations, always combined in specific forms, intervene as component elements in the realization of ideological class relations. For example, to adopt a historical hypothesis which requires verification, we have on the one hand the country sermon which is reproduced by the lower clergy, who live amongst the peasantry, and, on the other, the sermon the higher clergy preaches to the great nobles. We have, therefore, two discursive formations, with one subordinate to the other. Although they deal with the same 'things' (poverty, death, obedience, etc.), they do so in different ways (e.g., the need for the people to obey the nobility versus the need for the nobility to obey God). They also deal with different 'things' (e.g., work on the land versus the destiny of the nobility).

Finally, it must be stressed that, in historical terms, a discursive formation exists within given class relations; it may supply elements that can be integrated into new discursive formations, which may be constituted within new ideological relations involving new ideological formations. We could, for example, argue (again, subject to historical verification) that although the mentioned ideological formations disappeared as such, they provided ingredients which were remixed in various historical forms of bourgeois atheism and which were reappropriated in the shape of new discursive formations that helped to establish the ideological dominance of the bourgeois class (by incorporating, for example, certain parliamentary discourses of the revolution of 1789).

Here, a difficulty arises with which Marxist theorists are only too familiar: it is difficult to characterize the real boundaries of the real objects that correspond to the concepts we have introduced (in the present case, the concepts of ideological formation, discursive formation and conditions of production). The 'difficulty' is not the result of some

unfortunate accident, but an effect of the contradiction between the nature of these concepts and the spontaneously creationist and classificatory use to which we almost inevitably put them by asking apparently unavoidable questions such as 'How many ideological formations are there in a social formation?', 'How many discursive formations does each of them contain?', and so on. The *dialectical* character of the realities designated by these concepts in fact makes it quite impossible to arrive at such a discrete classification, unless the possibility of being transformed into another object is accepted as an inextricable characteristic of each of these objects, unless, that is, we denounce their discrete character as an illusion.

The relative exteriority of an ideological formation with regard to a discursive formation necessarily finds its expression within that discursive formation itself in the form of non-discursive elements such as representations or images bound up with practices. Indeed, it introduces a discrepancy reflecting that exteriority within the discursive. We refer to the discrepancy between two discursive formations, with the first acting, so to speak, as representational raw material for the second, as though the discursivity of this 'raw material' disappeared from the speaking subject's view (cf. Henry 1971 and 1974). We will refer to this phenomenon as 'forgetting' <sup>1</sup>; it is inevitably inherent in the subjective practice bound up with language. But at the same time, and this is another side of the same forgetting, the process whereby a concrete discursive sequence is produced or recognized as having a meaning for a subject disappears from the subject's view. By this we mean that, in our view, the production of meaning is strictly indissociable from the paraphrastic relationship<sup>2</sup> that exists between sequences, in such a way that a paraphrastic family of sequences constitutes what might be termed a matrix of meaning. This is equivalent to saying that it is on the basis of a relation internal to that family that both the meaning-effect and the relationship with a referent

---

<sup>1</sup> The term 'forgetting' does not refer to an individual's inability to remember something. Paradoxically, it designates something which *was never known* but which *concerns* the 'speaking subject' *intimately*; the subject is 'strangely familiar' with the causes of its determination... but knows nothing about them.

<sup>2</sup> So we do not begin by assuming that there is some 'identity of meaning' common to members of a paraphrastic family; on the contrary, we will posit the view that it is within a paraphrastic relationship that meaning and identity of meaning can be defined (cf. pp. 165, 175-177 below).

implicit in that effect are constituted.<sup>1</sup> It follows that the obviousness of the subjective reading that a text is biuniquely related to its meaning (due allowance being made for syntactic and/or semantic ambiguities) is an illusion which is constitutive of the subject-effect in respect of language and which contributes to the effect of subjection of the subject as an ideological subject. We are in fact arguing that the 'meaning' of a sequence is materially conceivable only insofar as that sequence is regarded as necessarily belonging to one and/or another discursive formation (and this explains, incidentally, why it can have several meanings).<sup>2</sup> This fact that any sequence which 'has meaning' necessarily belongs to a discursive formation that is repressed for (and perhaps by) the subject is concealed from the subject by the illusion that he himself is *at the source of meaning*, that he can grasp anew a pre-existing universal meaning (this explains, among other things, the eternal duality between individuality and universality that characterizes the subject's discursive illusion). Note in passing that the spontaneous hermeneutics that characterizes the subject-effect in respect of language is reproduced, without any fundamental

---

<sup>1</sup> To give an immediate example of what is meant in discursive terms by a 'paraphrastic family', we will describe a 'semantic domain' obtained in a recent study in which ADA69 was used:

Répartition	plus juste meilleure	équitable	des	biens richesses revenus
	More Better	equitable just	distribution of	goods wealth income

It will be seen later (pp. 165-166) that the relations which are represented here are to be interpreted either as symmetrical relations (vertical lines) or as non-symmetrical relations (arrows). It should also be noted that discursive paraphrase is not to be confused with what certain linguists term 'paraphrase' (e.g. passive transformation). We will return to this point below (pp. 176-178).

<sup>2</sup> We would stress that this conception is not to be identified with that of 'plural readings', which suggests the idea of a proliferation of significations in which every subject manifests its singularity. The latter conception obscures the materiality of the discursive, and Trognon (1972: 28) seems to lose sight of its materiality when he writes: "The meaning of a discourse is what we write about it in the problematic we have defined."

change in its nature, by the theoretical elaborations inherent in Chomskyan and post-Chomskyan conceptions of semantics (the inevitable recourse to a universal semantics mobilized by a logic of predicates, which is tantamount to assuming that the problem can be resolved by nullifying the distance between a discursive process and a logical formulation).

These remarks allow us to understand why the model of automatic discourse analysis basically precludes the idea of a semantic analysis of *a* text to the extent that it conforms to the conception of discourse theory which we have just outlined. ADA69 makes a distinction between the linguistic analysis of a discursive sequence and the automatic processing of a set of objects obtained by means of that analysis. This distinction appears to have escaped Fisher and Veron to the extent that they appear surprised to find that Pêcheux, "despite this warning [as to the impossibility of a semantic analysis of *a* text], applies his system of analysis to a single text" (Fisher and Veron 1973: 162-163); the expression 'system of analysis' provides the opportunity for a play on words which conflates linguistic analysis and discursive analysis.

Assuming that we take 'discursive process' to mean the paraphrastic relations internal to what we have termed the matrix of meaning inherent in a discursive formation, we can say that the ADA69 method provides an outline for a non-subjective analysis of meaning-effects (production or reading) in that it cuts through the illusion of the subject-effect and arrives, via a sort of controlled archaeology, at the discursive process. In its present state, the method supplies what might be termed *traces* of the discursive process in our object of study. As will be demonstrated below, the problem that has to be resolved here relates to the fact that there is no direct correspondence between a family of paraphrases (or rather different paraphrastic families or semantic domains) and a logical proposition (or a system of logical propositions). This is not, in our view, an accidental inadequation that can be overcome by adopting a more refined procedure; it relates to the distance between a logical proposition and a discursive process, to the very distance that is nullified in the imaginary by both the spontaneous philosophy of formal logic and positivist idealism in linguistics.

As we have just seen, discursive processes, as conceived here, cannot originate in the subject. Yet they are necessarily realized in that same subject. This apparent contradiction relates to the question of the constitution of the subject and to what we have termed its subjection. In this connection, certain clarifications have to be introduced with respect to

the ambiguous formulations of the 1969 text (Chapter 5). These clarifications concern the concept of 'conditions of production' in particular. The ambiguity arose from the fact that that concept designated both the effect of the positional relations in which the subject is inscribed, and its situation, in the concrete and empirical sense of that term (its material and institutional environment, the roles it adopts, consciously or otherwise, and so on). Ultimately, conditions of production, as defined in the latter sense, would appear to determine 'the situation as lived by the subject' in the sense that they determine the subjective variables (attitudes, representations, etc.) inherent in an experimental situation. We are now in a position to state that these two definitions are as different as the real is to the imaginary, and that what the 1969 text lacked was precisely a theory of the imaginary in relation to the real. In the absence of any theory of that relationship, it was inevitable that positional relations would be confused with the reflection of roles within an institution<sup>1</sup> (and that confusion did indeed arise), and that the term 'apparatus' (as in the concept of 'ideological state apparatus') would be erroneously conflated with the notion of 'institution'. In other words, what was lacking – and what is to some extent still lacking – was a non-subjective theory of the constitution of the subject in its concrete situation as enunciator.<sup>2</sup> The fact that this situation is basically an illusion does not prevent it from being necessary, and we therefore have to undertake the task of at least describing its structure (by providing an outline of the processes of enunciation), and perhaps also that of articulating our description of this illusion with what we have termed forgetting 1.

---

<sup>1</sup> Expressions of the  $I_A(A)$ ,  $I_A(B)$  type (cf. pp. 86ff. above) which were used in an attempt to characterize relations between imaginary formations, readily lend themselves to an 'interpersonal' interpretation of the system of conditions of production. The repercussions of this ambiguity can be seen in the work of a number of authors (Borel 1970, for example). Trognon's suggestion (1972: 164) to the effect that it is the function of automatic discourse analysis to distribute 'elements of discourse' or 'textual units' is quite alien to our project. Finally, we agree with Guespin's (1971) view that the introduction of more mechanisms does not really resolve the question.

<sup>2</sup> The effects of this on the constitution of the corpus are examined below.

## Linguistics as theory of syntactic mechanisms and of processes of enunciation

As we said earlier, the method of automatic discourse analysis (ADA) is designed to reveal traces of discursive processes.<sup>1</sup> Given that the starting point for automatic discourse analysis is a discursive corpus,<sup>2</sup> it is to be expected that it should include a phase of linguistic analysis, as the texts that belong to the corpus are obviously written in a 'natural language'. Developments in the automatic processing of texts show that it is impossible to rely on a statistical study of linearity alone (cf. Markovian processes).

But the adoption of one practice of linguistic analysis in preference to another presupposes a preliminary definition of the nature and role we are attributing to the language. In terms of discourse theory, what in fact is the relationship between discursive processes and the language? The overall perspective is as follows. As discursive processes are at the origin of the production of meaning-effects, the language constitutes the *material site* in which meaning-effects are realized. This specific materiality of the language relates to the idea of functioning (in the Saussurean sense), as opposed to 'function' (see p. 64). The whole problem of linguistics lies in the characterization of that materiality. As we shall see below, it is not enough to view the language as a given consisting of a lexicon and phonological, morphological and syntactic systems (this difficulty is emphasized by the reference to a 'lexical stock' in Haroche and Pêcheux 1972a: 14). We can, however, make use of this inadequate point of view by saying that, under these conditions, the task of the linguist is to characterize this lexicon and these systems of rules and to make them operationally useful, whilst ensuring that uncontrolled semantic considerations do not intervene. Failure to ensure this would imply lapsing into the subjec-

---

<sup>1</sup> It must be stressed from the outset that the terms 'discourse', 'discursive process', 'discursive formation' and 'text' (or sequence) are in no sense interchangeable. Precise definitions of the terms are given below.

<sup>2</sup> 'Discursive corpus' is to be understood as meaning a set of texts (or discursive sequences) of variable length deriving from conditions of production which may be regarded as being stable, or as a set of textual images bound up with a 'virtual' text (with, that is, the discursive process which dominates and generates the various discursive sequences belonging to the corpus). We will return to this question in our discussion of the construction of the corpus (p. 141 below).

tion effect of reading.

The non-subjective analysis of meaning-effects which ADA takes as its objective involves, as we have seen, a phase of linguistic analysis, and we will now show that the status of that phase is highly problematic. The question in fact revolves around *the role of semantics in linguistic analysis*. In the perspective defined above, we cannot start a linguistic analysis with what should result from the comparison of objects derived from this very analysis. In other words, the linguistic analysis on which ADA relies must be essentially morpho-syntactic. It must allow a specifically linguistic treatment of phenomena of subordination, coordination and determination, which results in a 'delinearization' (or 'desyntagmatization') of texts. There is, therefore, no question of introducing a 'world-view' based upon a universal and *a priori* semantics, as to do so would be tantamount to including within the functioning of the language historically determined discursive processes which cannot be regarded as being coextensive with the language unless we identify the language with ideology.<sup>1</sup>

Having said that, it remains true to say that at the moment the conditions for such a morpho-syntactic analysis have not been clearly defined, and that recourse to an implicit semantics cannot be ruled out. It is as though a morpho-syntactic analysis would necessarily bring into play elements which are normally described as being 'semantic'. The initial description of ADA ignored this aspect.<sup>2</sup> This can be justified both in terms of the precarious nature of the linguistic 'solutions' acknowledged then, and in terms of the priority given to the need to combat the idealist conception which sees the language as a worldview, a perception of the world, or even as the origin of the world.

In its extreme form, the linguistic position inherent in ADA69

---

<sup>1</sup> The fact that discourse is the site of the articulation between ideological processes and linguistic phenomena must not lead to the assumption that the language is an ideological superstructure. This caveat, which is one of ADA's theoretical starting points, has been seen by some critics as a prohibition (a normalization!) which restricts the linguist to subaltern tasks ('linguists keep out!'). It will be seen below that, on the contrary, the distinction between the language and ideology results in a helpful reformulation of the linguistic problematic in that it makes it possible to take into account processes of enunciation.

<sup>2</sup> The 'syntactic' analysis of relatives is a classic example of the surreptitious re-introduction of semantic considerations. This point is made by Fuchs, Milner and Le Goffic (1974); it is also discussed by Henry (1975) and Grésillon (1975).



would imply that syntax and semantics are two autonomous and well defined domains, and that the lexicon and grammar are two disjunctive domains as well. But this is obviously not the case. Indeed, the linguistic phase of ADA in its present state clearly illustrates the difficulties involved in any such claim. Far from preventing linguistic analysis from being contaminated by semantics, the syntactic rules that are applied surreptitiously introduce an uncontrolled appeal to meaning. But does this mean that this semantics, to which syntactic analysis cannot but appeal, is precisely the discursive semantics to which has been referred above? If that were the case, we would have to admit that linguistics has no theoretical autonomy at all, as we would end up with nothing more than our original inputs. We do not believe that this is in fact the case. The situation appears to us to be bound up with the philosophical heritage that is necessarily conveyed by grammatical categories, even in their most neutral, modern and technical form. What is lacking at the moment is *a theory of how the language functions materially in relation to itself*, or in other words a theory in which systematicity is not defined in opposition to the non-systematic (such as in the dichotomy between the language and speech) but articulated with processes. If we agree to call a theory of the material functioning of the language a 'formal semantics', we can say that what linguistics lacks is a formal semantics, and it is clear that this semantics is not the same thing as 'discursive semantics'. The expression 'formal semantics', which is borrowed from Culioli (1968), can be defined as being the ultimate level of linguistic analysis, and in that sense it attains the specific site within the language which corresponds to the construction of the subject-effect. This means that ADA, which is intended to cut across the subject-effect, must identify the *point* at which it cuts across it *within the language*. A desire not to reproduce that effect in the practice of objective analysis is a legitimate desire, but it would be a mistake to overlook its existence within our object of study.

This necessarily leads us to the question of *enunciation*, and the following remarks on that topic are not out of place, given the way in which idealism currently 'occupies' the question, and the various obstacles that arise as a result.

If enunciation is defined as the necessarily ever-present relationship between the subject of the enunciation and its utterance, then this clearly introduces a new form of the illusion that meaning originates in the subject or that the subject can be identified with the origin of meaning. This time, the illusion comes into play at the level of the language itself:

the discourse of the subject is organized by a direct or deferred reference (or by the absence of a reference) to the situation of enunciation as experienced subjectively by the speaker – the speaker's 'I here and now'; his points of origin on the axes of person, tense and localization. Any language activity requires these anchoring points to be stable for the subject: if that stability is lacking, both the very structure of the subject and his language activity will be affected.

When we speak of obstacles, we refer both to the subjective empiricist illusion which is reproduced in linguistic theory and to the formalist illusion which sees enunciation merely as a system of operations. In a discussion of the notions of the enunciating subject and of the enunciation situation, Fiala and Ridoux (1973: 44) write: "[...] nor should they be reduced to being a mere support for formal operations; we must always try to isolate their formal content, so as to avoid the ever-present traps of formalism". In an earlier paper, in which they discuss Benveniste's propositions, Hirsbrunner and Fiala note that:

Semiotics and semantics in fact appear to be a linguistic transposition of the philosophical categories of potentiality and actuality [...]. Once again, the mediation is effected with the help of an ambiguous notion, namely the notion of enunciation, which is defined formally [...] but justified in philosophical terms: "Enunciation causes the language to function thanks to an act on the part of the individual user". Here we find ourselves faced with the essential difficulty in the Saussurean approach; in our view, this is the major obstacle standing in the way of a Saussurean theory of discourse. The view that the language is simply a system of signs has certainly been superseded, but the price of its supersession is the introduction into linguistic theory of two notions which it attempted to reject when it constituted itself as a science; namely, the subject and the subject's relation with the social world. The paradox is, however, that whilst both these notions fill a gap in the conceptual apparatus, they have no specific theoretical status. By contrasting the freedom of the individual subject with the necessity of the system of the language, by assuming that the language is a mediation between the subject and the world, that the subject appropriates the

world through the intermediary of the language, and that it appropriates the language through the intermediary of the apparatus of enunciation, Benveniste simply translates into linguistic terms philosophical notions which, far from being neutral, belong directly to the idealist current. (Hirsbrunner and Fiala 1972: 26-27)

We will demonstrate below how we propose to escape both the problematic of enunciation and the circle of idealism.

Current difficulties in theories of enunciation stem from the fact that those theories usually reflect the necessary illusion<sup>1</sup> that constitutes the subject; in other words, they are content to *reproduce the illusion of the subject at a theoretical level* by following the tradition of Bally, Jakobson and Benveniste and by adopting the idea that the enunciating subject is endowed with choices, intentions and decisions (and here we come close to the notion of speech).<sup>2</sup>

The above remarks as to the material functioning of syntactic mechanisms in relation to themselves allows us to specify our understanding of enunciation as follows. We can state that processes of enunciation consist of a series of successive determinations which gradually constitute an utterance, and which are characterized by their ability to posit what is said and to reject what is not said. The enunciation therefore establishes boundaries between what is 'selected' and gradually specified (the elements which constitute 'the universe of discourse') and what is rejected. We thus obtain an inverted image of 'everything which it would have been possible for the subject to say (but which he did not say)' or of 'everything which conflicts with what the subject says'. The distance between the 'rejected' zone and the conscious mind varies, and questions an interlocutor may ask in order to get the subject to clarify 'what he means'

---

<sup>1</sup> The term 'necessary illusion' was first introduced by Le Goffic (cf. Fuchs, Milner and Le Goffic, 1974).

<sup>2</sup> As Robin (1973: 81) notes, this concept of enunciation means basing linguistics on the idealist 'psychological subject': "The linguistics of discourse has not succeeded in effecting the decentring of the subject of discourse because it has not succeeded in integrating either the ideological subject of historical materialism or the psychoanalytic subject into its theory of the subject."

may force him to reformulate its boundaries and to investigate it anew.<sup>1</sup> We propose to term this effect of partial occultation 'forgetting 2', and we will regard it as the origin of the subject's impression that he really expresses his thoughts ('I know what I am saying', 'I know what I am talking about').

It follows from the above that the study of the marks of enunciation is central to ADA's linguistic analysis, and that it results in major modifications of the concept of the language. Firstly, the lexicon cannot be regarded as a 'stock of lexical units', or as a mere list of morphemes which are unrelated to syntax; on the contrary, it must be regarded as a structured set of elements articulated with syntax. Secondly, syntax no longer constitutes a neutral domain of purely formal rules; it constitutes the mode of organization (specific to a given language) of identifiable traces of the enunciation. From this point of view, syntactic constructions have, therefore, a 'signification' which must be isolated.

From this perspective, it is important to specify the link between the two forgettings which we have described as forgetting 1 and forgetting 2 respectively: What is the relationship between a family of paraphrastic sequences which constitute meaning and the 'unspoken' elements, given that both are occluded?

### **The language, ideology and discourse**

We will now consider the phenomena we have described as forgetting 1 and forgetting 2. The two forgettings are profoundly different. The subject can *consciously enter* zone 2 and it in fact constantly does so by reflecting on its discourse, anticipating its effects and taking into account the discrepancies it introduces into the discourse of the other.<sup>2</sup> Insofar as the subject corrects itself in order to clarify for itself what it has said, to specify 'what it thinks' and to formulate this more accurately, we can say that zone 2, which is the zone of the processes of enunciation, is characterized by conscious-preconscious mechanisms. Forgetting 1, in contrast,

---

<sup>1</sup> Cf. the notion of 'anti-paraphrase' (Fisher and Veron, 1973).

<sup>2</sup> Zone 2 is the domain of what are sometimes known as 'discursive strategies'. The most important are rhetorical questions, tendentious reformulations, and the manipulative use of ambiguity.

relates to a zone to which the subject has no access and it is, for that very reason, constitutive of subjectivity within language. We can therefore advance the hypothesis that this repression (which affects both the discursive process itself and the interdiscourse<sup>1</sup> with which it is articulated by relations of contradiction, subservience and encroachment) is unconscious in character, in the sense that ideology is constitutionally unconscious of its own existence (and not merely absent-minded about itself or continuously escaping itself; cf. Haroche and Pêcheux 1972b).

The opposition between these two types of forgetting is not unrelated to the above-mentioned opposition between the empirical concrete situation in which the subject finds itself – a situation characterized by an imaginary identification in which the other is an other ego (other with a small o) – and the process of the interpellation and subjection of the subject – which goes back to what Lacan metaphorically terms the Other (with a capital O). In that sense, monologue is a particular type of dialogue and interpellation.

In other words, we are positing the view that the relationship between forgetting 1 and forgetting 2 relates to the relationship between, on the one hand, the non-subjective conditions of existence of the subjective illusion and, on the other, the subjective forms of its realization.<sup>2</sup>

In using Freudian terminology, which makes a distinction between the conscious-preconscious system on the one hand and the unconscious on the other, we are not claiming to have resolved the question of the relationship between ideology, the unconscious and discursivity; we are simply noting the fact that a discursive formation is constituted and bounded by what lies outside it, in other words by *something that cannot be formulated within it because it determines it*. We would also stress that this constitutive exteriority must on no account be confused with the subjective space of the enunciation, the latter being an imaginary space which allows the speaking subject its *displacements within the realm of what can be reformulated*, so that it continuously goes back over what it has formulated, and recognizes itself "in a reflexive or preconscious relationship with

---

<sup>1</sup> 'Interdiscourse' is to be understood as meaning the 'specific exterior of a given discursive process' (cf. p. 119 above). The term refers, that is, to processes that intervene in the constitution and organization of a discursive process.

<sup>2</sup> On this point, and on the distinction between 'unconscious law' and 'conscious/pre-conscious rule' in particular, cf. Herbert (1968). See also the discussion in Robin (1973: 100).

words, which makes them appear to us to be the expression of things" (Safouan 1973: 71). The term *preconscious* relates to Freud's first topography, and it is not present as such in the second topography. Yet it is primarily the second topography which provides the framework for Lacan's reworking of Freud, and it is to that reworking that we refer here. We will return to this theoretical inconsistency in a future study, where it will be explained, reworked and reduced (Pêcheux 1982a: 123-126).

The 'inequality' between the two forgettings corresponds to a relation of dominance which we can describe by saying that the non-asserted precedes and dominates the asserted (Culioli, Fuchs and Pêcheux 1970).

Moreover, we must not lose sight of the fact that the repression which characterizes forgetting 1 ultimately regulates the relationship between what is said and what is unsaid in the zone of forgetting 2, where the discursive sequence is structured. This is to be understood as meaning that, as Lacan puts it, 'all discourse is an occultation of the unconscious'.

To conclude this general introduction, let us state that two complementary errors are to be avoided in connection with the term 'discourse', as it is used in the expression 'discourse theory'. The first error is the conflation of discourse and speech (in the Saussurean sense of that term); according to this view, discourse is the realization of verbal acts on the part of a free subjectivity which escapes the system (of the language). We reject that interpretation by stating once more that the theory of discourse and the procedures it involves cannot be identified with a 'linguistics of speech'. The second error is the converse of the first in that it 'bends' the meaning of the term in the opposite direction, and sees discourse as a social supplement to an utterance and therefore as a particular element within the system of the language which has been overlooked by 'classical linguistics'. According to this view, the level of discourse can be integrated into the language by, for example, a particular type of competence which varies in accordance with the social position of the speaker. This is tantamount to accepting the idea of a plurality of languages, or to taking *literally* the expression 'Bosses and workers do not speak the same language'. The expression is correct in political terms, but in linguistic terms it is debatable.

Given these distortions of the reality designated by the term 'discourse', we find it useful to make a distinction between the linguistic

base and the discursive process which develops on that base.<sup>1</sup> In our view, it is that distinction alone which makes it possible to take into consideration relations of contradiction, antagonism, alliance, absorption and the like, between discursive formations belonging to different ideological formations without necessarily subscribing to the mythical existence of a plurality of languages belonging to those different formations.

## 6.2. Automatic discourse analysis: criticisms and new perspectives

### **Construction of a corpus on basis of its dominant conditions of production**

The above remarks clearly indicate that the 'conditions of production' of a discourse are not brakes or filters which interfere with the free functioning of language in the sense that, for example, the resistance of the air affects the trajectory of a moving body whose theoretical displacement – the trajectory it would follow if it were reduced to a point moving through a vacuum – could be predicted by kinematics. There is, in other words, no such thing as a socially empty space in which the laws of a general semantics (laws of 'communication', for example) can be developed and into which supplementary constraints of a social nature can be re-introduced as corrective parameters. Everything that has been said above is intended to make it clear that the discursive must be regarded as a social process whose specificity resides in the materiality of its base: the linguistic materiality.

The meaning of the expression 'conditions of production of a discourse' must therefore be explained in detail if we are to avoid certain errors of interpretation which have arisen as a result of the ambiguity of certain of our formulations. It is to be noted, first of all, that the term 'discourse' itself can refer both to what we have termed a 'discursive process'<sup>2</sup>, and to an oral or written sequence of variable dimensions (usually larger than a sentence). Because of its immediately 'concrete' character,

---

<sup>1</sup> Processes of enunciation constitute elements within the base which allow processes to develop in relation to that base.

<sup>2</sup> As, for example, used in the expression 'the discourse of a science'.

the latter reality has been designated a 'discursive surface' (in Chapter 5), but that expression has the double disadvantage of suggesting that sequences can be processed at the level of their 'surface' forms, in the Chomskyan sense of that term, and of being too brief a designation for what is in fact the linguistic surface of a discourse. Because of the misunderstanding as to the meaning of 'discursive surface', we would stress the need to make a distinction between two types of desyntagmatization which are, in our view, inherent in the linguistic domain and the discursive domain respectively. Linguistic desyntagmatization (or desuperficialization) refers to the material existence of the language, which is characterized by the non-linear structure of syntactic mechanisms and, at a deeper level, by all the elements affected by forgetting 1; discursive desyntagmatization can only begin to go beyond forgetting 2 if it is based upon the linguistic operation we have just mentioned. These remarks allow us to propose the following terminological distinctions:

- *Linguistic surface* is to be understood as meaning an oral or written sequence of variable dimensions, usually greater than a sentence. A linguistic surface is a 'concrete' discourse, in other words an empirical object affected by forgettings 1 and 2 to the extent that it is the site where they are realized in the form of a double illusion which is subjectively experienced as being both coherent and necessary.

- *Discursive object* is to be understood as meaning the result of the transformation of the linguistic surface of a concrete discourse into a theoretical object, into, that is, a linguistically desuperficialized object produced by a linguistic analysis designed to dispel the illusion of forgetting 2.

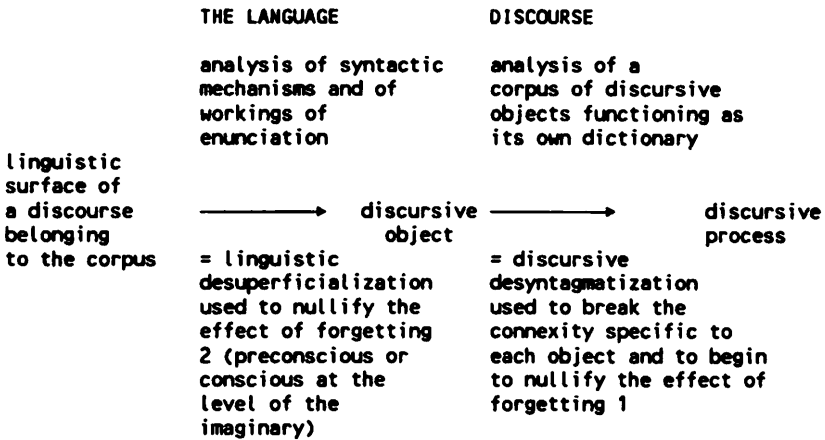
- *Discursive process* is to be understood as meaning the result of the controlled interrelating of *discursive objects* corresponding to *linguistic surfaces* that represent stable and homogeneous conditions of production. Access to the discursive process is gained through a desyntagmatization referring to the zone of forgetting 1.

It should, however, be stressed, that in order to overcome forgetting 1 completely we must do more than reveal the underlying discursive formation – the 'matrix of meaning', certain traces of which can be identified with current ADA procedures. We must also grasp the relations of discrepancy between that discursive formation and the interdiscourse which determines it – the 'operational' solution to this problem has not been reached yet.

The above terminological remarks can be summarized by the



following schema:



We can now return to our examination of the expression 'conditions of production of a discourse', which we described as being potentially ambiguous. In the light of the above, it becomes apparent that the expression can be taken to mean either the *determinations which characterize a discursive process* or the *multiple characteristics of a 'concrete situation'* which lead to the production – in either the linguistic or the psycholinguistic sense of that term – of the linguistic surface of a concrete empirical discourse. The ambiguity is of course the same as the ambiguity we noted in connection with the opposition between institution and apparatus: in both cases, what is at stake is the need to recognize the discrepancy between the imaginary, whose existence cannot be ignored simply on the grounds that it is imaginary, and the exterior which determines it. To that extent, it is apparent that what is really lacking is a theory of the 'concrete situation' as an ideological formation in which the 'lived experience' is informed and constituted by the structure of ideology, in which, that is, it becomes that structure in the form of a received interpellation, to use Althusser's formula.

A theory of the 'concrete situation', which could *in theoretical terms* relate determinations to their imaginary effect, would provide a valid starting point for the operations used in *constructing the corpus*. The

absence of such a theory<sup>1</sup> affects the foundation of our practice of constructing corpora (and procedures in which several corpora are combined) by leaving room for the empiricist temptation of trying to find an impossible articulation between an 'experimental' psychology and historical materialism. We can, however, state that the link between the imaginary and the exterior which determines it involves the concept of dominance. We will argue that a corpus consists of a series of either linguistic surfaces (concrete discourses) or discursive objects (which implies that linguistic analysis intervenes in the definition of the corpus in a different way than we will discuss here). Those surfaces are *dominated* by stable and homogeneous conditions of production. We are, that is, assuming that any 'concrete' discourse is in fact a *complex* of processes deriving from different conditions. Constructing a corpus with reference to this dominance therefore means eliminating all those individual elements which appear in one 'concrete' discourse and not in another, even though both are *dominated* by the same conditions, on the grounds that such elements are alien to the process we are studying. This does not, of course, preclude the possibility of taking differences as an object of study, but such differences must always be regarded as differences between conditions of production and never as individual differences.

The theoretical principles and practical considerations regulating the construction of a corpus are strictly *external* to the 'technical' principles and considerations of the ADA method itself. In other words the theoretical responsibility for constructing the corpus (or system of corpora) has, in theory, nothing in common with the specific responsibility for the ADA procedure: the responsibility for making a *non-subjective reading*. It must, however, be immediately added that responsibilities taken at the extra-discursive level (various sociological and historical hypotheses) which govern the construction of the corpus are not without their *effects* on the results obtained with ADA. Indeed, we might say that the results will reflect the hypotheses at the level of the discursive effects that are identified, but this does not mean that the results are purely and simply a *transparent* reflection of the extra-discursive hypotheses used to construct the corpus. Unless we make a distinction between these responsibilities, we will inevitably come to the conclusion that ADA is a *circular* process and

---

<sup>1</sup> This theory must be elaborated at a general level, and not at the level of conditions and findings of any particular discursive analysis, which would result in a vicious circle.

that it is therefore possible that "the result will be the very content that was introduced and organized by the categorization", as Borillo and Virbel suspect (1973: 1). Unless we wish to fall back on the aporia of an idealist theory of ideology, we must remember the important point that the discursive is irreducible to the linguistic or the ideological.

Once we have established that the materiality of ideology cannot be identified with discursive materiality (insofar as the latter is a particular element within the materiality of ideology, which — in the area that concerns us — implies that the conditions of the construction of a corpus cannot be *exclusively intra-discursive*), it is possible to take the criticisms formulated by Borillo and Virbel (1973) into consideration. Their primary criticism is the claim that the principle on which the construction of the corpus is based in Gayot and Pêcheux (1971) is rather unmotivated. Its construction consists of the selection of sentences containing a certain key word. In Gayot and Pêcheux (1971) this key word is 'circumstances'. The criticism is twofold:

(1) On the one hand, it is, *from a methodological point of view*, difficult to see in any detail how "justifications of an extremely variable nature" (Borillo and Virbel 1973: 10) lead to the selection of the word 'circumstances' rather than any other word.

(2) On the other hand, the decision to select sentences (sequences separated by two full stops) containing the desired term is also arbitrary, and this too helps to invalidate the procedure.

To deal with these two points in turn:

(1) The former criticism appears to be justified in general terms. It does not suffice to answer this criticism by stressing the non-methodological and purely theoretical character of the procedure which led to the selection of the term 'circumstances' (it derives, in fact, from historical materialist theory). A 'concrete analysis of a concrete situation'<sup>1</sup> should in fact have led us to identify *a system of interrelated observable points* which could be methodologically projected onto a procedure interrelating *several corpora* in order to investigate the differences between them. In

---

<sup>1</sup> The concrete analysis of a concrete situation implies, in our view, that discursive materiality in an ideological formation must be regarded as an articulation of processes. In this connection, it is worth recalling a comment made by Fiala and Ridoux (1973: 15): "The text is not a set of utterances bearing one or more meanings. It is primarily a process which develops in multiple forms within given social situations". For our part, we prefer to use the term 'discourse' rather than 'text'.

other words, it seems that it is now possible and necessary not to restrict ourselves to the analysis of *one* corpus which is arbitrarily constructed on the basis of *one* key word, but that we must systematically make use of an analysis of the internal differences that a procedure may reveal. At the methodological level, this presupposes that we have the means to combine *n* corpora into a single corpus in order to study the differences that are so induced. The so-called 'compaction' procedure introduced into the programme now makes it possible to do this (cf. p. 166 below). This procedure was not available at the time when the research of Gayot and Pêcheux was carried out. Our approach has evolved in this direction: access to the discursive process specific to a corpus seems to be guaranteed to a large extent by the study of its specificity within a system of hypotheses that are operationalized by treating a *complex of corpora* with the help of the compaction procedure we have just mentioned. Ultimately, this is a matter of both studying the productivity of a given hypothesis and of deducing its characteristics from the discursive process under examination. It should be added that a preliminary statistical analysis (such as the analysis of co-occurrence suggested by Geffroy et al. 1973) may be helpful to an initial formulation of hypotheses. It is also possible to envisage the application of *a priori* statistical checks on the homogeneity of any corpus submitted for analysis or on the rules governing the closure of a corpus.

(2) The second criticism deals with the relatively arbitrary character of the segmentation procedure, which is based upon the criterion of the sentence. It has to be said from the outset that this criticism, which is fully justified, indicates a major difficulty, and that it is at the moment impossible to say how it can be resolved. What are the empirical limits of a 'discourse' within a given complex sequence? Do combinations of processes necessarily correspond to juxtapositions within the linearity of the sequence? All that can be said on the basis of the theoretical presuppositions described above is that any 'literary' notion which refers to the 'internal unity' of an 'oeuvre', a text, or a section is null and void. The principle of an *expressive link* between the organic unity of form and the intentional unity of content – or, for that matter, a project or a meaning – is a literary myth (essential to the classic form of text interpretation) which reproduces the subjective illusions we discussed earlier. We can only note that this problem, which has also been raised by Provost-Chauveau (1970: 135), relates to the limitations of the linguistics of the sentence and to the urgent need for a *theory of the inter-sentence level*.

The remarks we have just made are relevant to two different procedures used to construct a corpus or a system of corpora that can be submitted for ADA: the experimental and the archival method. The experimental method 'stages' a setting that reproduces certain of the effects that characterize a 'concrete situation'<sup>1</sup> (and in which 'the imaginary' can vary). The problem of the segmentation of discourse arises with the archival method<sup>2</sup> but

---

<sup>1</sup> In terms of positional relations inscribed in class relations.

<sup>2</sup> The ADA procedure has been applied in its archival mode by the historian Gayot in various separate studies of eighteenth-century texts. He makes the following comments (personal communication):

"In all cases, the discursive processes identified corresponded to what a historian would expect to find from an informed reading of the texts. Two points are, however, worth stressing.

(1) In the 'archival' approach that we used, the corpus consisted of sequences extracted from a given text and containing a theme chosen because of the determinant role hypothetically ascribed to it by me as a historian.

Although the theme was chosen with great care, the results were sometimes 'poor' in that the discursive mechanisms brought into play in the chosen term's context provided nothing that could not have been obtained on the basis of a long familiarity with the texts analyzed (cf. Gayot 1973: 242). I think that this will always be the case when the corpus is *centred* on a term which is so powerful that it creates a vacuum around it and evokes only derivative and directly subordinate processes (in addition to the principal process it contains). We thus have an 'experimental' challenge to the truism that a word which was undeniably important to those who used it is necessarily productive in terms of the discursive processes with which it is bound up.

(2) Other studies, e.g. Gayot and Pêcheux (1971), on the other hand, showed that the *dominant* form of selection and combination of words bound up with the use of the chosen term left room for *relatively autonomous secondary discursive processes which a simple reading might regard as primary*, for example: Saint Martin's reflections on the progress made by the masses, and the reflections of eighteenth-century free masons on the theme of equality and fraternity between men. ADA in fact demonstrated that these secondary mechanisms were picked up, and integrated or digested, so to speak, into the general organization of the discourse which, for Saint Martin, was governed by the trust that was to be given to God's elect alone and which, for the free masons, was governed by the services the masonic brotherhood rendered to the established or traditional order, and not to some future order.

ADA thus allows a historian working on a collection of utterances to reconstruct and *differentiate* between the primary and secondary rules which produce those utterances. It is vital to be able to differentiate between them in this way if we are to avoid the traps set by secondary discursive processes which cast a zone of shadow around the dominant process. Personally, I feel that the principal advantage I gained from using ADA was the ability to avoid the traps laid by the curious conservative or counter-revolutionary ideology subtended by masonic and Martinist discourses in the eighteenth century, an ideology which has been

does not arise with this experimental method. The notion that it is preferable to approach a system of hypotheses operationalized in terms of a 'complex of corpora' is tantamount to the view that the archival method is a quasi-experimental approach. We would argue that the archival form is to be regarded as a derivative or 'bastardized' form of the experimental procedure. The point is worth stressing because certain formulations in the 1969 text (Chapter 5 above) gave rise to a number of 'non-directive' interpretations.

It should, however, be added that these guidelines do not in themselves resolve any of the basic problems as to the nature of materialist experimentation in the domain that concerns us. We will simply point out that the socio-historical practice which provides the inevitable reference point here is profoundly ambiguous: to be more accurate, that practice is highly 'unstable' in that there is no barrier to prevent it from lapsing into *historical materialism* on the one hand or *social psychology* on the other. It is, however, more likely to lapse into the latter than the former. If, in other words, the only barrier is 'the experimental method', we will almost inevitably lapse into a social psychology of situations and into the idealism which is inseparable from such a psychology.

## **Linguistic analysis**

### *Objectives of a linguistic analysis of discourse*

The expression 'linguistics of discourse' is sometimes used to designate a type of approach to language which can, at least partly, escape certain effects of the theoretical limitations of 'traditional' linguistics, whose primary shortcoming is its tendency to see its object within the framework of what classical grammar (and particularly Latin grammar) terms 'the sentence'. This implies both a fixation on the structure of the utterance and a sort of blindness to what is now called 'enunciation'. At the same time, discussion tends to centre on the inter-sentence level, a point to which we will return later (p. 159). As we shall see, it is of vital importance to our whole project that we take into account both the reality of *discourse* and the type of decentring it introduces into *linguistics itself*.

It would, however, be a mistake to think that our conception of discourse analysis simply implies the application of a new linguistics which has been stripped of the prejudices of 'traditional linguistics'. To take that view would imply that the change we are effecting is simply a matter of adopting a new approach to the same object and of introducing new research requirements. Whilst all this is true, it ignores the objective we are assigning to a *linguistic analysis of discourse*. The research techniques are an essential precondition for the realization of our project; there are, however, other preconditions that have yet to be specified, above all the mode of their articulation. To go straight to the main point, it seems useful to point out that linguistics (or 'pure linguistics') usually sees the result of its practice as a *theoretical discourse* which may take the form of either a general theory or a monograph, but which always concerns a more or less specific object and which takes the modality of a description of that object, an exposition of its workings or a theory of its mechanisms. Discourse analysis, as distinct from the practice of the linguist, has two specific features. The first is that this practice *necessarily* involves an algorithmic procedure<sup>1</sup>, and this leads to an essential difference in the results obtained with it (it should be noted that we are referring here to discourse *analysis*, and not to the *theory* of discourse it implies). In that sense, discourse analysis is, as we shall see, similar to what is sometimes called 'applied linguistics' in that, at the empirical level, both methods involve the use of the computer. But this first characteristic is not in itself enough, and it must immediately be added that discourse analysis displays a second specificity: the object it uses to obtain its results is not a linguistic object, but a socio-historical object and *the linguistic intervenes within it as a presupposition*.<sup>2</sup> And it is in our view this relation of

---

<sup>1</sup> *Algorithm*: a controlled sequence of operations which can be performed by a calculator within a finite space of time and which produces a determinate result on the basis of a given starting point.

<sup>2</sup> This is why it seems to us that it is erroneous to use the term 'linguistics of discourse' to designate what is in fact a *linguistics of texts* (or even of *a text*) just on the grounds that it goes beyond the framework of the analysis of the sentence. The expression linguistics of speech is also often used in the same sense. Our reluctance to use the expression has been explained earlier.

application<sup>1</sup> which determines the difference between linguistic practice and discourse analysis; for discourse analysis, linguistic analysis is an essential preliminary phase (there can be no analysis without a linguistic theory and practice), but it is not enough in itself in that it is subordinate to a second phase which involves a *change of terrain*. This is not an application of linguistics to linguistics (i.e. an internal application effected within a given theory, as when linguistics uses linguistic data as an intermediary stage within a procedure designed to *realize* a mechanism that has been revealed at the level of the theoretical discourse of linguistics; examples might include the use of an algorithm to generate syntactic forms, or a procedure for the automatic classification of the syntactic-semantic features of a list of verbs), but the application of linguistic theory to *an external field*. In these circumstances, it is quite understandable that the 'pure linguist' should react with a certain annoyance, rather as though he were an artisan who sees the content of his work escaping him; he inevitably feels that the constraints imposed by this 'external field' are *too great*. To that extent, discourse analysis meets with the same resistances and objections on the part of 'pure linguistics' as automatic documentation and translation, to which it is related at the theoretical level in that it is dependent on the same foundations. The common feature of discourse analysis and automatic documentation and translation is the need for a recognition grammar capable of meeting both the internal theoretical requirements of linguistics and the requirements of what we have termed the 'external field'.

We will argue that the recognition grammar required for discourse analysis must meet two requirements:

(a) The grammar must be able to produce an *algorithmic* representation of what has been described above as a *linguistic surface*, and the algorithmic representation must constitute what we have termed the

---

<sup>1</sup> As the use of the term 'application' may introduce a certain ambiguity, a distinction will be made between technical and theoretical applications.

- A *technical application* consists of using a theory and equipment to produce an effect, an object or a result in practice (e.g., the theory of semi-conductors is technically applied in the manufacture of transistors).

- A *theoretical application* refers to the intervention of one theoretical discipline within another (e.g., the application of mathematics in physics), or to the application of a discipline to itself. It should be noted that in the case of computational linguistics it is, unfortunately, not always possible to make a distinction between technical and theoretical applications.

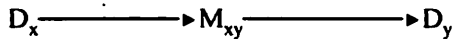


corresponding *discursive object*.

(b) The representation (the discursive object that has been produced) must be such that a *calculation* can be carried out on the relationship between the various discursive objects that have been produced; the objective of that calculation is to reconstruct traces of the semantic-discursive processes characteristic of the corpus under examination.

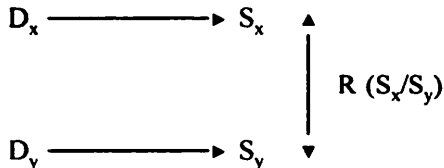
It is quite clear that, given the present state of developments in linguistics, many difficulties have to be overcome before this problem can be resolved. Grammars capable of recognizing specialist texts in English, Russian or French (for purposes of documentation or translation) have been in existence for some years, but they are vulnerable to the criticisms voiced by 'pure linguistics' (which are usually justified), and the solution we have proposed is as vulnerable to those criticisms as any other. To take the examples mentioned earlier, both translation and documentation involve the use of a recognition grammar, and they can be shown to correspond to the following schemata, whose stability is guaranteed by their very 'obviousness':

#### *Translation*



$D_x$  and  $D_y$  represent two 'versions' of the 'same' text in languages  $x$  and  $y$ , and  $M_{xy}$  designates the underlying 'deep representation' common to  $D_x$  and  $D_y$ .

#### *Documentation of texts*



$D_x$  and  $D_y$  represent two sequences that are to be related to one another,  $S_x$  and  $S_y$  the descriptions linguistic analysis provides of those sequences, and  $R(S_x/S_y)$  the result of relating them by means of the documentary or comparative method.

Appearances to the contrary notwithstanding, we would argue that

the second schema is *more elementary* than the first, and that, logically, it in fact exists prior to it. The term 'deep representation', which appears in the 'translation' schema, is not present in the second schema; it is replaced by the relationship between structures  $S_x$  and  $S_y$  which are not assumed to constitute 'deep representations' of  $D_x$  and  $D_y$  respectively. We would add that the idea of making a 'deep representation' correspond to a linguistic surface (as in the 'translation' schema) involves psycho-semantic presuppositions which are radically challenged by a discursive conception of signification. The 'translation' schema implies a purely extra-linguistic conception of the relationship between syntax and semantics, as the analysis represented by the horizontal arrow leads from the morpho-phonological surface of the text to its semantic structure or deep representation  $M_{xy}$ , which is assumed to be common to all languages and to constitute a *logical bridge* or 'pivot' (its logical structure is made up of arguments and predicates extracted from a universal logical language, or in other words from 'the structure of the human mind'). It is, on the other hand, clear that in the 'documentation' schema the 'deep representation', or rather the  $R(S_x/S_y)$  relationship which replaces it, is not regarded as the outcome of a linguistic analysis, but as the product of a specific operation *carried out on the preliminary results* of a linguistic analysis, namely  $S_x$  and  $S_y$ . When we speak of 'linguistic analysis', we are adopting the latter perspective, in which  $S_x$  does not designate a 'deep representation' but simply *the result of the application to the initial linguistic surface of a morpho-syntactic delinearization*, which may take an algorithmic form. In other words, the analysis we are describing as 'linguistic' and which constitutes the *second phase of the automatic discourse analysis* has as its sole goal the production of what was described above as a discursive object and it has nothing to do with 'deep representations'.

We can now define with greater precision the two requirements defined earlier.

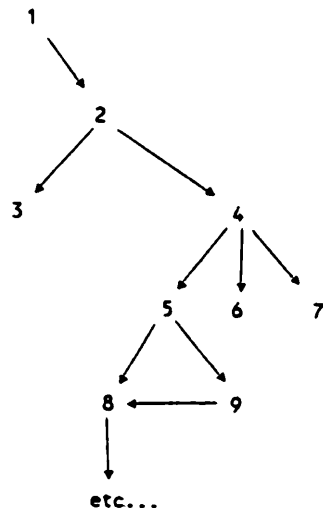
(a) Given linguistic surface  $D_x$ , the recognition grammar must be capable of producing a morpho-syntactically coherent representation ( $S_x$ ), that is, a linguistically delinearized representation which reconstructs the *syntactic non-linearity* making up the object of what we have designated as 'forgetting 2'. The coherence of this representation, which is based upon the relative autonomy of morpho-syntactic structures ('the functioning of the language with respect to itself'), presupposes that, given the representation supplied by the recognition grammar, it is possible, if not to reconstruct the initial text, to at least decide which of two linguistic surfaces

corresponds to representation  $S_x$  (i.e. to decide which it derives from) if their proximity is such that the present state of linguistic theory allows to distinguish between them.

(b) It is also essential that representation  $S_x$  should constitute a possible input into the comparison represented by the vertical arrow in the 'documentation' schema. It must immediately be added that this condition, which is external to the linguistic analysis as such, is the source of major difficulties which we can summarize by saying that it is very difficult to compare complex structures by using algorithmic procedures.

The 'linguistic analysis' phase of ADA has been described in Haroche and Pêcheux (1972a). We will not review the syntactic analysis procedure specific to this phase in any detail here. It does, however, seem necessary to recall briefly its main characteristics, which are basically similar to those outlined by Harris. We can summarize the procedure (which is designed to be amenable to the application of an algorithm) by saying that a representation of a given linguistic sequence of variable length is produced in the form of a graph with a single root.

This graph is composed of points and arrows. The points consist of *elementary utterances* of canonical dimensions (indicated by numbers). The arrows are relations connecting two utterances. The relations can be given various values (of determination, in the case of relatives, adjectives or noun complements; of subordination or coordination, in the case of the various temporal and/or logical relations that may affect a pair of utterances).



Unlike the trees, tree-structures or graphs typical of the transformational and generative grammars of Harris and Chomsky, the

nodes (or points) in this graph are neither syntactical nor lexical units, but what might be termed 'point-relations' or subgraphs which are reduced to points at the level of the structure of the graph representing the whole sequence. This means that we have *two* systems, with one imbricated in the other: a system of *utterances* and a system of *inter-utterance relations*. The objects of the first system thus become elements in the construction of the second system. It is to be noted that, within the framework of generative grammars, embedding (and, in more general terms, recursivity) is represented by using a more complex form of the graph of the utterance, so that the utterance gradually becomes *a matrix for the sentence* in the same way that a skeleton is an abstract, stripped-down form of a complete body. To use rather less picturesque terms, we can say that embedding is the condition which guarantees that there is a theoretical and methodological homogeneity between the utterance and the more complex formation; all relations therefore occur within the same system, which is characterized by the imbrication of the graph of the embedded utterance in the graph of the matrix-utterance. In the case of the procedure we are discussing, on the other hand, utterances are not reintroduced into another utterance, and this implies that the question of recursivity must be resolved in some other way (namely by the transition from the intra-utterance system to a second system of relations between utterances). It further implies that an utterance  $U_1$  can determine another utterance  $U_2$  by an asymmetrical relation which is equivalent to the embedding of  $U_1$  in  $U_2$ , and, more generally, that a portion of a graph organizing several utterances can be *equivalent* to a point in the general graph.

We will now elaborate this point in some detail by discussing relations between utterances as relations of dominance. Take the following sentence:

*It seems to me that the Church is still living in a world of its own and that it forgets the difficulties and problems of life*

The following utterances can be extracted from it:

1. *It seems to me that S (= something)*
2. *the Church is still living in a world of its own*
3. *the Church forgets the difficulties*
4. *the difficulties (are) of life*
5. *the Church forgets the problems*

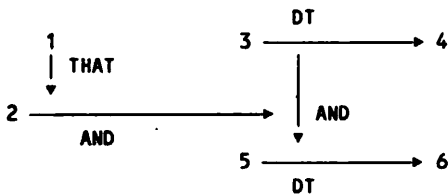
### 6. the problems (are) of life

If we then introduce the connectors THAT, AND and DT (the latter allows an utterance to determine the N in another utterance), we can represent the initial sequence as follows:

sequence	=	1 THAT /A/
/A/	=	2 AND /B/
/B/	=	/C/ AND /D/
/C/	=	3 DT 4
/D/	=	5 DT 6

We now find that the utterances are related constructions of utterances, indicated here by the capitals /A/, /B/, etc., and that they are imbricated in one another.

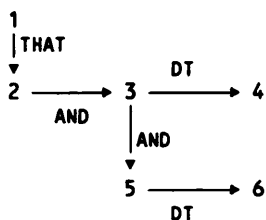
It is, however, possible to represent these dependencies by means of a schema which includes only *utterances* and *relations between utterances*. In this schema, the intermediary constructions /A/, /B/, etc., do not appear as such. In our view, this is an essential precondition for the third phase of ADA, the discursive desyntagmatization by computer. The schema, which takes a 'combinatory' form is as follows:



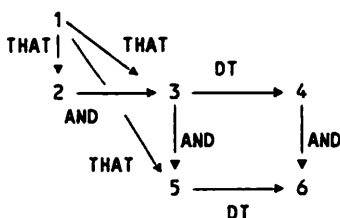
It will be noted that this schema is strictly equivalent to the bracketed notation:

1 THAT (2 AND ((3 DT 4) AND (5 DT 6)))

In our discussion of the following phase of the procedure, we will return to the question of whether or not an algorithmic comparative procedure can be carried out on representations of this type. For the moment, we will simply move from the above representation to a list of binary relations between utterances. First we transpose the above schema into the following:



Note that the distributive character of *THAT* with regard to the utterances in construction /A/ has disappeared; it is therefore necessary to re-establish the links that existed between 1 and 3 and 1 and 5 respectively. Links that are re-established in this way are known as saturations. We thus arrive at the following saturated graph:



The input data for the following phase (of discursive desyntagmatization) is therefore:

(a) a list of elementary utterances, each of which is a sequence of eight 'morpho-syntactic categories', each with its respective value (cf. p. 115 above):

F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
1) 0000	∅	S	SEEM	∅	TO	∅	EGO
2) 0000	T	CHURCH	LIVE	∅	IN	A	WORLD
3) 0000	R	CHURCH	FORGET	∅	*	T	DIFFICULTIES
4) 0040	R	DIFFICULTIES	E	∅	OF	∅	LIFE
5) 0000	R	CHURCH	FORGET	∅	*	T	PROBLEMS
6) 0040	R	PROBLEMS	E	∅	OF	∅	LIFE

*Note:* the R in the DET<sub>1</sub> column indicates the reprise of an earlier determination.

(b) a list of binary relations:

1	THAT	2
1	THAT	3
2	AND	3
1	THAT	5
3	AND	5
3	DT	4
5	DT	6
4	AND	6

### *Criticisms of the linguistic analysis phase of ADA*

Having recalled what we see as the objectives of the 'linguistic analysis' phase of ADA and the general manner in which we are currently attempting to realize them, we can now describe the various criticisms that have been addressed to it without any risk of confusing what are, in our view, *valid criticisms* (criticisms which it may be vital for discourse analysis to take into account in both its theory and its practice) and '*criticisms*' which *in reality mask a theoretical regression to a stage preceding discourse theory*. This distinction must be made primarily in relation to general criticisms addressed to the analytic procedure as a whole, and we will therefore begin with criticisms of this kind and, more specifically, with the question of the *implicit manipulation of the text out of hidden semantic considerations*. This point is raised independently, but in related ways, by a number of commentators, including Trognon (1972), Fisher and Veron (1973), and Borillo and Virbel (1973). These commentators insist that what are usually termed 'semantic' considerations *inevitably* intervene in the 'morpho-syntactic' analysis we have described, and that, as these considerations remain implicit, it is probable that they will affect any analysis and will result in inconsistencies which will mask phenomena or produce artifacts which will have detrimental effects on later phases in the ADA approach. We would like to make it quite clear that we accept that this criticism is *perfectly justified*; it does concern our problematic, and it relates to a vital point we made earlier when we described the first requirement that a recognition grammar has to meet (pp. 150-151 above). If 'undefined semantic operations' (Fisher and Veron 1973: 167) do intervene in the analysis, the coherence and stability of the results will suffer, and it cannot be guaranteed that the representation associated with a given sequence can be reproduced. This clearly affects the requirement for

biunicity as expressed above. In that sense, we are in complete agreement with Fisher and Veron (1973: 163) that "[i]f the method [of linguistic analysis] prevents certain properties from being identified, they can never be recuperated". In other words, a purely stenographic 'coding' of the surface, which filters what has to be retained *and* what can be left aside, cannot meet our 'first requirement'. It must be recognized that certain unresolved questions in linguistic theory do affect the analysis, and that they do not do so solely at a *peripheral level*; as we will demonstrate in a moment, they affect the very principle of the procedure.

We feel that, in these circumstances, we are more than justified in criticizing the criticisms that have been addressed to us at another level. These criticisms concern *the relationship between semantics and linguistics*, and we have already encountered them in our discussion of our double 'categorization' (to use the terminology adopted by Borillo and Virbel 1973). When, in their turn, Fisher and Veron dismiss our project on the grounds that it is "always associated with the *supposed* difference between syntax and semantics" (1973: 167, emphasis added), they speak *as if* the reinscription of semantics within the field of linguistics were a recent and decisive achievement of contemporary linguistics, and even as if semantics were an extra-linguistic reality. We obviously cannot accept the criticisms that have been addressed to us in this respect, and can only refer the reader to our earlier remarks on the subject (cf. pp. 129-131 above).<sup>1</sup>

A criticism which in turn we regard as being fully justified is the point raised by Provost-Chauveau (1970: 137) that the heterogeneity of the linguistic references reveals the absence of any overall theoretical reflection on syntactic phenomena. We were working very rapidly, and with the means at our disposal, and we were fully aware of the fact that the linguistic 'solutions' we were proposing were designed to provoke 'pure linguists' into making a transformational critique. We would, however, stress that this theoretical heterogeneity, which generated certain inconsistencies and artifacts, means that not all our 'solutions' are of the same 'theoretical age', and that it is therefore not to be confused with an inevitable conflation of *morphological* and *syntactic* characteristics, or with

---

<sup>1</sup> Logically enough, this leads our critics to accuse us of a 'sleight of hand' (Trognon 1972) and of passing off data as findings. Their position means that they fail to make a distinction between the 'linguistic' semantics which implicitly intervenes in syntactic analysis, and the semantic-discursive processes whose traces are recorded by the phase of discursive desyn-tagmatization of the ADA procedure.



the coexistence of *lexical terms* and '*meta-terms*' within the representation (and anyone who expresses surprise at this overlooks the fact that the language is 'its own metalanguage').

It is, moreover, obvious that our method of syntactic analysis is far from complete (which implies that, when it is implemented, it will probably take the form of a 'self-adjusting' automaton), and that certain 'local' inconsistencies will have to be eliminated. We certainly do not take the view that this heterogeneity will gradually be reduced thanks to a sort of 'reformism' which will patiently clear the field of unresolved problems. On the contrary, we think that the difficulties we will encounter (and which are encountered by all projects for a syntactic analysis) constitute a stubborn and consistent block based largely upon what might be called the theoretical dominance of the sentence. In that sense, it is not enough to make *an act of theoretical good will* that would allow for 'an unprejudiced approach' (Fisher and Veron 1973: 169). What is required is a transformation of the object of linguistics itself. The 1969 text evoked the need for a serious study of the *inter-sentence* rather than really engaging with it. Current developments in linguistic research and the link that is gradually being established between the inter-sentence [*inter-phrase*] and the paraphrase within the framework of 'text linguistics' suggest that linguistics is now on its way to solving a problem which is, as we have said, the key to many other problems. This also relates to the difficult question of *anaphora*, which necessarily involves complex syntactic-semantic operations which could not easily be resolved by automatic means (including the identification of links between pronouns and substantives, apposition, whole propositions presented by a new lexical term, and so on).

Without wishing to underestimate the extent of these difficulties, we do not think that they bring us to the central issue around which all the other difficulties revolve. The central issue – and it is here that the 'theoretical dominance of the sentence' comes into play – is in our view the theoretical assumption which links together *sentence*, *proposition* and *utterance*. The most frequent and pertinent criticisms concentrate on this issue, and rightly so in our view. Most critics basically insist that the 'linguistic analysis' phase coincides very closely with a traditional 'sentence analysis'. Admittedly, it relies upon the idea that the sentence is organized on the basis of clauses both hierarchically (in terms of main and subordinate clauses) and sequentially (in terms of coordination). This conception, which is based upon the notion of a *tissue of nodes* which constitute so many 'grains of enunciation', leads in practice to scruples.

Because of the arbitrary nature of the segmentation, the analyst hesitates between a desire to provide a faithful representation of linguistic reality and the need to 'fiddle' which may produce 'elementary utterances' which cannot be asserted, enunciated or even interpreted.

The 'theoretical dominance of the sentence' (and the implicit theory of the utterance that subtends it) has also considerable effects on relations between utterances; it would seem therefore that the distinction made between *connectors of determination* (such as DT, p. 155 above) and *connectors of coordination and subordination* (such as AND resp. THAT) has to be challenged on the grounds that transformations make it possible to move from one to the other, primarily by means of prepositions.

### **Analysis of the discursive process**

As indicated above, the third phase of ADA concerns operations of discursive 'desyntagmatization' once the linguistic delinearization is completed. We will begin by recalling the principal aspects of this phase, as implemented by our present calculation programme. We will then describe the various criticisms that have been addressed to the procedure, and will suggest possible ways in which they can be at least partially answered. Finally, we will raise the more general question of the relationship between the two phases in order to determine whether or not the two requirements we set earlier (pp. 150-151 above) have in fact been met by mechanisms which are both necessary and sufficient, and in order to determine what new solutions are available to us in the event of the answer to that question being negative.

#### *Principal aspects of the discursive desyntagmatization*

Assuming that phase 1 has determined the conditions of production of a corpus (or preferably, as we have seen, more than one corpus), that they have been materially realized in the form of a family of 'linguistic surfaces', and that phase 2 has supplied a *desuperficialized representation* corresponding to each of those linguistic sequences, we now have *n* corpora capable of providing inputs for phase 3 in that each corpus consists of *two lists*: a list of 'elementary utterances' (EUs) and a list of 'binary relations' (BRs) corresponding to the set of graphs making up the

corpus. Leaving aside all the precautions that have to be taken to ensure that the EU and BR data is free of errors (such as the nominalization of the data and the renumbering of the utterances), we will restrict ourselves here to giving an account of *the two basic stages of our approach*:

- (a) a comparison of the graph by means of a 'point-by-point' scanning designed to provide each point of comparison with a mathematical value expressing the result of the comparison;
- (b) a procedure for reconstructing sub-structures on the basis of the data obtained in stage (a).

We will now describe these sub-phases in turn, followed by a third aspect that must be mentioned (see (c) below).

- (a) The first sub-phase consists of a *systematic point-by-point comparison, taking the binary relation as the basis for the comparison*.

It should be noted that this scanning can take several forms, depending on which preliminary conditions are introduced. *All the binary relations in the corpus can be compared with one another*, but it may also be useful to restrict the comparison in two related ways. The first restriction consists of limiting the comparison to one *between discourses*; here, we take the opposite decision to Harris who, as we know, defines the procedure for discourse analysis with reference to a single text, in other words to *intra-textual mechanisms* (Harris 1952). In adopting this position, Harris advances the implicit hypothesis that a text repeats itself, and syntagmatically reproduces structures which can be superimposed by an analytic operation which organizes them into a paradigm. This explains why Harris attributes a particular methodological importance to propaganda or advertising texts and, more generally, to stereotypical forms of discursivity. In his perspective, this corresponds to the need for the text to be *its own dictionary*, whereas in the perspective we are describing it is the *corpus* which acts as its own dictionary. It is in fact in the transition from the intra-discursive to the inter-discursive (and here Harris's work paves the way) that we find the real conditions for the realization of the desubjectivization of discursivity. The methodological purpose of the restriction we have described is to make it possible to compare the results of an inter-discourse comparison (in which a discourse is compared with itself only in indirect terms, insofar as, that is, *two sub-sequences* are compared through the intermediary of a sub-sequence belonging to another discourse) *and* a comparison in which the sub-sequences under consideration are compared directly.

Then it is possible to study the conditions a discourse must meet if the selection of one or the other of these options is to have *no influence* on the results. The work of Del Vigna and Dupraz (1974) appears to be leading them in this direction, and their research should make it possible to specify certain of the formal aspects which characterize the autonomy of a discursive process in terms of its inter- and intra-repetitiveness.

The second restriction with respect to the point-by-point comparison of binary relations concerns the nature of the connectors linking the relations under consideration. It is possible either to compare all binary relations (be it within the limitations imposed by our first restriction) or to restrict the comparison to *pairs of binary relations with the same connector*. The latter option, which we originally proposed in ADA69, has been criticized on the grounds that treating connectors separately excludes them from the comparison. When two connectors are identical, their distance is considered *nil*. If they are different, however, their distance is *a priori* considered *too great* for permitting a comparison between the two binary relations.

We can now briefly explain what is meant by a point-by-point comparison. If we assume that neither restriction on the scanning applies, it will readily be seen that, if the list of BRs contains  $n$  binary relations, there will be  $n(n-1)/2$  comparisons, given that no BR is being compared with itself (which does not preclude the possibility of comparing two identical BRs!), and that the result of comparison  $C(BR_i/BR_j)$  will be identical to the result of  $C(BR_j/BR_i)$ . Each point of comparison takes the form:<sup>1</sup>

$$\begin{array}{ccc} U_m & K_i & U_n \\ U_p & K_j & U_q \end{array}$$

The proximity characteristic of a point of comparison is calculated as follows. Take the left-hand utterances ( $U_m/U_p$ ) and the right-hand utterances ( $U_n/U_q$ ). Each pair of utterances can be given a *Boolean vector* using a sequence of 0s and 1s to express the result of a column-by-column comparison of the literal content of each utterance, with 1 indicating

---

<sup>1</sup> [In the following schema, the binary relation  $BR_i$  is represented by two EUs,  $U_m$  and  $U_n$ , connected by connector  $K_i$ ;  $BR_j$  is represented similarly.]

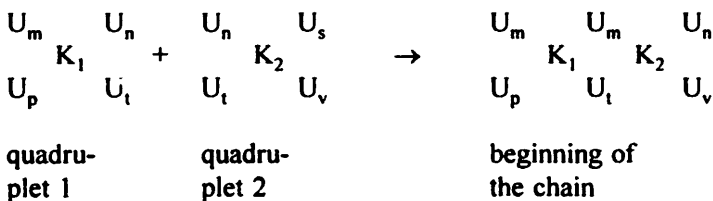
identity and 0 indicating difference. In the most general case, an additional 1 or 0 will be added to indicate the identity or non-identity of  $K_i$  and  $K_j$ . Each vector is multiplied by a factor assigning a certain weight (positive or nil) to each column. The sum of the weighted values so calculated expresses the 'proximity' between the two binary relations that are compared.

During this first stage of the data-processing operation, we give the  $n(n-1)/2$  points of comparison an arithmetic value. Our procedure consists of comparing these values with an arbitrarily fixed threshold value and thus breaking the set of points down into two subsets, one of which contains *all the points with an arithmetic value equal to or greater than the threshold value*. This set makes up a *table of quadruplets*.

(b) The second sub-phase consists of the *construction of 'semantic domains'*.

This sub-phase centres on a procedure for reconstructing the sub-structures characteristic of the corpus under analysis. The procedure is based, in principle, on the idea that it is possible to combine the syntagmatic and paradigmatic relations that are implicitly contained in the 'table of quadruplets' just defined. We will describe how '*chains*' are constructed from quadruplets, and *domains* from chains.

*Chains* correspond to the *syntagmatic* phase of reconstruction, and the procedure used to construct them is as follows. Given a table of quadruplets, each containing two left-hand utterances and two right-hand utterances, two quadruplets begin to form a chain if the left-hand utterances of the one are identical to the right-hand utterances of the other. For example:



We thus generate a 'table of chains' from a table of quadruplets. It should be pointed out that *residual* quadruplets which cannot be integrated into

chains are admitted individually into the set of chains at the end of this operation.

*Domains* correspond to the *paradigmatic* phase of reconstruction. The rule for their formation presupposes the intermediary definition of *the 'sequence' as semi-chain*. In the above chain, for example, a distinction can be made between the semi-chains ' $U_m, U_n, U_s$ ' and ' $U_p, U_r, U_v$ '. Having introduced this intermediary definition, we will now say that two chains belong to the same domain if they have a semi-chain in common. The application of this rule demonstrates that sequences which do not form a quadruplet together can be included in the same domain. In such cases, we can say that their inclusion is transitive, bearing in mind that the transitivity is imposed by the rule for the formation of domains, and that it is in no sense a 'demonstrable' property of the *domain as object*. Finally, chains which are not associated with another chain in order to form a domain will be integrated into the 'table of domains' at the end of this stage. It is not unusual for a quadruplet which remains syntagmatically and paradigmatically isolated to become a chain consisting of one quadruplet and then *a domain made up of two sequences*.

- (c) One final aspect of the present procedure should be mentioned before we turn to its principal results: the *representation of the discursive process by means of the semantic domains* whose mode of generation has just been described.

A 'table of relations between domains' provides two types of information. These can be defined as, respectively, *paradigmatic* relations defining relations of intersection and inclusion between domains, and *syntagmatic* relations characterizing the paths typical of the discursive process of the corpus. We will simply note here that the former lead to the construction of groups of domains (or 'hyperdomains'), whilst the latter allow us to plot a *graph of the discursive process*; this is a graph whose nodes consist of domains or hyperdomains.

We will now discuss the question of 'domains' as *basic elements* used to obtain a representation of the discursive process. Semantic domains are groupings of  $n$  sub-sequences extracted from discourses of a corpus during the syntactic delinearization effected in the linguistic phase. The relationship between these subsequences is *constructed* as one of *substitutability*, but our procedure cannot further specify it. We originally took the view that such substitutions were necessarily indices of equivalence, or in other words that the  $n$  sequences found within a discourse

constituted *n* semantic forms equivalent to a single proposition, in its logical sense (cf. pp. 103-106 above).

In Haroche and Pêcheux (1972a: 46-49) we drew attention to the fact that relations of substitution obtained in this way did not seem to be reducible to pure equivalence, and we therefore made a distinction between two basic types of substitution:

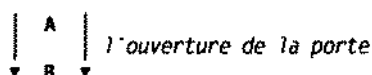
(1) Symmetrical substitutions. If A and B (which may be morphemes, syntagmas or utterances) can be substituted for one another, the path leading from A to B is identical to that leading from B to A; this implies that  $A = B$ , the equivalence between them being of the dictionary type or an equivalent contextual effect. In the first case, A may be contextually synonymous with B. Alternatively, it may be an adequate *metaphor* for B (and vice versa), in which case the equivalence is produced during the process itself, and cannot be referred to as a dictionary-type effect.

(2) Oriented substitutions. In this case, the path from A to B is not identical to that leading from B to A. The two are therefore not equivalent to one another, but we can move from one to the other or deduce one from the other. In other words, the relationship between A and B can be described, at its most general level, as being *metonymic*. The existence of this 'non-symmetrical relationship of deducibility' between A and B relates, in our view, to the possibility of the syntagmatization  $A + R + B$  (or  $B + R' + A$ ), where R (and R') indicate the existence of a syntactic relationship between elements A and B. Given the crude result

<p>A = <i>une catastrophe se produit</i> [a catastrophe occurs]</p> <p>B = <i>les personnes évitent</i> [people avoid]</p>	<p><i>l'ouverture de la porte</i> [opening the door]</p>
--	--

we can advance the hypothesis that there is an implicit syntagmatization between A and B, and that it takes the form: 'C'est parce que *une catastrophe se produit dès X que les personnes évitent X*' [*It is because a catastrophe occurs when X occurs that people avoid X*]. We must assume that it is formulated somewhere (not necessarily in the corpus we are studying), and we therefore have to ask what it is, in the 'specific exterior' of a given corpus, that intervenes in the substitutions that occur in it and

orients them.<sup>1</sup> The result can be represented as follows:



All these points are taken up and expanded in the discussion of the relationship between semantics and discursive processes in Pêcheux (1982a: 115-117).

To conclude this commentary on the current procedure for discursive desyntagmatization, we will summarize the meaning of several options which have recently been introduced.

- The procedure known as 'compaction' allows a corpus (A + B) to be submitted for analysis once A and B have been analyzed separately, and it therefore allows us to make a systematic study of the differences between the two corpora, and to determine which domains belong *specifically* to A and B respectively, which result from the agglomeration of domains A and B, and which are produced by the compaction of A + B.

- The distinction between two modes of comparing binary relations (interdiscourse alone respectively inter- and intradiscourse) appears to open up the way for important research which will allow the repetitive or stereotypical nature of a corpus to be characterized by a study of the extent to which it partly reproduces itself. This does not preclude the possibility that this problematic might be related to the *discourse of the other* within the speaker's own discourse.

- As we have already noted, the inclusion of connectors in the calculation of the distance between two binary relations raises a difficulty which has often been pointed out. The relationship between different types of connectors has yet to be studied in terms of their substitutability. The study of that relationship may have implications for the measurement of distance in that it may make it possible to give each pair of connectors a

---

<sup>1</sup> Fisher and Veron (1973: 166) refer to this example and criticize the above representation on the grounds that it "ignores the *surface* appearance of the expression '*c'est parce ...que*' in the sentence". Basically, the authors have been so careless as to mistake a hypothetical syntactic connection for a *sentence from the corpus analyzed*. They even criticize the way the sentence was syntactically segmented into two elements. For obvious reasons, no such segmentation was carried out!



value. This would presuppose the elaboration of a Cartesian table of connectors listing both the values given to each pair and their relations of compatibility and permutability (for example:  $U_a K_i U_b \rightarrow U_b K_j U_a$ ).

### *Criticisms of the present procedure*

It is the *relatively arbitrary* nature of the procedures carried out that has attracted most criticism. Borillo and Virbel (1973) note that many of the operations 'are regarded as natural', whereas they can be identified as belonging to a family of operations – which implies that the selection of our or another operation has to be justified. For his part, Trognon (1972) criticizes our mixture of *empiricism* and *formalism*. We wish to make it clear that, in principle, these criticisms seem to us to be quite valid in that the mathematical model which will allow our 'choices' to be identified and justified has yet to be constructed.

We will now look in turn at the two stages of the procedure we have described and at the criticisms they have provoked. Whenever possible, we will also outline ways to avoid similar criticisms in our future work.

(a) The problem of the value to be given to the comparison of binary relations.

Three criticisms, of varying importance, have been raised in connection with this problem.

Firstly, we would point out that, although the description of our 'points of comparison' has been described as 'natural', it is in fact fairly arbitrary; it would be an error to assume that, on the grounds that the procedure is carried out point by point, it includes *all possible comparisons*. That this is not the case can clearly be seen if we interpolate an utterance into a *sequence*. Take the quadruplet:

$$\begin{array}{cc} U_a & U_b \\ & K \\ U_m & U_n \end{array}$$

If the distances calculated for these components give an arithmetic value

greater than the threshold value, it is concluded that the sequences  $U_a K U_b$  and  $U_m K U_n$  belong to the same domain. If we assume that we have the sequence  $U_m K' U_p K' U_n$  instead of  $U_m K U_n$ , it appears that, because of the interpolation of sequence  $U_p$ , the result of the comparison we first postulated cannot be positive. In practice, it seems difficult to avoid this, unless we simultaneously adopt 'heuristic' solutions which allow to limit the procedure's field of extension by 'skipping' certain utterances in the sequence, and hence not to restrict the comparison to pairs of binary relations.

The second criticism is of greater immediate importance. It concerns the both empirical and arbitrary character of 'distance' (whatever its zone of application), that is, of the system of weighting used to multiply the Boolean vector obtained as a result of a column-by-column comparison of two utterances. It might be said that this is *arbitrary* in that the significance of the weighting principle has not been clearly defined in either linguistic or mathematical (statistical) terms, to say nothing of the significance of the difference in 'weight' between the columns. More specifically, it is unclear whether the identity or non-identity of two 'contents' has the same significance irrespective of what those 'contents' are.<sup>1</sup> This does not mean, of course, that we have to revise the principle of ADA itself, which, as we know, rules out the *a priori* constitution of classes of morphemes, syntagmas or utterances. It does mean that we have to investigate in linguistic terms the different workings of what are normally termed 'closed classes' and 'open classes'. Provost-Chauveau (1970: 138) discusses this aspect of the procedure, and says of the value attributed to each category that

the approximate determination of this value relies (at the moment) upon the notion of probability;  $DET_1$  which is 'chosen' from within a restricted set of terms (articles, demonstratives, etc.), has the value 2, whilst  $N_1$ , where lexemes are selected from within a much larger set, has the value 5.

In making this point, Provost-Chauveau 'is ahead of' ADA69, which

---

<sup>1</sup> Cf. homonyms such as 'to grasp' (to take and hold with your hand ≠ to understand intellectually), discussed by Michel Morin (1973: III, 12) in his comments on the results obtained with ADA.

makes no mention of probability, but it now seems to us that it would be useful to proceed along the lines that she is suggesting, provided that certain ambiguities are removed. We would therefore like to submit the following idea for discussion: rather than processing every co-occurrence in a homogeneous manner, regardless of the category in which it appears, might it not be more appropriate to make a distinction between two fundamentally different forms of processing, applying to 'closed' and 'open' categories respectively? One could, it would seem, readily envisage a form of processing applying to categories DET<sub>1</sub> and DET<sub>2</sub>, F, P and CONNECTOR in which every pair of elements (including, of course, pairs of identical elements) could be given an *a priori* value which could be integrated into the general calculation of distance. As for 'open classes' (essentially N, V and ADJ), it may be possible either to retain the present procedure or to establish a '*learning system*' in which substitutions which have already been identified are re-injected into the data in the form of a *meta-term* which takes the place of both substituents. This would allow us gradually to build up an 'auto-dictionary' which could record *n* levels of equivalence with a view to identifying other equivalences at level *n+1*. (It will be noted that the procedure suggested here is very similar to that used by Harris.) It should be added that it would be perfectly possible to combine this procedure with an *a priori* determination of key words (on the basis of preliminary statistical studies) and with the ascription of values to their co-occurrence.

The third point we would like to make relates to the use made of the distance that has been calculated. We have demonstrated above that this stage results in the final ascription of a value to each point of comparison. The question now arises as to whether the distribution of values ascribed to points of comparison might not be relevant to operations other than the dichotomization to which its application is currently restricted. We know that this is an 'all or nothing' procedure; a point of comparison is either *entered* or *not entered* in the table of quadruplets. We thus lose information which would allow us to differentiate the quadruplets according to the threshold value at which they are realized. It is not impossible that the study of this distribution in each corpus would provide valuable information, including, perhaps, an estimate of the optimal values of the threshold value for the corpus under consideration. It should also be noted that such a distribution would allow substitutions to be serialized, depending on whether they apply to a morpheme, a syntagma or an utterance. This might be of great importance for the realization of the 'maximum

context' algorithm described below (see p. 172). To bring this discussion of the definition of the distance between two objects to a provisional close (we are still dealing with the case of binary relations), it should be added that the arithmetical means that is currently in use is not the only possible type of measure that can be used. Mention should also be made of Lerman's work (1970) on these problems and on the more general question of methods of automatic classification involving the measurement of distance. One aspect of the construction of the mathematical model mentioned earlier will involve situating the specific requirements of ADA within the family of formally possible solutions.

(b) The problem of the construction of domains

Provost-Chauveau (1970) and Trognon (1972) independently raise what we see as the central question. It concerns the reference to semantics implied by the expression 'semantic domains'. Provost-Chauveau raises the issue of whether it is legitimate to argue that "substitutions do not change the semantic interpretation" (1970: 137), which relates our problematic to the problematic of paraphrase in transformational theory. For his part, Trognon challenges ADA69's statement to the effect that "two sequences belonging to the same domain receive the same interpretation". Although he had no knowledge of our results at the time of writing, and therefore did not know that we did in fact find incongruous comparisons and artifacts due to the *formal* (formalist) character of the procedure we used, he did foresee this problem. Is it necessary to go so far as to say that the relations involved concern only the proximity between sentences and that they have nothing to do with 'meaning-effects'? We now take the view that the question is more complex in that a distinction has to be made between *pure syntactic artifacts*, which can, in theory, be eliminated by correcting the syntactic analysis and/or modifying the weighting system which supplies the co-occurrence containing the artifacts, and *semantic phenomena* of substitution which, as we have already had occasion to say, cannot be reduced to a 'single semantic interpretation' in any circumstances. We have in fact drawn a distinction between two types of mechanism which should, in our view, both be described as semantic: a relationship of *equivalence-substitution* which relates to the logical stability of a formal metalinguistic system, and *oriented substitution* which constitutes the condition of possibility of a later equivalence or an equivalence in an embryonic state. This is tantamount to saying that equivalence results from the disappearance, forgetting or erasure of orientation, and that a logical

paraphrase (which preserves the meaning) is a particular example of the workings of a meaning-effect. To put it another way, we might say that metonymy (which is linked to syntagmatic orientation) always tends to be degraded to synonymy by the erasure of syntagmatization, though this obviously does not preclude the possibility of synonymies (or metaphors) being 'sutured' by new syntagmatic relations.

In the event, we are of the opinion that the principal difficulty does not stem from the need to justify the use of the term 'semantics', but from the fact that, given the present state of the procedure, neither type of relation is produced by an algorithm that can be computerized, as is obvious from the schema below.

### Example 1

Print out

```
674 0000 R REPARTITION E PLUS a a EQUITABLE 91 675 0000 R REPARTITION E a DE LS BIEN
921 0000 R REPARTITION E a a MEILLEUR 91 922 0000 R REPARTITION E a DE LS RICHESSE
921 0000 R REPARTITION E a a MEILLEUR 91 923 0000 R REPARTITION E a DE LS REVENU
970 0000 R REPARTITION E a a JUSTE 91 971 0000 R REPARTITION E a DE LS BIEN
```

Interpretation

Répartition	plus juste meilleure	équitable	des	biens richesses revenus
	More Better	equitable just	distribution of	goods wealth income

### Example 2

Print out

```
479 0000 a X AUGMENTER a LS POSSIBILITE 92
480 0000 R POSSIBILITE E a DE EPANOUISSEMENT 92
481 0000 a EPANOUISSEMENT E a a INTELLECTUEL
723 0000 R GOUVERNEMENT OFFRIR a TTLS POSSIBILITE 92
724 0000 R POSSIBILITE E a DE CROISSANCE 92
725 0000 R CROISSANCE E a a INTELLECTUEL
```

Interpretation

le gouvernement offre X augmenter	les toutes les	possibilités de	épanouissement croissance	intellectuel
the government offers X to increase	the every	possibility for	intellectual development	growth

We can, however, state that two conditions must be met if such an

algorithm is to be realized:

- Zones of substitution must be identified by an automatic process; this has yet to be done.
- Orientation must be defined by identifying syntagmatizations attested in other domains belonging to the results, or in the corpus of data, or even in the interdiscourse of the corpus or corpora under examination.

In the absence of the elements that would allow us to deal with the problem, we can say no more about the second condition. We can, however, further specify the first condition.

A 'semantic domain' which the present programme can calculate may be considered to be a list of sequences of the same length which are assumed to exhibit relations of substitution; the first operation to be carried out is the reconstitution of these zones of substitution (or, to be more accurate, the prevention of the loss of data that occurs during the dichotomization of the set of 'possible quadruplets'). We define the 'maximum context' of a domain as the set of  $n$  elements that can be found in the same position in at least two sequences of the domain, such that  $n$  has a maximum value for the domain in question. In any domain under consideration, the counterpart of this maximum context is a zone of commutation containing at least two elements.

Where there is only one maximum context, the algorithmic operation we envisage consists of replacing the zone of commutation with an indexed meta-term, i.e., of regarding the sequences (at least two) as though they were one, and of re-applying the algorithm. In cases where several maximum contexts appear simultaneously, the successive results obtained in this manner will be integrated into a union. It goes without saying that, as we indicated earlier, the availability of the information produced when calculating distances for all points of comparison would represent a major step towards the practical realization of the algorithm. This would then raise *pertinent questions* providing the basis for research into relations of syntagmatization (the second condition mentioned above).<sup>1</sup>

(c) Our remaining remarks on the subject of the construction of domains are more general, and apply to the *statistical validity* of the results obtained, and to their *presentation*.

---

<sup>1</sup> Obviously, there is a relation between the 'maximum context' algorithm and the 're-injection' procedure mentioned earlier (p. 169).

As far as the results obtained in the form of domains and relations of domains are statistically valid, we will simply point out once more that the principle of ADA itself lies *outside the field of statistical considerations*. This obviously does not preclude the possibility of carrying out certain statistical investigations. We will now briefly describe the most useful of those investigations.

(c1) Regarding *elementary utterances and binary relations*:

- calculation of the relative frequency of lexical items in each morpho-syntactic category of utterances, including 'form of utterance' (F);
- identification of the combinations of  $N_1$  and  $N_2$  in a matrix of incidence;
- analysis of the structure of the graph: relative frequency of connectors and of relations of saturation; average length and complexity of both lateral sequences and central sequence.

This investigation should, in our view, be carried out in relation to both statistical properties common to the whole corpus and differential properties of corpora. It also seems appropriate to examine possible similarities between these results and those obtained with a corpus of discourses in their 'natural' state.

(c2) Regarding the *results obtained from the ADA processing*:

By relating the results to the number of utterances, relations and discourses, we propose to study:

- the distribution of values of paradigmatic distance on the basis of the threshold value and pattern, the density (ratio of number of quadruplets retained to number of possible quadruplets);
- the number of quadruplets, chains and domains;
- the structure of domains (number of discourses intervening in their construction, 'optional' discourses, i.e., discourses which do not intervene in any domain) and the structure of dependency between domains.

(c3) Finally, we would point out that the comparison of the characteristics of a corpus (both in its 'natural' form and as a set of discursive objects) and the characteristics of results obtained from processing may lead to the elaboration of a *preliminary evaluation procedure* which will allow us to predict whether or not ADA processing is applicable to a given corpus. As far as the presentation of results is concerned, it would seem that the major criticism that has been raised relates to the 'table of relations between domains'; this requires subsequent manual operations which can easily be computerized as they consist of grouping domains linked by intersections or inclusions into 'hyperdomains', and of reconstructing the

graph representing the paths between the domains and hyperdomains belonging to the corpus. It is to be noted that one of the theoretical effects of the constitution of hyperdomains is the possibility of grouping together sequences of unequal length, and that this may lead to the renewal of the problematic of substitution.

*Towards a transformation of inputs of the discourse analytic phase*

To conclude this discussion, it seems desirable to return briefly to the general question of the relationship between the linguistic and the discourse analytic phase of the procedure. It will be recalled that this relationship is articulated around two requirements (cf. pp. 152-153 above). The representation of the linguistic surface must, that is,

- (1) restore the non-linearity of syntactic structures and, more generally, display the guarantees of stability and coherence described earlier, and
- (2) be so structured as to allow a comparison-reconstruction calculation to be made on the basis described in the discourse analytic phase.

It would seem that the representation we are currently using, and which we will describe as a EU-BR representation, has a major defect in that it does not fully meet our first requirement and, furthermore, that the second requirement does not seem to demand the transformation of the 'combinatory' form of the graph (cf. p. 155 above) into the EU-BR form. On the other hand, as Dupraz (1972) notes, the representation currently entered as data in the discourse analytic phase reveals no trace of structural hierarchical relations between utterances; this may explain the appearance of a type of artifact characterized by a heterogeneity at the structural level of the utterances that make up the sequences of a domain. If this is in fact the case, there would appear to be two possible solutions. In the relatively short term, it seems possible to attempt to refine the EU-BR representation in such a way that the connectors can integrate parenthetical relations between utterances. In the longer term, our objective must be to transform the discourse analytic phase so as to allow the processing of graphs (in a combinatory form). This would in fact be a prelude to still more radical changes in that it would make it possible to compare structures containing the same *lexis*.



### 6.3. Conclusion

In her above-mentioned review article, Provost-Chauveau (1970: 136-137) rightly notes that one of the preconditions for the successful conclusion of our project must be "the possibility of effecting, in certain contexts, certain substitutions between terms *x* and *y* without changing the 'semantic interpretation' of the utterance".

We will take her comment as a starting point for a discussion of the theoretical problems currently raised by the 'interpretation of results' phase. As we have already had occasion to point out, this phase leads the analyst to reconstruct the various *types of substitution* to be found in sequences belonging to a single domain. In order to indicate the full import of the problem, on which rest both the validity of discourse analysis and the possibility of future developments, we will go back for a moment to the assumption we made when writing ADA69, in order to dispel, both for ourselves and for the reader, an illusion concerning the structure of semantic processes. The text centred on the idea that 'substitutions do not alter meaning' provided that a minimal identity of content can be ensured. We believed that this idea took the possibilities opened up by Harris to their logical conclusion in that it could provide a stricter definition of the 'equivalence classes' he introduces into the analysis procedure, and on whose subject he remains curiously vague:

The formal findings of this kind of analysis do *more* than state the distribution of classes, or the structure of intervals, or even the distribution of interval types. They can *also* reveal peculiarities within the structure, relative to the rest of the structure. They *can* show in what respects certain structures are similar or dissimilar to others. They can lead to *a great many* statements about the text. All this, however, is still distinct from an *interpretation*<sup>1</sup> of the findings, which must take the meanings of the morphemes into consideration, and ask what the author was about when he produced the text. Such interpretation is *obviously quite separate from the formal findings, although it might follow closely in the directions which the*

---

<sup>1</sup> [Emphasis by Harris; the other emphases have been added by Pêcheux]

*formal findings indicate.* (Harris 1952: 29)

We had assumed that, if the identity of the context was guaranteed, not only by the construction of the material (insofar as it forced the text to be repeated in the form of a corpus which was homogeneous in terms of its conditions of production), but also by the internal comparative workings of the elements of that corpus, we could overcome the hesitations which Harris expresses here, and which appear to go hand in hand with a surrender to a sort of intuitive semantics of the speaking subject and his intentions ('what the author was about ...'). We were and are highly critical of this. In referring to 'hesitations', we mean that *everything* combines to lead Harris to argue that the 'formal findings' he reaches are in fact the only findings that a non-subjective reading can supply, even though he does add the proviso that further (and equally 'formal') findings may be reached. We in fact still take the view that, given the choice between what Leeman calls a 'strong semantics', or the study of relations between utterances and extra-linguistic reality, and a 'weak semantics' characterized by the fact that "it refers to semantic equivalence between utterances without raising the question of what those utterances mean" (1973c: 85), the correct solution will be found by taking *the latter option*. If we further recall that, as Leeman also notes, paraphrase is a "basic concept of a weak semantics" (1973c: 85), we can conclude that Harris's 'hesitations' imply a reluctance to make a direct link between *paraphrase*, *substitutability* and *synonymy*. We believe that such a link must inevitably be posited if full use is to be made of Harris's insights, which are characterized primarily by a rejection of any extra-linguistic analysis of 'meaning'. In 1969 that link led us to the idea of a propositional invariant *underlying a paraphrastic family*; although our theoretical motives were very different, we arrived at a position similar to that of Gochet (1972). Our objective was, indeed, to arrive at 'semantic nexuses' which make up the common content of a set of propositions by using the analytic procedure we had outlined.<sup>1</sup> The result can also be described as a 'basic proposition'.

According to this view, the outcome of the analysis would be a

---

<sup>1</sup> The perspective we attempted to develop is in certain respects similar to Mel'cuk's views on the question of paraphrase and of the relationship between meaning and text (Zolkovskij and Mel'cuk 1971).

non-valued graph whose nodes consist of 'basic propositions'.<sup>1</sup> I think it necessary to explain this view in some detail if only because it is not currently possible to determine whether it is valid to some extent. At this point we encounter a problem in formal logic which might be formulated as follows: given a graph of a set of 'basic propositions' associated with a determinate discursive process, how to define the rules which make it possible

- (1) to construct a set of basic propositions, and only basic propositions, on the basis of the lexicon of predicates and arguments;
- (2) to construct a graph linking the basic propositions that have been so defined.

Such a system of rules would in fact construct what could rightly be termed the 'process of production' of the discourse corresponding to the corpus under analysis.

Two remarks seem to be required. The first is a simple observation whose generality cannot be guaranteed as such: as we have already noted, the 'semantic domains' obtained with the ADA procedure are not reducible to a family of utterances which can all be paraphrased by a single basic proposition. We therefore made a distinction between *two types of relations of substitution* (p. 165 above). Our second remark follows from the first, and attempts to explain it on the basis of the distinction between 'strong' and 'weak' semantics by showing that, because it took a 'weak semantics' as a starting point, our position in fact implied attempting to construct a 'strong semantics' on the basis of a 'weak semantics'. The idea that there is a correspondence between an invariant (the basic proposition) and a series of variations representing that invariant is *homologous* with the distinction between 'deep' and 'surface' structures. But these distinctions are themselves based upon a third and more general distinction, namely the distinction between a domain to which the values of truth apply ('objective information' or denotation) and a domain of

---

<sup>1</sup> In the following semantic domain, for example, we can construct the basic proposition R(A,B), in which R = give, guarantee, etc., A = X, the State, etc., and B = a minimum income, etc.

X	give	
	guarantee	
the State	guarantees	a minimum income

expression typified by the subjective character of 'messages' and 'connotations'. This distinction also allows us to endorse the rhetorical theory of 'deviation from the norm'; in a 'weak semantics' this becomes deviation from the invariant. Whereas we originally posited the view that metaphor is primary and constitutive rather than secondary and derivative, the logic of this perspective necessarily results in the relegation of metaphor to a category of 'surface phenomena' *accompanying* meaning, and in the assumption that meaning itself has already been constituted.<sup>1</sup> Ultimately, the question now arises as to whether it is still possible to discriminate between a 'weak semantics' and a 'strong semantics'. The system of oppositions we have just described in fact basically presupposes that meaning exists *as an object*, which would imply that the stability of the object (the real object or referent) is primary, and that *processes* are to be regarded as *objects that are set in motion*, and that are displaced across the representation that is given of them.

If, on the contrary, we accept the materialist thesis that 'objects' are not primary invariants, but points around which processes stabilize, the perspective changes considerably, particularly with regard to the principle that 'substitutions do not alter meanings'. It is apparent that neither this principle nor the corresponding principle applying to transformations ('transformations do not alter meanings') can be *posited as universal*. The materialist thesis would in fact presuppose the existence of a much broader field in which there is no *a priori* guarantee that substitutions and transformations do not alter meanings. This obviously means that we have to specify once more what we mean by 'paraphrase' by relating that concept to the concept of substitution and synonymy on the one hand, and the concept of transformation (defined in linguistic terms) on the other.

Leeman (1973a: 42) traces the evolution of Harris's concepts as follows:

---

<sup>1</sup> The question of metaphor and of the metaphoric effect is in our view decisive. In saying that metaphor is primary and not derivative, we do not intend to *invert* the relationship between literal meaning (kernel of meaning, denotation, basis of the logical proposition) and figurative meaning (periphery of meaning, manner of speaking, connotation and 'style') by suggesting that all meaning is figurative and peripheral, as the 'plural readings' perspective might imply. On the contrary, we wish to get rid of the duality between kernel and periphery and to argue that metaphor is the *movement* between two signifiers which gives them their meaning. The orientation of that movement effects an *asymmetry* which is the condition for the appearance of a meaning as either 'literal' or 'figurative'.

Initially, we have a non-ordered set of purely paraphrastic transformations defined in terms of co-occurrence, even though the term *paraphrase* itself is not used [... . At a later stage,] we get two types of operators, both with characteristics which can be described in grammatical terms: incremental operators and paraphrastic operators.

It is characteristic of paraphrastic transformations

that they do not in general determine any change in the meaning of their operand; they provide no supplementary information about their operand. (Leeman 1973b: 43)

It is characteristic of the second type of transformations that

they provide certain information and can therefore be interpreted as being predicative. (1973b: 51)

Although the formulation is not devoid of ambiguity, we think that this distinction between two types of transformation corresponds to our distinction between 'symmetrical substitution' and 'oriented substitution' (p. 165 above). We call Harris's formulation ambiguous because it is reminiscent of the assumptions of 'strong semantics', in particular of *the distinction between objective and subjective*, and of their linguistic implications in terms of the distinction between the language and speech. The latter distinction is, as we know, one between a *system* and *acts* which, at the same time, exist prior to it, constitute it and 'merge into it'. The pertinence of the parallel we are drawing may be gauged from the following passage:

The difference between the increment system and the paraphrastic T system is roughly that between the directly useable activities of life and the institutional apparatus which channelizes these activities. Like social institutions, the T system structures, facilitates, slants and petrifies the activities-for-use of the {I} system, and is inflexible, conventional, and in part historically accidental. (Harris 1970: 677)

It would appear, then, that no appreciable progress has been made, as this new distinction takes us back to the very presuppositions we wished to get away from. We propose, however, to show that this distinction in fact paves the way for a new conception which comes closer to meeting the theoretical requirement we have formulated, provided that we make a distinction between *three, and not two types of transformation* (or relations between pairs of sequences). The third type absorbs the other two, provided that the conditions we are about to define are met. We make the following distinction:

(1) Transformations involving constant lexical units. Current linguistic research relies to a great extent on the hypothesis of transformations of this type (for example: *The Romans resolved to destroy Carthage* → *The destruction of Carthage was resolved upon by the Romans*). These are what might be termed purely syntactic transformations, or substitution-transformations that, in theory, do not alter the meaning in that they involve the conversion of one set of morphemes into another. We will reserve judgement as to the basis of the hypothesis of constant meaning (which definitely relies upon a logical presupposition in that it necessarily implies the opposition between competence and performance) and as to its compatibility with the concept of enunciation outlined earlier.

(2) Substitution-transformations which 'change meanings' in that it is impossible to regard the substituents as equivalents. These are what we have described as 'oriented' substitutions. They involve, that is, a lexical change and a relationship of syntagmatization between the commutated elements. They correspond to Harris's 'incremental' transformations (1970: 616-617).

(3) Finally, and this is in our view the decisive point, we propose to introduce a third type of relation: non-oriented substitution involving lexical change. We refer to the *relation of synonymy* which, as Leeman (1973b: 49) stresses, is

a relation of equivalence between sentences which differs from the transformational relationship: the registration of semantic similitude is immediate and non-empirical; in other words, one has not yet found the linguistic means of deriving synonymy from a linguistic operation.

We would like to advance the hypothesis that the difficulty in arriving at a linguistic analysis of synonymy arises because synonymy is being thought

of in terms of the category of syntactic transformations as an attenuated equivalence, and not in terms of the oriented substitution, which appears to be antithetical to the very notion of synonymy. In our view, however, it seems productive to see *synonymy as an erasure of orientation*, and not as a lexical extension of syntactic equivalence.

If we now return to the concrete problems posed by the interpretation of the semantic domains produced by ADA, we can state that the most urgent problem is that of the criteria for the identification between commutatives. We know that the principle of their identification involves a search for constructions linking commutatives by means of syntagmatization which is, so to speak, perpendicular to the sequences that are commutated. These constructions may include temporal sequences of a narrative type as well as logical relations such as deducibility. In practice, however, the boundaries of the corpus pose an obstacle to their identification: there is no proof that the type of data we are looking for is *discursively homogeneous* with the zone in which the commutations are effected. In fact, all research suggests that the reverse is true. This therefore takes us back to the theoretical problem of the relationship between a discourse and its 'interdiscourse', in other words all the other processes which intervene within it in order to constitute it (by supplying it with 'preconstructs'; cf. Henry 1974 and Pêcheux 1982a) and orient it (by acting as a *transverse* discourse or, as we have just said, a *perpendicular discourse*).

This question relates directly to the problematic of the two 'forgettings' which we outlined at the beginning of this chapter (pp. 138-140). Forgetting 2, which is analogically related to the preconscious-conscious system, and which concerns the point of articulation between linguistics and discourse theory, corresponds primarily to the mechanisms of 'syntactic' paraphrase and to 'natural' linguistic synonymies or, in other words, to synonymies in which orientation has been erased. The entire system of self-paraphrase by means of which any discourse can clarify itself by divorcing what can be said from what is consciously rejected (and what is rejected is to a large extent the presence of the other – with a small o – within the speaker's discourse) relates to substitution-transformations of types (1) and (3).

As for transformations of type (2), they too might relate in part to forgetting 2 to the precise extent that the speaker is capable of evoking the discursive processes which allow substitutions to be oriented. What we have termed forgetting 1 is, then, definitely characterized by the fact that *the speaker-subject can have no access to the processes that constitute the*

*transverse discourses and preconstructs of its own discourse.* He has no access, that is, to anything implied by the expression 'the discourse of the Other (with a capital O).<sup>1</sup> We can therefore readily advance the hypothesis that the semantic domains which can currently be identified by ADA are not homogeneous with respect to the distinction between the two forgettings. The question of whether or not this disparity is reproduced by a further heterogeneity resulting from the impossibility to disassociate the combined processes remains open.

Given the present nature of our outputs, results derived from different corpora should be compared in different ways, depending on whether the difference between those corpora pertain primarily to the zone of forgetting 1 or that of forgetting 2.

Finally, the question of the criteria that allow us to identify both the autonomy of a process and the limits of its autonomy also remains unresolved: until this question is resolved, there will be uncertainties as to the possible relations between oriented substitutions and combinations of

---

<sup>1</sup> The 'acts of the speaking subject' in a given 'situation' and in the presence of given 'interlocutors' — in other words, the subjective illusion which certain theories of enunciation take at face value — are in reality an *effect of relations between discursive processes*. The fact that one specific (phonic or graphic) sequence — rather than another — is constantly being 'filtered' or 'selected' is not the result of an *act of choice on the part of the speaker*. It reflects the intervention within a given discursive formation (with paraphrastic mechanisms of its own) of other formations which make relations between the elements in play in its mechanisms *uneven*, and which therefore *order* them in such a way that one of them always has the 'privilege' of seeming to be the 'right' word or expression. In what is conventionally known as the domain of literature, this 'privilege' takes the form of the *obvious* impossibility of paraphrasing a 'brilliant text' ('There is no other way of putting it'). This point — and we can do no more than outline it here — seems to enable us to overturn the problematic which sees 'literal meaning' as a natural link between 'language' and 'thought'. It is also therefore a challenge to literary theories which see style as a deviation from a norm. Such theories usually refer to the *unique* character of the linguistic sequence (or to the irreplaceable nature of its words, expressions or turns of phrase), and it is often believed that its *sustained deviation* from the norm (in the sense in which we can speak of a sustained note) results from a more or less 'brilliant' will considered unique. In the perspective outlined here, its 'unicity' is the overdetermined product of contradictory and uneven relations between discursive formations: the phonological and morpho-syntactic materiality of the sequence (the signifier) is therefore determined as being unique amongst the paraphrastic multiplicities which support its 'meaning', just as the existence of a play on words imposes, by its very materiality, one formulation (rather than another logically equivalent paraphrase), and thus maintains a compromise between two discursive formations and realizes the phenomenon we call overdetermination.



processes. We have not in fact abandoned the idea that orientation is to be regarded as the effect of an articulation between processes which involve relations of erasure, subordination and subjection. If that hypothesis is correct, the autonomy of a process is certainly marked by the existence of inter-paraphrastic families in which all orientation has been erased, and the nature of our current findings results from the fact that we have not yet succeeded in *isolating* an individual process. If that is so, it may be possible to improve our findings by increasing, perhaps to a considerable degree, the dimensions of the corpus (which may not yet have reached a state of critical mass) and by raising the threshold value which establishes the minimal proximity between the sequences under comparison.

To conclude, we will go back to the problem of relations between semantics and syntax, which is in fact basic to the debate whose major axes have been outlined above. It seems that we can at the moment formulate three hypotheses relating to this question.

(1) *The field of linguistics is confined to the domain of those facts of syntax* whose mechanisms can be calculated, and it is an extension of phonology and morphology; semantics, on the other hand, belongs to the subjective domain of meaning. Originally, this hypothesis – which is inscribed within the structuralist tradition attempting to describe systems of forms – was based upon the postulate that syntax is independent of semantics. This solution has been adopted both by the behaviourists and by the theorists of distributionalism and formalism. It is also the viewpoint initially adopted by Chomsky in *Syntactic Structures* (1957).

(2) *The whole of semantics belongs within the field of linguistics.* Paradoxically, this second hypothesis derives from the first. The study of semantics is seen as a natural extension of the facts of syntax and as aiming to explain them (once distributional methods have defined their limits). Chomsky, among others, adopts this position in *Aspects of the Theory of Syntax* (1965), where the semantic component is used to interpret syntax. It is to be noted that the integration of semantics into this model of language is justified by the assumption, which remains largely unspoken, that meaning is a fact of language. From a methodological point of view, the procedure used for the semantic analysis of units is comparable to that used by phonology (the decomposition of features, cf. Katz and Fodor 1964). This solution relies upon a theory which is at once a 'theory of knowledge' and a psychology of the human subject; the construction of the world is related to a neutral, ideal subject.

Despite the polemical tone of the debate between Chomsky and the supporters of generative semantics, the latter current is also related to hypothesis (2). Ultimately, the whole field of semantics is a topic for linguistic studies, even though its supporters disagree with Chomsky in that they do not divorce semantics from syntax (cf. their underlying 'logico-semantic-syntactic' structures), and even though certain semantic phenomena (such as 'presuppositions') are dealt with within the framework of a 'theory of worlds' which is intended to destroy the unicity of the subject. And yet their 'subjects' are still neutral subjects and sources of meaning, and they are not referred to their objective determinations. This is why generative semantics can be regarded as a further step along the road that leads to the confusion of ideology, discourse and language.

(3) *Only certain semantic facts come within the field of linguistic study.* From this third hypothesis can originate two mutually exclusive solutions. (It may lead to confusion in that both solutions refer to 'enunciation'.)

(3a) Benveniste's (1966) solution is to make a distinction between meaning and reference (or designation) within the process of signification itself. The most immediate interpretation of this distinction consists of seeing it as a distinction between a linguistic semantics and a non-linguistic semantics. This distinction appears to apply to the study of the language, which is seen as a structured and hierarchical system of signs: "The *meaning* of a unit is the fact that it has a meaning. [...] This is equivalent to identifying it in terms of its capacity to fill a 'propositional function'" (p. 127). This is a matter for linguistic analysis. The sign's *reference*, on the other hand, relates to "the world of general or particular objects acquired from experience or forged by a linguistic community" (p. 128). For, for Benveniste, the language is not simply a system of signs; it is also "an instrument of communication, and discourse is its expression" (p. 130). The system of signs and discourse constitute "two different universes, although they embrace the same reality, and they give rise to two different forms of linguistics" (p. 130). The articulation between these two linguistics is effected at the level of the sentence, and at that point, the study of reference is re-introduced into the field of linguistics; the *meaning* of the sentence (its function of predicate) is in fact described by analyzing the relationship between the signs that make it up, whilst the *reference* of the sentence (i.e. 'concrete and empirical situations', p. 128) is a matter for discursive linguistics (theory of enunciation). This second line of argument suggests the idea that discourse-speech is the locus of individual creativity;

for Benveniste, the modalities of the sentence express three 'inter-human' characteristics of discourse, each corresponding to an attitude on the part of the speaker, namely "transmitting an element of knowledge, requesting information, and giving an order" (p. 130). In our view, the characteristic feature of this solution is its inability to escape the ideological duality of system and creativity (a system of signs as opposed to individual creativity). Here, 'discourse' is no more than a new avatar of 'speech'.

(3b) According to another solution, the boundary between the linguistic and the non-linguistic lies within semantic phenomena; it lies on the borderline between formal semantics and discursive semantics. We once again encounter the notion of 'enunciation', but it is now defined as the theory of the subjective illusion of speech (the theory of the 'verbal body'), and not simply as a reproduction of that illusion. Everything that has been said here indicates that we have constantly been trying to work within the framework of this hypothesis, but that obviously does not mean that every aspect of ADA is immune to criticisms from this very point of view.

Whilst we can claim that phenomena such as nominalization, preconstructs and so on have already been identified in our findings,<sup>1</sup> it is

---

<sup>1</sup> We present here some examples of the linguistic phenomena which can be identified in the findings currently obtained with ADA (taken from Pêcheux et al. 1979).

*preconstructs*

minimum income  
cultural development  
distribution of goods  
development of people

*modalities*

It should ... / the action of the State must ...

*instantiation*

X	give guarantee	a minimum income
the State	guarantees	
decline of the well-being of	each person l'individu	

still true to say that all the criticisms we have repeated or formulated with respect to our current modes of morpho-syntactic analysis are valid and call for far-reaching changes. We would therefore like to conclude by raising the question of the relationship between the linguistic and the discourse analytic phase. It is not inconceivable that, rather than being juxtaposed sequentially, these phases could determine one another in such a way that the discourse analytic phase has a retroactive effect on the linguistic one. Nothing prevents us from elaborating a mode of reading which applies to several levels, which, that is, begins with a minimal morpho-syntactic system. The data obtained from this first reading – and its effects – could then be gradually integrated into the linguistic phase.

It seems possible that intermediary findings as to relations of synonymy, syntactic paraphrase and syntagmatization between commutatives could be re-invested in a morpho-syntactic analysis which takes place at a higher level, and which deals specifically with the determination of the inter-sentence phenomena that are bound up with the identification of textual zones affected by the possibilities of reformulation and paraphrase offered by specific syntactic complexes on the one hand and relations of syntagmatization on the other hand. We hope our long-term research into the articulation between linguistics and discourse theory will lead in this direction.

---

*determiners*

a / the stabilization is expected  
States / the States

*nominalizations*

humanity encounters danger of	shortage of	raw materials goods
we have shortages scarcity restriction limitation ▼ distribution problem ▼	of consumer goods	

### **PART III**

## **EXPERIENCES AND PERSPECTIVES**



## Theory, procedures, findings and perspectives

Michel Pêcheux, Jacqueline Léon,  
Simone Bonnafous and Jean-Marie Marandin

### 7.1. Theoretical presentation of ADA69

The theoretical references (positive and negative) which, from 1969 onwards, presided over the construction of the ADA apparatus were inscribed within the space of the philosophical structuralism of the 1960s, and centred on the question of ideology. To be more specific, they centred on the question of reading ideological discourses. The structuralist problematic, which was beginning to condense around the names of Lévi-Strauss, Foucault, Barthes and Althusser, was as much a polemical device to combat the dominant ideas of the period as a programme of work. By the 'dominant ideas of the period' we mean:

- 'relics' of a philosophical spiritualism and of religious concepts of reading, ranging from literary hermeneutics and the search for the 'themes' of 'works' to the phenomenological conception of a 'project' as the projection of meaning onto verbal material by the constitutive power of the reader-subject, and the idea that the meaning of a text is a correlate to a reading consciousness dwelling in an unbounded interpretative subjectivity (all these 'relics' were alive and well, and still cling stubbornly to life);
- secularized and more mundane forms of this spontaneous practice of reading in the shape of the multifarious forms of 'content analysis' which were then invading the human sciences;
- a quantitative objectivism which reacted against this impressionistic spiritualism by referring to the seriousness of the sciences, and to information theory in particular, and by elaborating projects in which texts were treated as though they were populations of words amenable to a sort of statistical textual demography (some of these projects were realized in

lexicometric studies).

The philosophical structuralism of the 1960s set of to battle against these various (spontaneous and 'scientific') forms of the empirical obviousness of reading, emblazoning its banners with concepts of 'symptomatic reading' and 'discourse theory' and with slogans like 'Identify the efficacy of a structure on its effects, through its effects' (cf. Althusser 1970). Marx, Nietzsche, Freud and Saussure were all recruited to fight in the same battle, which centred on the question of what is meant by speaking, listening and reading. As Althusser (1970: 16) put it:

Only since Freud have we begun to suspect what listening, and hence what speaking (and keeping silent) *means* [*veut dire*]; that this '*meaning*' [*vouloir dire*] of speaking and listening reveals beneath the innocence of speech and hearing the culpable depth of a second, *quite different* discourse, the discourse of the unconscious.

And so, we openly asked linguistic structuralism for strategic support. If we had to analyse the 'unconscious discourse' of ideologies, structural linguistics, which was the 'modern science' of the day, provided the scientific means to displace the question from the domain of the quantitative to that of the qualitative, from a statistical description to a quasi-algebraic theory of structures, and at the same time a means to reject the 'anything goes' attitude of 'literary readings'.

If ideological discourses were indeed the true myths of our societies, and if they were comparable to the myths studied by Propp and Lévi-Strauss, then it must be possible to elaborate effective procedures capable of reconstructing a trace of the invariant structure (the system of their 'functions') that lay beneath the combinatory series of empirical surface variations and, therefore, to reconstruct something of a 'structure' present in the series of its effects.

The development of ADA69 was one of a number of attempts to implement this programme, and to take seriously 'modern linguistics', and especially the work of an American linguist who had written a text with the providential title 'Discourse analysis' (Harris 1952). For a long time, this was a concrete scientific reference for linguists who had followed the example set by Jean Dubois by beginning to work in the field of discourse analysis.

From this point of view, the specificity of the 1969 version of



ADA, as opposed to other work on discourse analysis, lay in its attempt to take Harris's linguistics to its logical conclusion on the basis of the theoretical position described above.

If the meaning of a textual surface lies in the play of the relations (of equivalence, commutation, paraphrase ...) which are necessarily established between it and other specific textual surfaces, it follows that the study of discursive processes (inherent in the underlying structure that is to be studied) presupposes a reference to sets of surfaces (or discursive corpora). The apparatus will organize the set into a state of potential self-paraphrase, and can then investigate its structure by extending to corpora (which can be identified by their 'socio-historical conditions of production') the processes Harris applies to sequences marked by internal repetition and stereotypes, as in the famous example 'Millions can't be wrong' (Harris 1952).

The order and arrangement of the procedures used by ADA69 were established on this basis. ADA69 necessarily comprized three phases:

- the socio-historical construction of the corpus to be analyzed;
- a 'Harrisian' phase involving the syntactic delinearization of the textual surfaces of the corpus in order to isolate elementary utterances and linguistic relations between those utterances;
- a phase in which the data produced by the syntactic analysis is processed automatically. It is this phase which justifies ADA's claims to being 'automatic'. The objectivity of a process which functions by itself was deliberately sought in order to escape the 'subjective obviousness' of reading in the hope of bringing to light traces of the famous 'underlying structure' of the textual corpus under examination.

The application of effective algorithmic procedures was therefore an essential precondition for this undertaking. That is still the case today, despite the radical restructuring of various aspects of ADA on which we are now working.

## 7.2. Presentation of the procedure of automatic discourse analysis

This presentation is of purely historical interest. It is a sort of balance sheet — or even a final account — of the use that has been made of the discourse analysis method developed by Michel Pêcheux in 1969. The present text is a résumé, or almost a banalization, of the method in that it

does not deal with the above-mentioned theoretical aspects which presided over the method's elaboration. It may serve as a reference point for any reader who is interested in the new algorithms for discourse analysis, or in the new prospects they open up, which are described at the end of this chapter (see pp. 209-213).

In the new project, the structuration of the data described below is transformed completely, as the old structuration proved in retrospect to be unacceptable from a linguistic point of view. However, the notion of a paradigmatic algorithm for the establishment of paraphrases or for the identification of the associated discursive paths is not invalid. If interpreted differently, it may constitute one aspect of the new algorithm.

## The data

### *Corpus*

In order to illustrate the ADA69 algorithm, we have taken part of the corpus constructed by Bonnafous on the basis of motions put to the *Parti Socialiste's* 1979 Congress in Metz. This corpus is of the 'archival'<sup>1</sup> type, and is constructed on the basis of three congress motions: motion A (Mitterrand), motion C (Rocard) and motion E (CERES). The motions were selected on the basis of Bonnafous's politico-historical hypotheses. As each motion dealt with two themes (the union of the left and the economy), it was possible to construct six basic corpora:

	Economy	Union of the left
Motion A (Mitterrand)	ACE	ACU
Motion C (Rocard)	CCE	CCU
Motion E (CERES)	ECE	ECU

In the following we will present the results of our treatment of ACU (i.e.

---

<sup>1</sup> The corpora processed by ADA69 are of two types: the experimental and the archival (see Chapter 6, pp. 147-148).

the corpus that consists of the part of motion A that concerns the union of the left).

### *Autonomous discursive sequences*

The first manual stage of the procedure consists of dividing the corpus into autonomous discursive sequences, which are the maximal units of comparison. The idea of segmenting the corpus into autonomous discursive sequences is consistent with ADA's principal hypotheses as to the production process of a meaning-effect. It is in fact by relating two sequences – by studying their capacity for commutation, substitution and equivalence – that we can reveal discursive processes. Autonomous discursive sequences are, however, the product of a selection process which wrenches fragments of a text out of their context, and which makes it impossible to process texts in their sequentiality.

The procedure of segmenting a corpus into autonomous discursive sequences is based on a careful linguistic analysis of inter-sentence links (anaphora<sup>1</sup>, ellipsis, connectors) and questions of mode, aspect, tense and determination. Autonomous discursive sequences, which consist of a minimum of one sentence, are defined by their thematic unity.

- Take a sentence i. If the next sentence j begins with a connector such as 'but', there is no segmentation; there is a thematic continuity between i and j, and sentences i and j belong to the same autonomous discursive sequence.

- If sentence j contains an anaphora whose referent is present in sentence i (a simple anaphora such as in 'John ... He'), sentences i and j belong to the same autonomous discursive sequence to the extent that the anaphora establishes a thematic unity between them.

These two conditions are therefore the primary criteria for non-segmentation.

### *Syntactic analysis*

The syntactic analysis of autonomous discursive sequences corresponds to

---

<sup>1</sup> Not all types of anaphora are examined here; in certain cases, they can pose complex segmentation problems.

a manual delinearization of the surface of the text. Each autonomous discursive sequence is analyzed by means of a graph whose *points* represent elementary utterances (propositions as defined by traditional grammar) and whose *arrows* represent connections between those elementary utterances (cf., e.g., p. 153). The graph is represented by a list of binary relations linking elementary utterances by means of a connector. The ACU corpus is divided up into 25 autonomous discursive sequences of which, for the purpose of this exposition, we will take ACU5 and ACU9 as examples of the syntactic analysis.

Elementary utterances comprise eight morpho-syntactic positions:

F	form of the utterance, containing indications as to grammatical voice, status, mode, modality, tense and emphasis
DET <sub>1</sub>	determiner of N <sub>1</sub>
N <sub>1</sub>	generally a noun, or meta-term S in a subject position
V	verb, or meta-term E indicating the presence of a complex noun phrase
ADV	adjectival, verbal or phrasal adverb
P	preposition governed by a verb or introducing an adverbial adjunct
DET <sub>2</sub>	determiner of N <sub>2</sub>
N <sub>2</sub>	noun in object position, adjective, or meta-term S reflecting an objective clause or free adjunct

First we take autonomous discursive sequence ACU5:

*Le Parti communiste n'a participé (avec de Gaulle, Gouin, Bidault et Ramadier) qu'à des gouvernements d'union nationale de concentration républicaine*

[The Communist Party has only participated (with De Gaulle, Gouin, Bidault and Ramadier) in predominantly republican governments of national unity]

Syntactic analysis results in the following list of utterances:

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
ACU054041	0003	L	PC	PARTICIPER	SEUL	A	DS	GOVERNEMENT
ACU054042	0000	R	GOVERNEMENT	E	0	DE	0	UNION
ACU054043	0000	R	UNION	E	0		0	NATIONAL
ACU054044	0000	R	GOVERNEMENT	E	0	DE	0	CONCENTRATION
ACU054045	0000	R	CONCENTRATION	E	0		0	REPUBLICAIN
ACU054046	0003	R	PC	PARTICIPER	SEUL	AVEC	0	DE GAULLE
ACU054047	0003	R	PC	PARTICIPER	SEUL	AVEC	0	GOUIN
ACU054048	0003	R	PC	PARTICIPER	SEUL	AVEC	0	BIDAULT
ACU054049	0003	R	PC	PARTICIPER	SEUL	AVEC	0	RAMADIER

The list of binary relations is the following:

4041 92 4042  
4042 92 4043  
4041 92 4044  
4042 40 4044  
4044 92 4045  
4041 09 4046  
4041 09 4047  
4041 09 4048  
4041 09 4049  
4046 40 4047  
4047 40 4048  
4048 40 4049

Let us now take autonomous discursive sequence ACU9:

*Le point qui nous importe aujourd'hui est de savoir s'il est imaginable que le PC change d'attitude, cesse bientôt de considérer les socialistes comme ses adversaires principaux, et de préférer le gouvernement de la droite et du grand capital à la victoire des travailleurs. Rien ne le montre.*

[The issue that concerns us today is whether or not it is possible to imagine the CP changing its attitude of considering the socialists to be its main enemies and of preferring the government of the right and of big capital to the victory of the workers. There is nothing to show that this is possible.]

Here the list of utterances is as follows:

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
ACU094117	0000	0	S	IMPORTER	AUJ	A	0	NOUS
ACU094118	0000	0	NOUS	SAVOIR	0	•	0	S
ACU094119	0000	0	S	ETRE	0	0	0	IMAGINABLE
ACU094120	0020	L	PC	CHANGER	0	DE	0	ATTITUDE
ACU094121	0020	R	PC	CESSER	OTOT	0	0	0
ACU094122	0000	R	PC	CONSIDERER	0	•	LS	SOCIALISTE
ACU094123	0000	R	PC	CONSIDERER	0	COMM	SES	ADVERSAIRE
ACU094124	0000	R	ADVERSAIRE	E	0	0	0	PRINCIPAL
ACU094125	0000	R	PC	PREFERER	0	•	L	GOVERNEMENT
ACU094126	0000	R	GOVERNEMENT	E	0	DE	L	DROITE
ACU094127	0000	R	GOVERNEMENT	E	0	DE	L	CAPITAL
ACU094128	0000	R	CAPITAL	E	0	0	0	GRAND
ACU094129	0000	R	PC	PREFERER	0	A	L	VICTOIRE
ACU094130	0000	R	VICTOIRE	E	0	DE	LS	TRAVAILLEUR
ACU094131	0100	0	RIEN	MONTRER	0	•	0	CELA

And the list of binary relations:

4117 05 4118  
4118 13 4119  
4119 05 4120  
4119 05 4121  
4120 40 4121  
4121 06 4122  
4122 10 4123  
4123 92 4124  
4121 06 4125  
4122 40 4125  
4125 92 4126  
4125 92 4127  
4126 40 4127  
4127 92 4128  
4125 08 4129  
4129 92 4130  
4117 60 4131

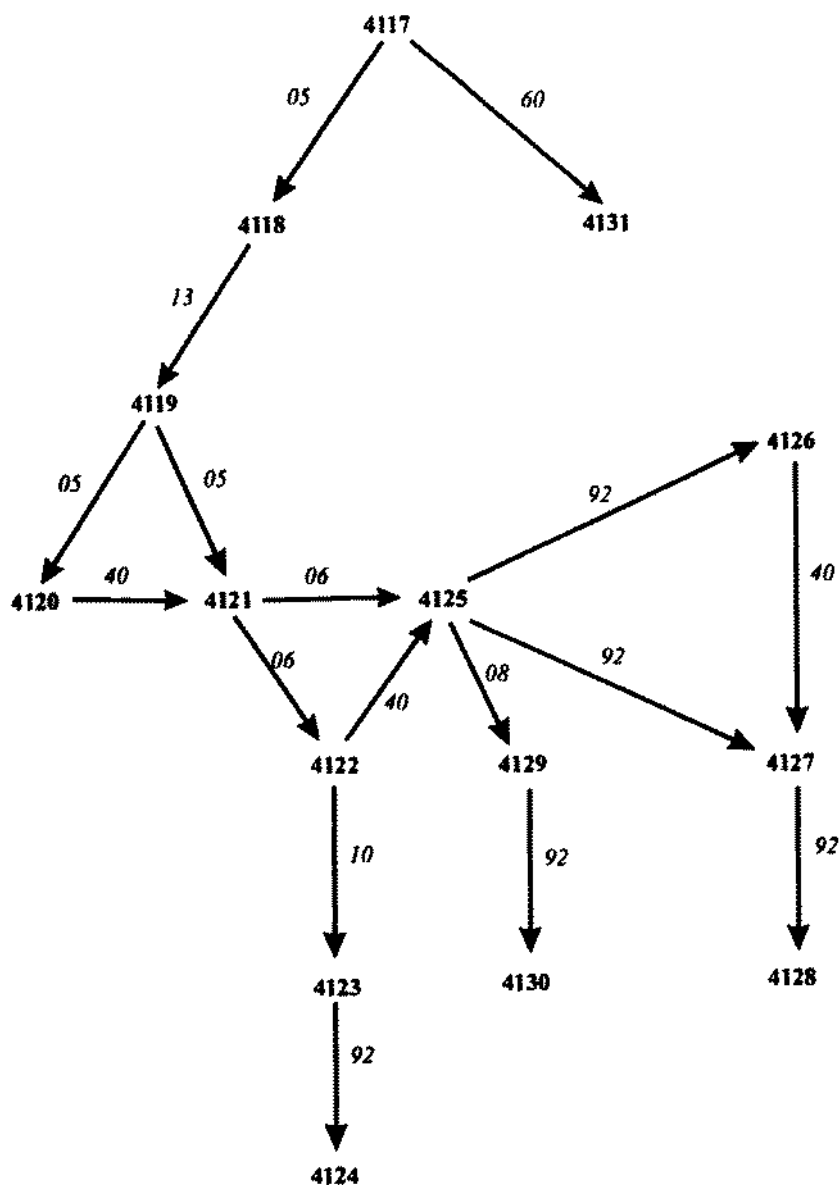
This list is represented by the graph on page 197.

### **The algorithm**

The FORTRAN IV programme developed by Philippe Duval and Michel Pêcheux in 1971-72 made it possible to realize the algorithm described below.

#### *Formation of quadruplets*

The procedure is as follows: the programme compares all binary relations in an autonomous discursive sequence with all binary relations in all other autonomous discursive sequences in the corpus. Binary relations in the autonomous discursive sequence itself are excluded because an autonomous discursive sequence may display internal repetitions which might obscure the paraphrase-effects identifiable in the corpus.



Graph of Discursive Sequence ACU9

We will illustrate the paraphrase procedure by using binary relation 4041 92 4042 (from ACU5) and binary relation 4125 92 4126 (from ACU9). The two binary relations are compared category by category. A Boolean vector is then constructed: 1 in case of lexical identity in the category which is being explored, and 0 in case of lexical difference. The vector is then multiplied by a 'pattern' which is introduced into the data and which gives each morpho-syntactic category an empirically-determined<sup>1</sup> weight. The sum of their weights is compared with a threshold value, which is also introduced into the data and established empirically<sup>2</sup>. If the sum is equal to or greater than the threshold value, the pair of binary relations is retained, and it constitutes what we call a quadruplet:

	F	DET, N <sub>i</sub>	V	ADV P	DET <sub>j</sub>	N <sub>j</sub>	CO
ACU54041	0003	L PC	PARTICIPER	SEUL A	DS	GOUVERNEMENT	92
ACU94125	0000	R PC	PREFERER	0 *	L	GOUVERNEMENT	92
vector	0	0	1	0 0	0	1	1
weight	5	0	6	0 5	0	6	6
result	0	0	6	0 0	0	6	6

	F	DET, N <sub>i</sub>	V	ADV P	DET <sub>j</sub>	N <sub>j</sub>	
ACU54042	0000	R GOUVERNEMENT	E	0 DE	0	UNION	
ACU94126	0000	R GOUVERNEMENT	E	0 DE	L	DROITE	
vector	1	1	1	1 1	0	0	
weight	5	0	6	0 5	0	6	
result	5	0	6	0 5	0	0	

$\Sigma$  vector x weights = 37

The results are obtained by multiplying the vectors by their weights. The sum of these results is 37. As the threshold value entered as a parameter is 36, the quadruplet formed by the two binary relations is retained.<sup>3</sup>

The comparison of binary relations in ACU5 and ACU9 gives seven quadruplets:

<sup>1</sup> The term 'empirical', as applied to the weighting system, refers to both an intuitive hierarchy (which is obviously highly debatable) established between syntactic categories, and to the fact that that hierarchy produces optimal results, at least in quantitative terms.

<sup>2</sup> The term 'empirical', as applied to the threshold value, indicates that the decision to accept certain values of variables is based on a comparison of results obtained in different conditions, and not on any procedure of statistical validation.

<sup>3</sup> The maximum value of the sum of weights, obtained for two identical binary relations, is 62, on condition that both V-positions are occupied by a verb (weight = 6), not by an E (weight = 3).



4041 92 4042  
4125 92 4126

4041 92 4042  
4125 92 4127

4041 92 4044  
4125 92 4126

4041 92 4044  
4125 92 4127

4042 92 4043  
4127 92 4128

4042 40 4044  
4126 40 4127

4044 92 4045  
4127 92 4128

The comparison of binary relations in all the autonomous discursive sequences in the corpus gives twenty-five quadruplets, which constitute the first quantitatively exhaustive stage of the processing of findings for the ACU corpus. These results are now transitively organized on two axes, vertically in chains and horizontally in domains.

### Formation of chains

If the numbers of the right-hand utterances in a quadruplet are identical to the numbers of the left-hand utterances in another quadruplet, they form a chain by transitivity. The seven quadruplets derived from the comparison of ACU5 and ACU9 give the first four chains by horizontal transitivity; quadruplet 3 remains isolated. Chains 1 and 4 are then grouped together to give chain 5. The stage of forming chains is complete when all recursive groupings have been effected:

```

quadruplets 1 and 6 → chain 1      4041 92 4042 40 4044
                                     4125 92 4126 40 4127

```

```

quadruplets 2 and 5 -> chain 2      4041 92 4042 92 4043
                                       4125 92 4127 92 4128

```

quadruplets 4 and 7 → chain 3      4041 92 4044 92 4045  
4125 92 4127 92 4128

quadruplets 6 and 7 → chain 4	4042 40 4044 92 4045 4126 40 4127 92 4128
chains 1 and 4 → chain 5	4041 92 4042 40 4044 92 4045 4125 92 4126 40 4127 92 4128
quadruplet 3 → quadruplet 3	4041 92 4044 4125 92 4127

*Formation of domains*

If two chains have a sequence in common, if, that is, the sequence of the numbers of the utterances and higher connectors in one chain is identical to their sequence in another chain, we form a chain by vertical transitivity. Only chains of similar length are grouped together. If domains are formed on the basis of ACU5 and ACU9, we obtain the following domain:

chains 2 and 3 → domain 1	4041 92 4042 92 4043 4125 92 4127 92 4128 4041 92 4044 92 4045
---------------------------	--

Chain 5 (containing chains 1 and 4) and quadruplet 3 remain isolated. They are accepted as domains in their own right.

Paraphrastic relations between sequences are therefore realized at both the vertical and the horizontal level. The domains contain equivalence classes which have to be interpreted on the basis of the invariants which allowed the sequences to be grouped together.

The whole of the ACU corpus gives 13 domains (see p. 201).

*Relations between domains*

The list of domains having been established, relations between domains now provide new findings. These relations are of two kinds: paradigmatic relations result in the construction of hyperdomains, whilst syntagmatic relations establish relations of dependency between domains and hyperdomains respectively.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	149
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

*Constitution of hyperdomains*

Hyperdomains are formed on the basis of the numbers of left-hand utterances in domains, and on the basis of three types of relation: inclusion (*ICL*), intersection (*INT*), and identity of origin (*IRG*).

- *ICL*: all numbers of left-hand utterances in domain  $D_i$  also belong to  $D_j$ . For example:  $D_1$  *ICL*  $D_2$ ; domain 1 is included in domain 2.  $D_1$  and  $D_2$  form a hyperdomain.
- *INT*: domain  $D_i$  and domain  $D_j$  have one or more left-hand utterances in common. For example:  $D_{12}$  *INT*  $D_1$ .  $D_{12}$  and  $D_1$  form a hyperdomain.
- *IRG*: the numbers of all left-hand utterances are common to domains  $D_i$  and  $D_j$ . For example:  $D_4$  *IRG*  $D_1$ .  $D_1$  and  $D_4$  form a hyperdomain.

We can, by transitivity, form a hyperdomain taking in domains  $D_1$ ,  $D_2$ ,  $D_4$  and  $D_{12}$ :  $HD_1 = (D_1, D_2, D_4, D_{12})$ . Hyperdomains constitute equivalence classes of dimensions greater than domains, and are formed on the basis of paraphrastic relations established on the paradigmatic axis.

*Relations of dependency between domains*

The relations between domains are relations of dependency if those domains contain utterances belonging to a single path in the same autonomous discursive sequence. For example, a left-hand utterance in  $D_{10}$  is utterance 4122, and a left-hand utterance in  $D_4$  is 4125. These utterances belong to the same path  $4117 \rightarrow 4118 \rightarrow 4119 \rightarrow 4120 \rightarrow 4121 \rightarrow 4122 \rightarrow 4125 \rightarrow 4126 \rightarrow 4127 \rightarrow 4128$  (see the graph on p. 197). Utterance 4122 dominates utterance 4125 because it has a higher position in the graph of ACU9. Therefore,  $D_{10} \rightarrow D_4$ .  $D_{10}$  dominates  $D_4$ .

A higher position of an utterance in the graph of an autonomous discursive sequence corresponds to the relative position of the utterance in the sentence or the autonomous discursive sequence (e.g. principal vs. subordinate clause, determined vs. determining, succession of two principal clauses in two different sentences).

It might be said that these relations of dependency are similar to a 'micro-argument' inside an autonomous discursive sequence.

Relations of dependency thus group domains along a syntagmatic axis (belonging to a single path). But they also introduce a paradigmatic relation between the other parts of the domains which do not necessarily belong to the same path, such as binary relations 4041 92 4044 (of  $D_4$ ) and 4094 40 4095 (of  $D_{10}$ ). A question thus arises as to whether we can

interpret the entirety of the relations between the domains of a corpus as the 'underlying argument' of that corpus.

All these relations of dependency can be represented by a graph which will be discussed below when we analyse the findings for the ACU corpus as compared with the findings for the CCU corpus.

### **Findings of the ADA procedure for two sub-corpora**

For anyone carrying out research on discourse analysis, the real work begins when the entire process described above (segmentation into autonomous discursive sequences, manual syntactic analysis and automatic discourse analysis) has been completed. The input for this work consists of the list of domains (and hyperdomains) corresponding to each corpus, and of the relations of dependency which link the various domains.

#### *Two possible lines of research*

- We could have chosen to make a detailed study of the organization of each domain in order to make a distinction between contextual synonymy, implication and contradiction (cf. Pêcheux et al. 1979). That would have allowed us to classify domains on the basis of themes, to study the emergence of meaning within each corpus, and to specify relations of identification, divergence and contradiction between corpora. That was not, however, our primary goal.

- Rather than looking at the domains in detail, we concentrated on the relations of dependency that are established between domains. This made it possible to construct a discursive trajectory for each corpus. In our view, domains are simply a means of gaining access to discursive trajectories.

If trajectories are to be established in a readable fashion, the content of each domain must be presented in a single table. We have therefore attempted to represent each domain or hyperdomain by means of a basic proposition which accurately summarizes it. The role played by basic propositions in this study therefore differs from the role they play in other ADA studies. Insofar as those studies are concerned primarily with relations internal to each domain and only secondarily with the general structure of discursive structures, the content of the domain is rarely

mentioned during the latter phase, and it is mentioned only for the record. Here, we take the view that the basic propositions shown in the graphs representing discursive trajectories make it possible to reconstruct what might be termed a second text. Our interpretation is based upon a reading of that second text.

The principles governing the elaboration of our basic propositions are simple: coordination or juxtaposition of elements which have the same position in the domain; elimination of redundancies; parenthetical adjunction of contextual elements absolutely essential to the intelligibility of the proposition.

Take for example hyperdomain HD1 in ACU. It is made up of domains D1, D2, D4 and D12, which we use to form the following basic proposition:

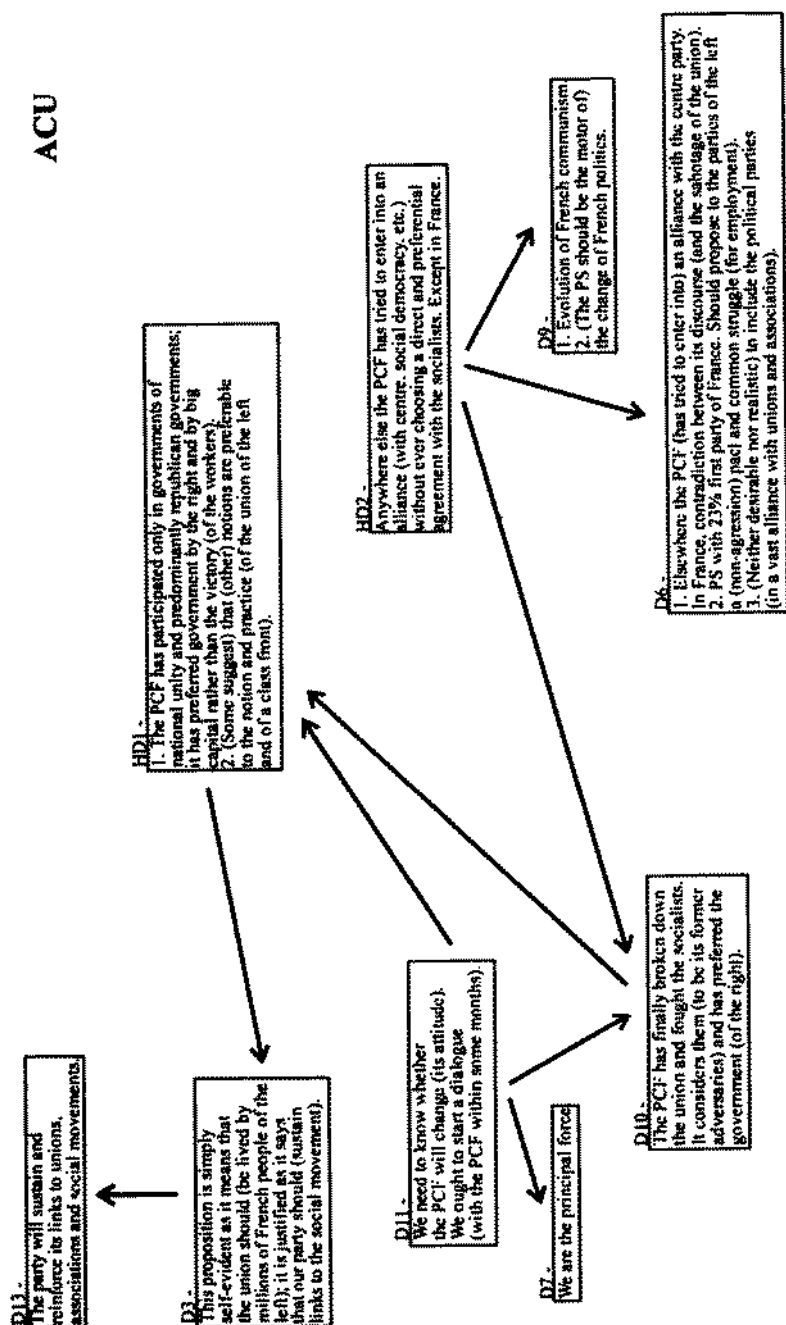
- HD1 = 1. *The PCF has participated only in governments of national unity and predominantly republican governments; it has preferred government by the right and by big capital rather than the victory (of the workers).*  
 2. *(Some suggest) that (other) notions are preferable to the notion and practice (of the union of the left and of a class front).*

Every domain or hyperdomain in ACU is summarised in the same way, and their contents are linked by arrows; the latter symbolize relations of dependency. We thus obtain a discursive trajectory with 'source', 'destination' and 'transition' domains; these allow us to identify, within each text's argument, starting-points, conclusions and obligatory transitions. The interest of this manoeuvre is, of course, that it allows us to compare the two graphs. This is why both ACU and CCU, which are both very short and convenient, are discussed here.

### *Analysis of the discursive trajectory of ACU*

The schema of motion A (see p. 205) leads us to one immediate conclusion: it gives a primary role to parties (we will see that motion C is very different in this respect). 'The party', 'our party', 'we, the PS', 'socialists' and, on the other hand, 'the PCF' or 'communism' constitute two series of terms which are dialectically related by the notions of dialogue, agreement, alliance, pact and union. Relations between the PS and the PCF are clearly

# ACU

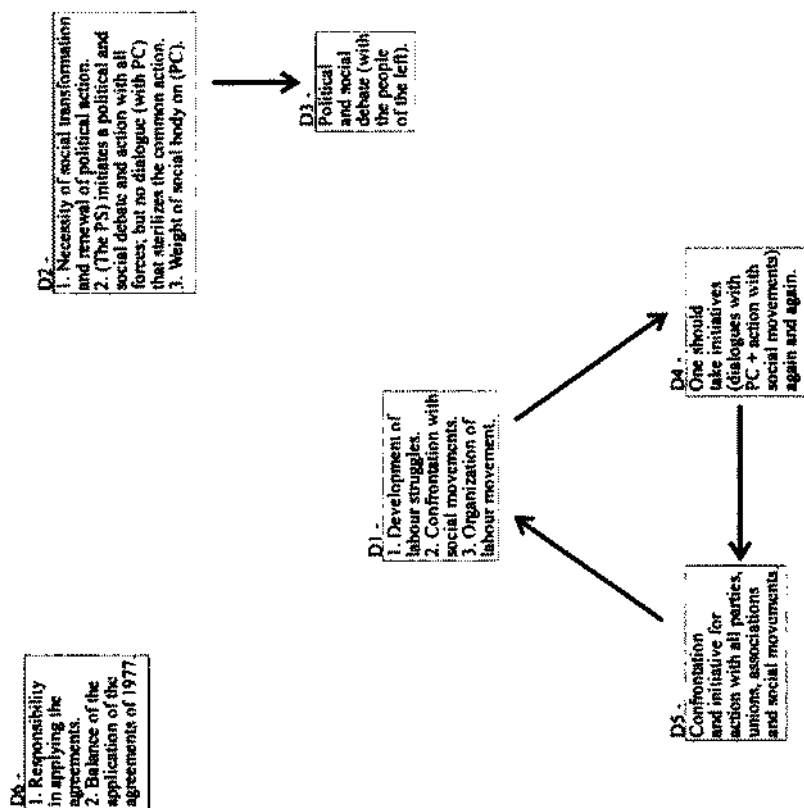


posited as being the central problem. It should, however, be noted that the two series are not symmetrical. On the one hand we have 'socialists'; on the other we have only 'the PCF' or 'communism'. This relates to the analysis made by current A, and indeed by the whole PS, and to the distinction it makes between the leadership and the base of the PCF (cf. the preamble to motion A: *"We all know under what conditions the Communist leaders kept the right in power, even though they could not prevent the Parti Socialiste from becoming the most important party in France"*).

Although motion A centres on relations between the PS and the PCF, it does not give both parties the same importance. The entire argument centres on the PCF: both source domains (D11 and HD2) and centre domains (D10 and HD1) are devoted to an analysis of its attitude, whilst destination domains (D13, D7, D9 and D6) and lateral domains (D3) deal mainly with the strategy of the PS. The PCF is therefore shown to be the blind spot on which everything, and especially the future of the PS, depends. If we move from HD2 to D11, passing through D10, where the two meet, we have a short history of the PCF. HD2 reminds us of the preferential agreement reached by Communists and Socialists in France, whilst D10 mentions the breakdown of that agreement and D11 raises the question whether the PCF will change its attitude. Two paths are articulated with the central kernel. The first, which runs through HD1, D3 and D13, and partly through D6, is a refutation of the suggestions put forward by 'some' to 'include political parties in a vast alliance with unions and associations'. The other, which leads to destination domains D7, D9 and D6, uses the strength of the PCF as an argument in favour of an offensive policy of making proposals directed at the parties of the left in general (D9, D6) and the PCF in particular (D11).

Two major absences result from this polarization of relations between the PS and the PCF on the one hand, and within these relations of the PCF on the other hand. The first concerns the PS's share of responsibility for the breakdown of the earlier agreement; this is not even mentioned. The second concerns 'struggles'. Both these themes will play a major role in the arguments put forward in motion C.





*Analysis of the discursive trajectory of CCU*

Although motion C makes a point of asserting support for the 'union of the left' strategy, a delinearization of the text – which is basic to ADA – allows us to conclude that that is not the recurrent theme of the motion (see the schema on p. 207). 'PS', 'PC' and 'union of the left' do not appear as such in any domain; the political agreements mentioned are past electoral agreements, and they, moreover, appear in an isolated domain (D6) which is outside the main circuit. The two central circuits are devoted largely to the idea of a debate and joint action with 'social movements', 'parties, unions, associations and social movements', or in short with 'all forces', but the idea of a dialogue with the PC (D2) is treated with great suspicion. Although we have not included the paragraph on 'social movements' in our ADA corpus on the grounds that it is specific to this motion alone, and although we have only selected paragraphs centred on the 'union of popular forces' and 'the union of the left', it is interesting to note that, even so, 'social movements' do figure in two domains (D1 and D5). The characteristic feature of motion C is the priority given to the 'social' rather than the 'political': the search for a PC-PS agreement is postponed for better days, and priority is given to 'the union of popular forces'. The expression itself never appears in the domains identified by ADA, but it subtends the entire circuit of the argument, and is clearly developed in D5 ('confrontation and initiative for action with all parties, unions, associations and social movements').

We can see from these two simple examples that a comparison of the argumentative threads of two or more corpora is not without its interest; by the delinearization of these texts which results from the constitution of domains, the uniformity of their references (to the union of the left in particular) fades away and it becomes apparent that the motions take very divergent paths.

Because of the political past and ambiguous position of many members of the PS, the functional ambivalence of the motions, and the very special atmosphere at the Metz conference, our corpus consists of texts which divide the subject, and of which the reader intuitively senses that 'things could go in either direction'. ADA allows us to remove the discursive ambiguity and to see clearly where things are going.

### 7.3. From ADA69 to ADA80: new perspectives

The ADA69 procedure is dated in the sense that it corresponds closely to the problematic described in the introduction to this chapter. Now we will draw lessons from the experiences it permitted and outline, in very broad terms, a new project.

#### **ADA69**

##### *Findings*

As we have seen, the ADA69 programme produces two kinds of findings: a list of semantic domains and a list of relations of dependency between domains. The problem lies in deciding what these lists contain.

##### *Semantic domains*

Semantic domains are seen and interpreted as paradigms in which the value of lexical items is differentially determined in an equivalent distributional context. This view (like the whole procedure) is based upon a specific interpretation of Harris's analysis, in which equivalence is taken to mean identity of meaning (lexical synonymy and paraphrase). When compared with the discourse from which the corpus has been extracted, the semantic domains that have been produced appear to identify objects of discourse: referential elements scattered across lexical realizations. If that is the case, a problematic which regards discursive strings<sup>1</sup> as traces of a discursive practice or a manner of speaking should concentrate on the specific modalities of the construction of those objects (as actors, processes, situations, etc.; cf. Courtine 1981: 113f.). The semantic interpretation of semantic domains distracts from the apparent object of discourse analysis.

Semantic domains frequently contain modalized propositions such as 'X do Y', 'X must do Y', 'X not do Y' ... . Discourse analysts have always paid attention to this kind of findings, which they see as either

---

<sup>1</sup> A discursive string is a fragment of discourse as perceived by a naive reading; a discursive sequence is a system constructed by a description which integrates a string into a whole. The construction of sequences is not equivalent to the segmentation of strings; a single string may belong to several sequences.

points of division within a discursive formation or points at which different discursive formations intersect. Whilst this interpretation appears to be well founded (in terms of the hypotheses of discourse analysis), it can only be verified by taking into account all the modalities of enunciation. This is not simply because they mark the distance between enunciation and utterance (as defined by Dubois 1969), but because various zones within a discursive formation can be realized in different positions and forms of enunciation. In other words, the study of enunciation may provide a possible approach for a description of a discursive formation. But enunciation is beyond the scope of ADA69.

#### *Relations of dependency between domains*

Relations of dependency between domains are represented by graphs, and we interpret them by imagining the existence of deep argumentative paths (which diverge from the sequential order). The modality of their construction calls for further specification. The syntactic analysis (the encoding in the form of binary relations made up of two eight-position utterances linked by a connector) and the subsequent comparison homogenize (or flatten) and lump together complex noun phrases, verb phrases, propositions and clauses (in a relation of coordination or subordination). Here we are faced with the interpenetration between different discursive levels (sentence, sub-sentence, inter-sentence) and, more specifically, the relationship between the construction of complex noun phrases (determination and nominalization) and the surface concatenation of the propositions that make up a narrative or argument. The fact that this important phenomenon escaped the notice of ADA users relates no doubt to the absence of a sequential element in the procedure (or in its problematic). It would have opened up a whole field of description. What, for example, is the degree of dependency between the definition-description of the protagonists of a narrative and the unfolding of that narrative, or between the definition of objects and an argument?

The confrontation of the findings of ADA69 and their interpretations with the corpora under examination cannot but lead us to define a new strategy of description and analysis.

## *Critique*

Our three main criticisms of the present procedure can be summarized as follows.

### *Role and definition of paraphrase*

Although paraphrase has a central role in discourse analysis (not only in ADA69), it has not been studied in its own right. In fact, discourse analysts, like many linguists, have a very naive notion of paraphrase.<sup>1</sup> ADA69 defines paraphrase in compositional fashion: the relationship between two sentences is said to be one of paraphrase if the sum of their parts constitutes the same meaning through lexical identity or equivalence (cf. the comparison phase in ADA). We must therefore:

- relativize the role of paraphrase by recognizing that discourse is not restricted to the production of significations by lexical substitution;
- approach the problem of paraphrase in a new way which corresponds to the overall approach of our project of studying 'the other in the same', in which we study those relations between syntactic structures which allow a propositional content which is stable (in terms of its discursive construction) to be invested in different meanings (shifts in terms of lexicon, enunciation, aspect, etc.).

### *Role and definition of syntactic analysis*

In ADA69, the autonomy of syntax is taken to mean that a universal (propositional) form exists prior to any discourse. Hence the form and role of syntactic analysis: discursive series are coded in (or forced into) a syntactic format, and syntactic analysis is given absolute priority over the textual algorithms. This also explains why discourses which this concept of syntax would define as 'deviant', 'non-standard', 'oral', etc., cannot be processed. We therefore have to adopt a new approach to syntactic analysis and to syntax itself. To mention only two starting points:

- A phase of syntactic analysis which exists prior to and independent of the discourse analytic phase is a theoretical and procedural impossibility. We have only to evoke the examples of ambiguity well-known from transformational syntax (*I make a picture of the children in front of the bank*, etc.) to see that syntactic analysis cannot succeed unless it resorts

---

<sup>1</sup> For a different – and exemplary position – see Fuchs (1979).

to discursive knowledge (definition of process, construction of protagonists, etc.). At a deeper level, these examples of 'ambiguity' are not accidental, and may be characteristic of a discursive mechanism. We should therefore conceive of an interactive syntactic analysis which returns to an initial minimal analysis in order to refine it.

- Forms of syntax must be investigated. We have adopted a recently developed parser (DÉREDEC, see Chapter 8) as a minimal syntax for the construction of new algorithms. In the longer term we are working towards an analysis that can articulate two systems: a system of hierarchical rules and a system of sequential rules.

#### *General aim of the procedure*

ADA69 is a machine for producing identity. The construction of the corpus freezes its conditions of production, whilst the interpretation of the findings gives discourses a reified identity such as 'socialist (or communist) discourse', or even 'the socialist discursive formation'. Discourse analysis has always been a typology which disregards rhetorical forms (or reduces them to oversimplified schemata such as basic propositions). Its basic aim was to discover typical discourses which amalgamate and freeze a political apparatus, a doctrinal content, and the different types of memory that are mobilized in speaking.

#### **Prospects of ADA80**

The construction of algorithms for discourse analysis is no more than one aspect of a wider problematic. There are two reasons why this is so:

- Only zones which are well-known or hypotheses which are general enough to be tested against a fairly wide corpus can be calculated. Given the inadequacy of current research into enunciation and concatenation, it is difficult to go beyond this stage at the moment.

- The questions the analyst formulates with respect to the corpus and his interpretations cannot, by definition, be calculated.

In order to arrive at a description of discursive strings, we envisage the construction of two spaces and of a number of algorithms corresponding to different objects within those spaces:

- A 'vertical' space which refers to the historical dimension of the discourse in question and which governs the algorithms used to group the units extracted from discursive strings. Every utterance is caught up in a

concatenation of utterances belonging to other discursive sequences (either earlier or contemporary) that constitute its conditions of existence.

- A 'horizontal' space referring to the 'thread of the discourse', to the complex unity in which discursive strings become sequential. Every utterance is caught up in a concatenation of utterances which is organized in terms of a system of sites of enunciation and of several rhetorical disposition systems.<sup>1</sup>

- The meaning of the term *discourse analysis* must definitely be altered. It designates both the breakdown of discursive strings and a form of regression which reduces the complexity of those strings to a law or over-simplified model of representation. It is vested with the hermeneutic project of uncovering the system of thought, the representations or the attitudes of those who produced the strings.

We wish to abandon this practice and to replace it with a contradictory practice which combines elements of morphology and reading:

- It is morphological in that the analyst describes forms which group together or distribute heterogeneous elements of all discursive sequences.

- It is a form of reading in that the analyst controls and writes this description. There is in fact a close analogy between the act of reading and that of description. Any reading segments a text, privileges certain elements at the expense of others, brings together elements that have been scattered, and disperses elements that have been brought together.<sup>2</sup> It is our belief that these interventions, which are made in what might be termed a 'wild' or 'unconscious' way in 'spontaneous reading', can be transformed into rule-governed interventions which break down the object of reading along the very axes that structure it. Discourse analysis will no longer be an aid to reading (cf. Pêcheux 1981: 17). It will be a stimulus to reading.

---

<sup>1</sup> These two spaces are not separated from one another, and nor are they unrelated. Sites of enunciation and dispositions are not constantly being invented by the subject of the enunciation; they are governed by a discursive formation. By contracting the thread of the discourse, vertical algorithms make it possible to compare any point on the thread with the whole thread; they function as a memory.

<sup>2</sup> Cf. Compagnon's (1979) fine description of reading.





## Strategies for textual description

Alain Lecomte, Jacqueline Léon and  
Jean-Marie Marandin

### Introduction

Reading or describing texts means performing material operations on words, utterances, sequences and configurations that make them up; we upset or distort the linear order of a sequence of utterances in order to grasp certain component elements (lexical, syntagmatic or phrasal elements), which we then recombine in a different space and in accordance with different laws. Alternatively, we may keep to the linear order in order to isolate sequences defining islands of coherence. Compagnon (1979: 37) compares the act of reading to the actions of someone who cuts quotations out of a text and reassembles them to make a new text:

Reading (solicitation or stimulus) and writing (re-writing) have nothing to do with meaning; they are ways of handling or manipulating a text, and if at the end of the process a meaning emerges, for better or worse, that is a different matter.

This formulation defines a type of reading which is defined only by the play of pleasure and the play of the signifier. A morphological approach to texts (or to textual material in socio-historical research), however, can be defined in terms of a set of 'manipulations'; its specificity, as opposed to that of the reading described by Compagnon, lies in the need to control the manipulations by carrying them out on the basis of formal markers which we can hypothetically define as the lines of fracture and of cohesion which inform both the intratext and the reading we make of it. Certain of these manipulations are so controlled that they can be used in algorithmic procedures which can be applied to texts (or textual material).

DÉREDEC appears to us to be the most appropriate software for this purpose. In section 8.1. we will describe certain of DÉREDEC's formal properties, and in section 8.2. we will explain why this software is an appropriate instrument for the objectives we wish to attain in discourse analysis.<sup>1</sup>

## 8.1. Summary of formal properties of DÉREDEC

DÉREDEC, which was designed by Plante (1981), is a software programme written in LISP, a language which is particularly suitable for the processing of symbolic material. DÉREDEC retains all LISP's properties: the inputs, outputs and programmes are tree-structures. Like LISP, DÉREDEC makes extensive use of recursivity.

The formalist conventions Plante adopts to represent and manipulate these tree-structures are similar to Woods's Augmented Transition Networks (ATNs) (Woods 1970). They make it possible to add information to the nodes of the tree-structures, to store it, analyse it, and to modify it in the course of subsequent operations.

Grammars based on ATNs have traditionally worked in a top-down direction. They require an exhaustive description of the paths which 'descend' from general syntactic symbols (such as S, NP or VP) to lexical items. Plante, on the contrary, adopts a bottom-up approach, constructing structured syntagmas step by step on the basis of lexical items. This imposes fewer constraints, and means that the parser (the surface grammar, for instance) is a more robust system which always produces results (see pp. 222-224 below for a discussion of using a bottom-up approach in discourse analysis).

DÉREDEC therefore allows us to construct descriptive grammars of texts (which may or may not be linguistic) and systems for exploring texts. These two procedures and the objects they manipulate have all the

---

<sup>1</sup> This chapter is a sequel to Chapter 7 and to Plante (1983). Our thanks are due to Dominique Bègue who worked with us on an earlier version of this text and provided valuable information on DÉREDEC's position in the history of software for processing texts in natural languages (cf. Bègue 1983).

same syntax: they are tree-structures<sup>1</sup>. All the system's functions can therefore be applied at any point of the procedure.

Descriptive grammars of texts and exploration systems implement descriptive functions and exploratory functions respectively.

### Text-descriptive functions

A text description implies the construction of tree-structures on the basis of atomic expressions. For example: a syntactic description of a text (in a natural language) takes the form of a hierarchical series of analyses of the component elements of sentences in that text. DÉREDEC uses the term EXFAD (*expression de forme admissible à la description*, expression of a form that is admissible to description) to refer to both atomic expressions (lexical items) and nodes (syntagmatic components). DÉREDEC also makes it possible to construct horizontal relations between component elements; these are referred to as relations of contextual dependency.

Moreover, information of a paradigmatic nature can be added to atomic expressions in the input sequence. This takes the form of an EXFAL (*expression de forme atomique liée*, expression of an associated atomic form). EXFALs convey paradigmatic information about the behaviour of an atomic expression in sequences other than the sequence under examination. During a syntactic description of a text, syntactic or semantic features can be associated with lexical items and can be given the EXFAL form. Features of strict sub-categorization can, for example, be associated with verbs, and syntactic features such as gender can be associated with substantives.

### Text-exploratory functions

Text descriptions produced by automata will be analyzed by exploratory functions whose arguments are pattern-matching structures which attempt to make a given structure correspond to the structure of the text. These are known as exploration models.

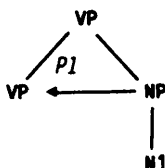
---

<sup>1</sup> This is a minimal characterization of what makes possible the composition of retention structures and functions of manipulation (description and exploration) as different objects in DÉREDEC (Plante 1984).

An exploration model is a parenthetic representation of a part of a tree-structure. Take the model:

((VP ((VP) (P1 -) (NP((=N1))))))

In the text under analysis, this model corresponds to a tree-structure of the following type:



The element we are looking for is preceded by the sign '='. We look for only one element in each model. The element may be a syntagmatic category, a basic category or an atomic expression (a lexical item), and it may or may not be accompanied by one or more EXFALs. The above model is read from right to left; we see a substantive (N1) that is part of an NP which is the object that belongs to a verb phrase (VP). The VP furthest to the left is the category which dominates the two component elements that are linked by the relation of contextual dependency P1. Relations of contextual dependency are always oriented: the sign '-' indicates that the NP is the destination of the relation (the sign '+' would indicate that it is its source). Orientation refines our description and opens up new areas of research, notably by allowing a distinction to be made between 'determiner' and 'determined' within an NP.

### Schema for the application of DÉREDEC

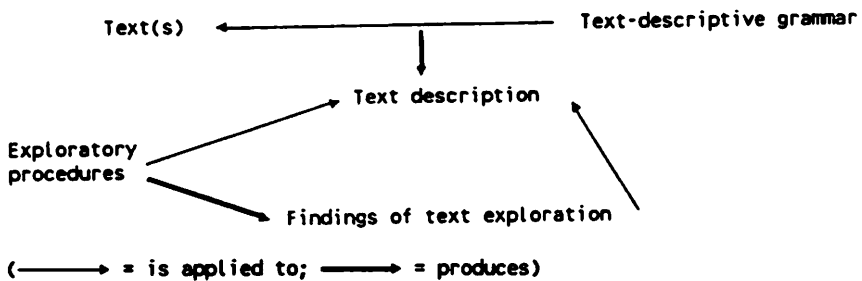
In order to use DÉREDEC, it is usually necessary to apply descriptive functions before applying exploratory functions. The procedure therefore involves two successive stages:

- application of descriptive functions to provide a text description (a syntactic analysis, for example);
- application of exploratory functions to that description.

The use of DÉREDEC is not, however, reducible to a linear sequence of

description and exploration. Descriptive and exploratory phases can be combined by using exploratory findings in further descriptions which can in turn become the object of further explorations. Translation functions ensure that the various levels interact. They make it possible to transform any EXFAD (any input or output of descriptive functions) into a constraint on an exploration model; inversely they make it possible to transform an exploration model, or an argument in an exploratory function, into a constraint on an automaton or on an argument in a descriptive function. Any finding at any level can therefore be used as an input into a different level of the system, as tree-structures are found at every level.

The application of DÉREDEC can be represented by the following schema:



It is obvious from the schema that a text description can still be the object of multiple explorations which will 'enrich' it by providing data for a subsequent analysis. Similarly, a text which has been so enriched by various explorations can become a text which can be submitted for a new descriptive text grammar.

Neither the data nor the results are completely static and they can always be used again in a further analysis.

### **Surface grammar**

We use the syntactic analysis described by Plante (1983). This parser is a bottom-up recognition grammar for French. It comprises some 1600 rules for some 60 automata which construct syntagmatic categories and relations of contextual dependency (RCDs) between basic categories. It produces a

hierarchical analysis of sentences, as in the following example:<sup>1</sup>

*J'ai voulu faire insérer un article dans un journal régional pour  
la continuation de la peine de mort*  
[I wanted to include an article in a regional newspaper in favour  
of the retention of the death penalty]

The analysis provided by the surface grammar is represented on p. 221. Note that certain RCDs are marked '?', which indicates that they are in competition with other RCDs.

### **Ambiguity**

Without going into the analyses produced by the surface grammar in detail, we will look at the handling of some of the ambiguities it may encounter.

#### *Ambiguities in functional form*

Certain syntactically ambiguous functional forms are given a temporary category by the categorization procedure before the syntactic parser is applied. The removal of the ambiguity of such categories (for example *le* as article and *le* as pronoun) is effected by the surface grammar's automata before syntagmatic groupings are made.

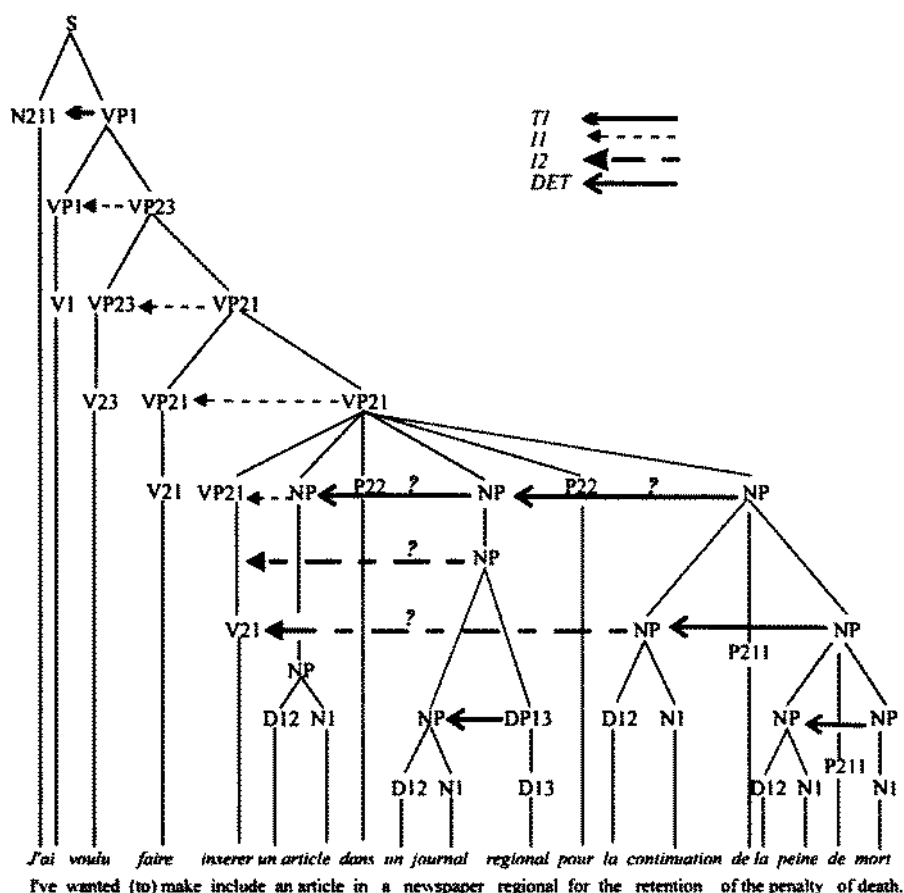
#### *Ambiguity in the analysis of components*

The surface grammar uses RCDs to link syntagmatic categories. There are five types of RCDs:

- the 'theme-issue' relation (TI) links any non-propositional noun syntagma to the left of a conjugated verb to the rest of the proposition;
- 'development of the issue' relations (II and I2) correspond to complements to the verb;
- the 'determination' relation (DET) groups together complex noun phrases, or a relative clause and its antecedent;
- the 'coordination' relation (CO) groups together enumerations or coordinate clauses.

---

<sup>1</sup> This example is taken from a corpus of letters collected by R. Dulong.



Basic categories:

N211 subject (personal pronoun)  
 N1 substantive  
 V1 conjugated verb  
 V21 infinitive  
 V23 past participle  
 D12 article  
 D13 adjective  
 P211, P22 prepositions

Syntagmatic categories:

S sentence/proposition  
 NP noun phrase  
 VP1, VP21, VP23  
 verb phrases

Relations of contextual dependency:

T1 theme-issue  
 I1 development of the issue; direct  
 complement to the verb  
 I2 development of the issue; other  
 complements to the verb  
 DET determination  
 ? ambiguity

If ambiguities are structurally possible in the analysis of component elements, the surface grammar generates two sets of RCDs. Thus, in our example, the surface grammar generates on the one hand two RCDs starting from the syntagma *pour la continuation de la peine de mort* [in favour of the retention of the death penalty]:

- DET? connects the syntagma with *dans un journal régional* [in a regional newspaper] (cf. 'a newspaper in favour of the retention of the death penalty');

- I2? connects the syntagma with *insérer* [include] (cf. 'include an article in favour of the retention ...'). 'In favour of the retention' can therefore be regarded as an adverbial complement.

In the same way, the surface grammar generates on the other hand the following RCDs starting from the syntagma *dans un journal régional* [in a regional newspaper]:

- DET? connects the syntagma with *un article* [an article] (cf. 'an article in a newspaper');

- I2? connects the syntagma with *insérer* [include] (cf. 'include in a newspaper').

The question marks indicate that the syntactic analysis is not yet finished. They explicate the ambiguity, which may be dealt with in a later stage of processing. If the grammar cannot construct an analysis for a given sentence, it stores it and continues the analysis of other sequences.

## 8.2. DÉREDEC and discourse analysis

The properties described above mean that DÉREDEC is a highly appropriate instrument for the choices and objectives of a morphological description of discourse.<sup>1</sup> Not only does it allow us to overcome the obstacles encountered in attempts to computerize automatic discourse analysis; it also opens up new possibilities for an approach to texts which is not reducible to a description of a single level of structuration. The ADA69 procedure only dealt with constituent phrases (NP or VP; in practice, priority was given to NP) by producing zones of lexical substitution and by approaching phenomena of determination. This reductive

---

<sup>1</sup> An account of discursive morphology can be found in Chapter 7. See also Marandin (1983) and Lecomte and Marandin (1986).



approach, which is characteristic of much discourse analysis, must be abandoned in favour of a stricter definition both of the discipline and of its object. DÉREDEC allows us to make progress in this direction by making computerized descriptive montages of texts.

Plante's general intentions in conceiving and constructing DÉREDEC were very close to the viewpoint we wish to elaborate with respect to a morphological approach to discourse (Plante 1981). We can note three points of agreement:

(1) Any description of a text must be based upon a description of that text's linguistic structuration, and of its syntactic structuration in particular. A text is written<sup>1</sup> in a given language. Taking into account the syntactic dimension alone will obviously not exhaust the linguistic materiality of a text (the hypostatic view of syntax taken by much French discourse analysis, and by ADA69 in particular, notwithstanding). We have yet to realize a description of forms of enunciation, lexical forms and forms of sequentiality which can be used for discourse analysis. DÉREDEC makes it possible to programme certain algorithms which can approach these forms in texts.

(2) A description of a text is not to be confused with an analysis which reduces it by isolating either a textual zone which can be submitted for description (and ignoring the rest of the text) or a mechanism which is assumed to represent the workings of discourse (syntax, enunciation, lexicon ...). A description of such a complex object calls for a multiplicity of descriptive viewpoints. This is why Plante refuses to use DÉREDEC for such specialist purposes as the statistical processing of the phenomena of co-occurrence or the simulation of arguments by using semantic representations. The possibility of adopting these procedures and of using them to enrich a description remains open (Plante 1981: 4-5). The need for a multiplicity of descriptive viewpoints stems not only from the complexity of the object, but also from the very nature of the act of describing texts (cf. Todorov 1984: 14ff.). If any description implies interpretation (and if any morphology is already caught up in a reading; cf. Chapter 7), no description can be univocal. It remains for us to regulate these plural viewpoints by recognizing that the levels at which language is structured are heterogeneous and that the manner in which they are interwoven in

---

<sup>1</sup> It will be recalled that we are only describing written texts. Our current understanding of the structurations of oral language is inadequate, and we cannot describe a corpus that has been transcribed on its basis.

discourse has a destratifying effect.

(3) The choice of a bottom-up strategy for data-processing also reflects, finally, a prejudice in favour of description. It minimalizes the initial inputs, and privileges a constructivist approach in which the stages of the analysis are as important as its findings. It also produces objects which can be remodelled. A bottom-up description is not a local specification of a theory which must be able to predict everything. It is an unstable construct trapped between a generality which is acquired through repetition and a singularity which might disturb that generality or escape it. Many of the criticisms and self-criticisms that have been addressed to discourse analysis arise from this point. The debates that took place over the notion of the corpus are, in this respect, exemplary. Discourse analysis has frequently been criticized for being no more than a digression in that the findings of the analysis reproduce the knowledge about the exterior of a discourse (conditions of production, enunciation situation, history of ideas) which was used to effect its closure to begin with. This criticism (which is justified when one thinks of interpretative discourses on particular corpora) clearly indicates the difficulty discourse analysis has in extricating itself from the knowledge (of the language, history or science) which it mobilized, 'articulated' and projected onto texts which were powerless to resist.<sup>1</sup>

We will now enumerate and describe certain of DÉREDEC's properties, functions and objects which were exploited in conceptualizing and programming algorithms for the description of texts.

### **Surface grammar and discourse analysis**

The absence of a syntactic parser was the major obstacle preventing the development of a computer-assisted discourse analysis. More specifically, it made it impossible to analyse a corpus of any great length (which explains why discourse analysis limited descriptions to a restricted corpus which was claimed to represent a larger corpus). That parser does now exist, and it has two major characteristics:

- It is designed to produce different states of a syntactic analysis.

---

<sup>1</sup> In terms of the corpus, note that Plante provides for the following possibilities: description of a text; description of a part of a text; comparison of several texts; investigation of one text on the basis of another.

- The syntactic analysis procedure is not blocked if it cannot construct an analysis of an atypical construction.

In other words, syntactic analysis is not seen as an isolated, perfect or primal realm (this was the interpretation that discourse analysis made of the Chomskyan principle of the autonomy of syntax), but as a provisional phase. It is therefore possible to

- alter its terms in accordance with the use to be made of it in describing a text;

- enrich and refine analyses by exploiting data produced elsewhere or introduced interactively.

Two examples will illustrate these possibilities.

### *Modification of terms of analysis supplied by the surface grammar*

The surface grammar generates a 'theme-issue' RCD between a conjugated verb and a (non-propositional) syntagma to the left of that verb. This is not, strictly speaking, a syntactic relation.<sup>1</sup> If we wish to introduce syntactic relations ('subject of', in our example), it is possible to make a new analysis of the text description produced by the surface grammar by using exploratory models. The procedure is as follows.

### *Analysis of text descriptions*

The surface grammar gives the following syntagmas in the 'theme'-position. They are either the subjects of a verb or in apposition (to the subject):

- (1) *John* eats
- (2) *Hurried*, *John* runs
- (3) *As a schoolboy*, he used to lie
- (4) *The bachelor*, a village *teacher*, speaks

---

<sup>1</sup> This parser has often been criticized because of its use of 'theme-issue' analysis. It is clear that it is neither a syntactic analysis nor a functional analysis of the types developed in contemporary linguistics. It might be compared with the analysis proposed by Halliday (1967).

The analyses supplied by the surface grammar do not depend upon this relation to any great extent. It is possible to programme a parser which is either partly or completely different to that proposed by Plante. The 1984 version of the surface grammar is different from the 1981 version.

The terms italicized are the heads of noun phrases (NPs) connected to the verb by the 'theme-issue' RCD. Subjects are in most cases in a 'theme'-position, but in cases of inversion, subjects can also be related by I1:

(5) Then came *the war*

### *Method*

In order to identify subjects to the left of the verb, we must locate the NP which is in the 'theme'-position. In example (2) we must eliminate 'hurried', which is categorized in D  REDEC as VP23 (past participle phrase). We must then make a distinction between 'NP subject' and 'NP in apposition' by discovering whether there is a personal pronoun (N211) – in which case the co-occurent NP is in apposition, like 'as a school-boy' in example (3) – and by studying the distribution of determiners in both NPs – in example (4), e.g., 'the bachelor' is the subject, and 'teacher' is in apposition.

In order to identify subjects to the right of the verb, like in example (5), we must identify those sentences which have no theme, and then the NPs which stand in an I1 relationship to the verb.

### *Search procedure*

Each search procedure is executed by means of an exploratory model. Examples include:

(=S ((X) (TI –) (X)))

This model identifies sentences with a theme-issue relation (and, by applying it negatively, sentences without such a relation).

((VP) (((VP) (I1 –) (NP ((=N1))))))

This model identifies N1 in NPs which are in a I1-relation with VPs.

(S ((VP (TI –) (N211)) ((VP (TI –) (NP ((=N1))))))

This model is conjunctive. It identifies N1s in 'theme'-position, co-occurring with N211s (personal pronouns) also in 'theme'-position. Its 'outputs' are N1s which appear when both structural constraints are realized (N211 in 'theme'-position in VP; N1 in NP in 'theme'-position in VP).

These models will provide inputs for an algorithm which is now under construction, called *Variation syntaxique d'un item lexical* (VSIL; syntactic variation of a lexical item). This algorithm provides a syntactic definition of every item (full word) in a given text (or a sub-set identified by items in that text); it supplies, that is, its syntactic position and function. The definition may be synchronic, or it may follow the thread of the text ('dynamic VSIL' algorithm). This procedure will make it possible to study a text's syntactic handling of the lexicon and to compare one text with another on that basis.

*Refinement of the analysis supplied by the surface grammar*

Let us return to the analysis of the sentence

*J'ai voulu faire insérer un article dans un journal régional pour la continuation de la peine de mort*

[I wanted to include an article in a regional newspaper in favour of the retention of the death penalty]

Without going into how we approach ambiguity in discourse in detail, it is possible to envisage the construction of a procedure which will allow us at least to privilege<sup>1</sup> one reading, if not to choose between two readings.

Regarding the two ambiguous RCDs discussed above (p. 222), two types of information make it possible to privilege a single reading.

(1) Discursive information concerning the constitution of noun phrases in the discursive sequence. If we find groups such as *newspaper against the street urchins ...* or *article in favour of the government ...*, we can give greater weight to DET? in our analysis of *newspaper in favour of the retention* or *article in a newspaper* in the utterance.

(2) Lexical information concerning the strict sub-categorization of *include* (*include* + N + preposition of place + N). As a result of this sub-categorization, we will privilege the I2 relation for *in a newspaper*.

This leaves us with interesting cases of 'conflict'. Assume that we find *article in favour of the government* and *article in Paris-Match* in the

---

<sup>1</sup> The term 'privilege' is used to suggest a possible approach; the possibility of a non-privileged meaning has not been excluded. It is a discursive index of a possible transformation of the object, or of our view of that object.

textual sequence. It would seem that here we have to privilege an I2 relation between *include* and *in a newspaper* rather than the DET relation between *an article* and *in a newspaper*. The problem remains unresolved, but it provides an excellent anchoring point for the study of the constitution of noun phrases in a text, and of the modalities of the transition from the domain of the sentence to that of the noun.

### *Extension of the possibility of sub-categorization*

We have just looked at the possibility of a sub-categorization of basic category V (verb). The possibility of constructing sub-categorizations of all basic categories is another interesting property of the surface grammar. The tree-structure of the basic categories makes it possible to introduce particular information which may be pertinent to a particular search-operation without overloading the syntactic analysis phase (defined in its strict sense). For example, substantives are categorized as N1. Proper nouns have no particular characteristics in terms of the bracketing of syntactic components or the establishment of syntactic functions. They may, however, be an important index for the constitution of a thematic sequence. If N1s are sub-categorized as N11s (proper nouns) and N12s (other substantives) respectively, the syntactic parser will only take into account the first level of the sub-categorization (N1), whereas the sequential-segmentation automaton will work with the second level (N11; N12) (cf. pp. 233-234 below). The possibility of sub-categorization is a further example of the general ability of the surface grammar to supply different states of a description.

### **EXFAD and EXFAL**

An EXFAD is a tree-structure which encodes the description of a sequence (components and relations between components, in the present case). An EXFAL is a tree-structure associated with atomic expressions which conveys information that does not appear in the sequence itself. To simplify matters, we can say that an EXFAD supplies a state of a syntactic analysis, and that an EXFAL makes it possible to introduce information

about co-texts<sup>1</sup>. The only restrictions upon the nature of the informations are those imposed by our ability to imagine what is pertinent to the discursive description of a sequence in a text.

It is possible to envisage the EXFAL-ization of a given item in different ways:

- (1) one could add information about other lexical items with which it co-occurs in the text's propositions (we would then have the item as referring to the discourse itself);
- (2) to a sub-set of such items, one could add information about syntagmas which determine them (we would then have access to the discursive construction of the objects they denote and a means to study the impact of that determination on the construction of the propositions);
- (3) one could add information about syntagmas which are equivalent to it, in Harris's sense of that term, and which appear in a context which is lexically and syntactically identical (we would then have a discursive paradigm very similar to those constructed by ADA69); and so on.

The essential point is that this information is not a dead weight which encumbers the components of a sequence. It plays an active role in the formulation of exploratory models: a model may, depending on how it is defined, take into consideration a lexical item in a sequence, the union of a lexical item and the EXFAL attached to it, an isolated EXFAL or one of the EXFALS attached to an item. In other words, a lexical item can be treated in isolation, as an element in a paradigm, as the index of a paradigm or as a member of a paradigm.

This duality of structure, together with the syntax of the exploratory models, gives us now the practical possibility of regarding a lexical item or an utterance in a text as a point within a network of formulations or utterances. This corresponds to the function we ascribed to the distinction between vertical and horizontal algorithms: comparing each point (or a selected sub-set of points) of a textual sequence with what comes before it and what comes after it, respectively with sequences belonging to other texts, in a strategy involving a constructed co-text. This point is of crucial importance to the approach we are outlining. Whilst everyone who describes texts or strings of utterances now accepts that a sequence is

---

<sup>1</sup> The co-text can be defined as the text from which the sequences in question have been extracted, or as a corpus made up of other texts. Depending on how the description is oriented, it is possible to introduce information which is, to a greater or lesser extent, discursively constructed (synonyms, semantic fields, etc.).

lacunary, that it only 'holds together' thanks to something which is not materially realized, the nature of that non-material element remains to be defined. To describe it as an argument (Ducrot and others), an inference (Bellert 1970), or as a body of propositions which are regarded as true in the universe of belief (pragmatics), presupposes that we can reconstruct the absent element by calculations of utterances which are co-present in the consciousness of their co-enunciators. The omnipresence of the subject implies the adoption of a position which we do not share. The absent utterances may be invested in the interpretation or reading given by co-enunciators, as it is inscribed in the interdiscourse which subtends the text that is being read or interpreted. It must, then, be possible to track down this interdiscourse and to produce traces of it. The position we are adopting must be further developed by the production of descriptive montages. DEREDEC appears to offer us a possible means to realize them.

### **Exploratory models**

An exploratory model looks for an element in a textual sequence and/or in the EXFALs attached to lexical items in that sequence by using a pattern-matching procedure. The element in question can be defined in terms of the conjunction or disjunction of constraints of various orders. We will illustrate this possibility with two models carrying out a simple search in a text:

(S ((VP) (TI-) (NP ((NP:patron) (DET+) (NP ((=N1)))))))

This model combines two orders of constraints: a syntactic constraint (the element we are looking for (N1) is in the 'theme'-position in a sentence) and a discursive constraint (N1 is determined by a set of substantives gathered in a category named 'patron', and that set is constructed in the textual sequence). This model may provide a starting point for research into the intertwining of thematic threads, into how, and at what point in the text, a thematic thread becomes attached to objects that are introduced elsewhere in the text. The second model:

(=S ((N2111))  
 ((N213))  
 ((NP ((D1311))))))



This very simple model searches for all the propositions in which there is an explicit reference to a personal pronoun of a specific kind (N21...), or a possessive adjective (D1311). It therefore allows us to make a preliminary 'mapping' of the textual sequence in terms of its enunciation.

The syntax of the exploratory models allows us to combine characteristics relating to syntax, lexicon, co-occurrence and membership of a paradigm. And this corresponds to the definition of a morphological approach in discourse analysis: a search for configurations of elements belonging to differentiated levels of the language. These configurations are our object of study, and DEREDEC allows us to locate their presence in a text. We call them forms. A form is a unit into which formal components of the language which are distributed regularly along the thread of a discourse are integrated. We are currently studying 'theme of discourse' and 'sequence' forms.

### **Intradiscourse**

The fact that inputs to exploratory models consist of text descriptions referring to the domain of the sentence does not mean that our description must remain restricted to that level. As we have seen, EXFAL-ization introduces qualified elements of a co-text into the sentence. This is one way of making the boundaries of the sentence 'porous'. The notion of form which governs our descriptive project disregards, moreover, the sentence as an autonomous domain. More radically, it does not regard it as a pattern or model for a discourse or text; our forms are neither basic sentences, nor macro-propositions, nor strings of propositions. We can, finally, characterize a sequence – in intuitive terms – as a dynamic space in which reformulations and transformations take place. The classic example is that of the retroactive constitution of a sequence, which, through the cataphoric reprise of a series of utterances (in the form of a noun syntagma that is their co-referent) and by naming them, provides a framework for their interpretation and amalgamates them. We are by no means in a position to describe these phenomena, and the ground for their description has yet to be prepared. In order to do that, we must be able to

- (1) replace any element that has been momentarily isolated from a text in the text that precedes it and comes after it (this is a different procedure from the EXFAL-ization procedure, which transforms the text into a list);

(2) work on sequential units larger than sentences, on, that is, units in which boundaries between sentences are merely elements like any other. DÉREDEC opens up both these possibilities.

*Diachronic description of the intradiscourse*

The study of a phenomenon in terms of what comes before and after it in the text is one of the objectives Plante (1981: 7) describes for DÉREDEC:

We can require a description of a textual event to take into account the particular position of that event within the text as a whole, and to differentiate between that part of the text which follows that event and that part which precedes it.

We will take as an example our definition of VSIL (see p. 227 above) and of one of the objects it allows us to study: the phenomenon of nominalization in discourse. In discourse analysis, nominalization has usually been studied in paradigmatic terms. Studying it in its textual dimension means identifying the moment at which we move from the sentence to the substantive.<sup>1</sup> It may be important to know whether the utterance corresponding to the nominalization precedes it or follows it. Intuition suggests that when the utterance (or utterances) follows the nominalization, it introduces a specification or definition which can be invested in the narrative, or a description which can be invested in a reformulation, and which particularizes or exemplifies it. When the utterance (or utterances) precedes it, the nominalization limits a sequence and introduces it into what Lecomte (1981: 74f.) calls a 'levelling-down'. If several nominalizations appear in the text at the same moment, we presumably have an index of a specific textual mechanism: it may be possible to compare several texts or several zones within a single text on the basis of that mechanism.

---

<sup>1</sup> The present discussion is restricted to examples of nominalization in which the relationship between verb and substantive is morphologically marked, or in which nominalization occurs in a free syntagma.

### *Segmentation of intradiscourse into sequences*

As we said earlier, the use of DÉREDEC in discourse analysis means that we can operate on the frontiers which define the domain of the sentence. Operating on these frontiers is equivalent to asking: what does 'making a sequence' mean? How does it come about that the formulations of a text or discursive surface 'hold together'? If we refer back to the concepts established in other areas of discourse analysis, it would seem that, in order to answer these questions, we must explore the constitution of an intradiscourse. Our hypotheses suggest that we should represent an intradiscourse as a space whose structure is based upon morphemes and syntactic regularities which can be identified by using exploratory models. Relations of anaphorization, co-reference and deixis in particular can be thematized in this way. To be brief about it, we can say that these relations homogenize series of component elements. Anaphorization and cataphorization give a sequence a coherence-effect or, to put it another way, ensure the stability of a position within a discourse. Deictics establish a reality-effect which designates a sequence as referring to an object-of-discourse, whilst co-references lead to the stratification of the sequence, and certain expressions therefore appear to sum up or to condense a number of earlier formulations and are in a position to refer to the same object-of-discourse.

It is possible to describe this phenomenon of 'levelling-down' by co-reference in the production of explanatory or demonstrative texts or, in other words, in arguments (Lecomte 1981). We will take as an example:

*La lumière du soleil est blanche; après avoir traversé un prisme, elle montre toutes les couleurs qui existent dans le monde visible. La nature reproduit le même résultat dans la belle gamme des couleurs de l'arc-en-ciel. Les tentatives d'expliquer ce phénomène sont très anciennes.*

[Sunlight is white; after passing through a prism, it shows all the colours that exist in the visible world. Nature reproduces the same results in the rainbow's beautiful range of colours. Attempts to explain this phenomenon have been made since ancient times.]

We have here a combination of

- anaphorization: 'Sunlight ... it'
- co-reference: 'Nature reproduces the same results'

'Attempts to explain *this phenomenon* ...'

- deicticity: 'the same results ... *this phenomenon*'

Together, these phenomena construct the discursive homogeneity of the sequence by reformulating the series *Sunlight is white; after passing through a prism, it shows all the colours that exist in the visible world* and by using nouns (*results, phenomenon*) to referentialize it.

DÉREDEC is then used to construct sequences. In the present case, we have to create new automata to be added to those that make up the body of the surface grammar. These new automata which introduce supplementary marks and categories, are to be applied later. To begin with, for example, we simply describe the simplest anaphoric relations and use them to categorize a string of sentences as a sequence, and to place the supplementary punctuation marks which delineate the sequence in a text description. Depending on the goals one sets oneself in the analysis, it is now possible to 're-establish the anaphora'. To do so, it is sufficient to construct an automaton with a LISP function which allows to replace an anaphoric element, e.g., a personal pronoun, with the NP to which it refers.

Given our first experiences with DÉREDEC, it is possible to envisage an extension of the method. Having broken the text down into sequences on a first level, we can look for constructional regularities in sequences, group them into sequences of a second level, reveal relations of substitutability between NPs (between, for example, *sunlight, nature* and *rainbow*), and, finally, connect the latter and deictics by means of the relations of lexical repetition and substitution we have detected. If an index of deicticity appears in a sequence which has been agglomerated in this way (in, that is, a sequence which has been constructed on the basis of recognition of homomorphism or of an intertextual link, and not by segmentation), we can assume 'levelling-down' or referentialization.

Our aim is to uncover the schema which organizes sequences of the first level. When that schema has been discovered, our syntactic description explodes, or 'decompacts' so to speak. As a result of this operation, it is possible to see in which points of the text the element which establishes relations between formulations circulates.

## Three stages of discourse analysis

Michel Pêcheux

### **First stage of discourse analysis: DA-1 as a methodological exploitation of the notion of structural discursive machinery**

#### *Theoretical position*

- A process of discursive production is regarded as a self-determining and self-contained machine, as a structural subject which determines subjects as producers of their discourse. Subjects believe that they 'use' their discourse, whereas they are in fact subjected to it. They are its 'serfs' or 'supports'.

- A natural language (in the linguistic sense of that term) constitutes an invariant base upon which a heterogeneous multiplicity of juxtaposed discursive processes is deployed.

The adaption of this 'structuralist' position, which will become less pronounced after DA-1, involves the rejection of the assumption that a universal metalanguage is innately inscribed in the human mind, and of the notion that an intentional subject can be at the origin of the discourse it enunciates (this rejection will not change from DA-1 to DA-3).

DA-1 presupposes the possibility of making two successive gestures:

- (1) assembling a set of empirical discursive traces (a 'corpus of discursive sequences') on the basis of the hypothesis that the production of those traces was in fact dominated by a single discursive machine (a myth, an ideology, an episteme);
- (2) constructing, on the basis of that set of traces and using linguistically stable procedures, the combinatory space within which empirical variations on the traces are distributed. The actual construction of this space constitutes the epistemological gesture of 'working backwards' towards the discursive machine which is assumed to have generated them.

*Procedural implications*

- The starting point for a DA-I is a closed corpus of discursive sequences selected (usually on the basis of their proximity to a key word referring to a theme) from a discursive space which is assumed to be dominated by stable and homogeneous *conditions of production*. This is why the various examples of DA-I focus upon textual discursivities which are themselves self-stabilizing, such as political discourses which take a doctrinal or theoretical form.

- The linguistic analysis of each sequence is an essential *preliminary* to the discursive analysis of the corpus. Linguistic analysis is regarded as an autonomous operation which can be carried out exhaustively and once and for all. It assumes the neutrality and independence of syntax. Enunciation and the underlying constraints on discourse remain opaque to it (in other words, it takes them into account implicitly).

- The discursive analysis of the corpus initially comprises the identification and construction of the position of intersequential paraphrastic identities (which occur between fragments of sequences deriving from different discursive surfaces). Insofar as they are points within a combinatory variation, these paraphrastic identities form the site of the inscription of the *basic propositions* characteristic of the discursive process under consideration. Identification of the trajectories linking these propositions may lead to the extension of the analysis. Ultimately, the idea (and it is no more than an idea!) is to arrive at a discursive algebra which will make it possible to construct formally the generative structure associated with the corpus on the basis of the set of arguments, predicates and operators which governs the construction and transformation of propositions.

- The 'interpretation' consists of re-inscribing the results of this analysis within the initial discursive space, as though they were 'answers' to the questions which organize it thematically. The interpretation usually takes the differential form of a structural comparison of juxtaposed and heterogeneous discursive processes.

*Conclusion*

DA-I is a procedure involving a set series of stages, and it is theoretically and methodologically restricted to a predetermined beginning and end. It operates in a space in which discursive 'machines' constitute juxtaposed

units. The existence of the *other* is therefore subordinate to the primacy of the *same*:

- At the level of 'empirical' discursive alterity, the other is reduced to being either 'the same' or a residue, as it is the combinatory basis for a single discursive process.
- At the level of 'structural' alterity, the other is in fact no more than an incommensurate difference between 'machines' (each of which is self-identical and self-contained). It is, that is, a difference between machines which are 'the same'.<sup>1</sup>

## DA-2: From the juxtaposition of discursive processes to the thematization of their uneven intrication

The theoretical displacement which inaugurates the second period results from a philosophical shift of perspective. Discourse analysis now takes as its object *relations between* discursive machines. For DA-2 these relations are *unequal power relations* between discursive processes which structure the entire field as 'a complex whole in dominance' consisting of 'mechanisms' with an unequal hold over one another: the notion of a *discursive formation*, which is borrowed from Michel Foucault, begins to explode the notion of a closed structural machine in that the mechanism of a discursive formation stands in a paradoxical relation to its 'exterior'. A discursive formation is not a closed structural space in that it is constantly being invaded by elements from elsewhere (i.e. from other discursive formations). They are reproduced within it and provide it with its basic discursive truths (in the shape of 'preconstructs' or 'transverse discourses').

The notion of interdiscourse is introduced to designate the 'specific exterior' which irrupts into a discursive formation and constitutes it as a site of discursive obviousness subject to the law of closed structural repetition. The closure of the machine is therefore retained, but it is now seen as the paradoxical result of the irruption of an external and pre-existing 'elsewhere'.

It follows that the subject of discourse is still seen as a pure effect of subjection to the machinery of the discursive formation with which the subject identifies. At the level of DA-2, the question of the 'subject of the

---

<sup>1</sup> References on AD-1: *Langages* 11 (1968), 13 (1969) and 23 (1971), Pêcheux (1969a).

enunciation' can be raised only in terms of the illusion of the 'I' that results from subjection. (Cf. the Althusserian problematic of ideological state apparatuses, which is haunted by the Spinozist theme of the subjective illusion that arises from 'our ignorance of the causes that determine us'.)

But, because it posits a relationship of unequal intrication between a discursive formation and its exterior, the DA-2 problematic can at the same time identify points of polemical conflict on the internal boundaries of a discursive formation. These zones are traversed by a whole series of discursive effects which are thematized as effects of ideological ambiguity, division, and 'strategic' replies and rejoinders. Ultimately, this problematic gives rise to the idea of a sort of discursive vacillation which affects the sequences situated on the boundaries of a discursive formation to such an extent that it proves impossible to determine precisely which discursive formation generates them.

The insistence of alterity within discursive identity thus represents a challenge to the notion of the closure of identity, to the notion of structural discursive machinery and perhaps to the notion of the discursive formation itself. In procedural terms, however, DA-2 makes few innovations. The displacement is visible primarily at the level of the *construction* of discursive corpora. It facilitates systematic work on the uneven internal hold they have over one another, and allows us to go beyond the level of contrasted juxtaposition.<sup>1</sup>

### **Emergence of new DA-procedures through the deconstruction of discursive machineries: DA-3**

It would be futile to attempt to describe current developments as an object. Within the limitations of this chapter we can do no more than indicate certain trends within the process of the investigation-negation-deconstruction of the notions called into play by discourse analysis, and reveal a few fragments of the new constructions.

---

<sup>1</sup> References on AD-2: Robin (1973), Pêcheux (1975 (=1982a)), Henry (1977), Guilhaumou and Maldidier (1979), *Langages* 37 (1975), 55 (1979) and 62 (1981).



*Some reference points ...*

(1) The theoretical primacy of the *other* over the *same* is so accentuated as to problematize the notion of a structural discursive machine. This is in fact a precondition for the construction of new algorithms or 'paradoxical machines'.

(2) The procedure of using a set order of stages is finally exploded

- by the destabilization of the socio-historical guarantees which were assumed to ensure *a priori* the theoretical and procedural pertinence of the empirical construction of a corpus reflecting those guarantees; and
- by the cumulative interaction between alternating moments of *linguistic analysis* (their main feature being the use of a syntactic surface analyzer<sup>1</sup>) and moments of *discursive analysis* ('vertical' paradigmatic algorithms and 'horizontal' syntagmatic/sequential algorithms). This interaction is a reflection of procedures designed to take into account the incessant discursive destabilization of the 'body' of syntactic rules<sup>2</sup> and 'obvious' forms of sequentiality (e.g. narrative, descriptive or argumentative forms). It presupposes the *reinscription*<sup>3</sup> of traces of these partial analyses within *the discursive field that is being analyzed as a corpus*, gives that field a new configuration, and opens it up to a new phase of linguistic-discursive analysis. The spiralling production of new configurations of the corpus punctuates the process by producing a series of *interpretations* of the field under analysis. What is the necessary position of 'the same' within an analytic process of this kind?

(3) At the level of DA-1, the divorce between linguistic analysis (of each sequence) and intersequential discursive analysis (of a corpus of sequences) made the notion of the discursive analysis of a sequence in its singularity meaningless. And yet a linguistic analysis of the DA-1 type implicitly assumed that every sequence analyzed was homogeneous at the

---

<sup>1</sup> I refer to the surface grammar developed with DÉREDEC software, and perfected by Plante (1983) (cf. Chapter 8).

<sup>2</sup> By, for example, shifts in the frontiers between component elements, lexico-semantic displacements of the acceptability of constructions, or grammatical ambiguities (as to the status of the infinitive, for example).

<sup>3</sup> With DÉREDEC, this reinscription can be carried out by constructing EXFALS: "networks of atomic expressions connected by oriented relations whose depth and complexity are not subject to any formal limits" (Plante 1983).

level of enunciation, because the register of enunciation and of the constraints of sequentiality remained opaque to it. Current developments in the many research projects being carried out on intradiscursive or 'inter-sentence' sequences allow DA-3 to begin to study both the *construction* of discursive objects and events, and the *construction* of 'points of view' and sites of enunciation within the intradiscursive thread.

At the same time, certain theoretical developments pertaining to the question of the heterogeneity of enunciation lead to the thematization of the linguistic-discursive forms of the *discourse of the other*:

- the discourse of an 'other' staged by the subject, or the discourse of the subject staged as 'other' (cf. the various forms of 'conspicuous heterogeneity'; see Authier-Revuz 1982);

- and, and this is more important, the insistence of an interdiscursive 'elsewhere' which both *structures* this 'staging' at a level which escapes the functional self-control exercised by the 'I' or by the enunciator-strategist who stages the sequence (it structures it at the identity points where the 'I' takes up its abode), and *destabilizes* it (at the points where the subject drifts into the other and loses his strategic control over his discourse).

*... and a lot of questions*

(1) How, within what we are still calling the 'subject of enunciation' can we separate the functional position of the 'I', of the subjected strategist (the active intentional subject theorized by phenomenology), from the emergence of a *subject-position*? What is the paradoxical relationship between its emergence and the stumbling block of the unexpected irruption of an 'other' discourse, the crack in the subject's self-mastery? Could it be that the 'subject' is something which intermittently emerges at the point where the 'I' vacillates? How can we inscribe the implications of this question in the concrete procedures of analysis?

(2) If discourse analysis claims to be a (new) way of 'reading' written and oral materialities, what is the new relationship that has to be constructed between reading, interlocution, memory and thought? Why is it that texts and oral sequences can, at a specific given moment, intersect, converge or diverge? How can we use these intersections, convergences and divergences to reconstruct the *memory-space* of a socio-historical body of discursive traces traversed by heterogeneous divisions, breaks and

contradictions? How can such an *interdiscursive body of traces* be inscribed *across* a language, that is, not only by it, but also in it? If a thought-process confronting a 'theme' from a certain point of view is a *position within a network of questions*, how is that position inscribed within both the figures of conversational 'exchange' (which range from dialogue to altercation and which include all the intervening forms of conflict) and in the figures which establish perspective; how can it be inscribed therein as a gesture which structures a field of readings (identification of relations or 'thematic trajectories' which convene heterogeneous textual series)? What is it that *brings together* a space of interlocution, a memory-space and a network of questions?

(3) How can we conceive of a *discourse analytic process* as a 'spiral' interaction which combines intersections, convergences and divergences between textual series (oral or written), constructions of networks of questions, structurations of memory-networks and productions of writing?

How can writing punctuate this process by producing an *interpretation-effect* within it?

How does the *reading subject* emerge from within this punctuation?

What does *interpretation* mean within this process?

On what conditions is it possible (or impossible) for an interpretation to make an *intervention*? Can we (re-)define a 'politics' of discourse analysis?<sup>1</sup>

---

<sup>1</sup> References on AD-3: Concein et al. (1981), Plante (1983).



# Bibliography

- Achard, P. (1991), *Compte rendu de D. Maldidier, ed., L'inquiétude du discours, Langage et société* 56, 101-104.
- Achard, P., M.-P. Gruenais and D. Jaulin, eds. (1984), *Histoire et linguistique*. Paris: Maison des Sciences de l'Homme.
- Ackermann, W., and R. Zygouris (1966), 'Représentation et assimilation de connaissances'. Paris: Centre d'Études et Recherches Psychotechniques (C.E.R.P.) (unpublished).
- Althusser, L. (1970), 'From Capital to Marx's philosophy'. In: L. Althusser and E. Balibar, *Reading Capital*. London: New Left Books, 11-69.
- Althusser, L. (1971a), 'Ideology and ideological state apparatuses'. In: L. Althusser, *Lenin and Philosophy*. London: New Left Books, 127-186.
- Althusser, L. (1971b), 'Freud and Lacan'. In: L. Althusser, *Lenin and Philosophy*. London: New Left Books, 189-219.
- Althusser, L. (1976), 'Elements of self-criticism'. In: L. Althusser, *Essays in Self-Criticism*. London: New Left Books, 101-150.
- Althusser, L. (1990), 'Philosophy and the spontaneous philosophy of the scientists (1967)'. In: L. Althusser, *Philosophy and the Spontaneous Philosophy of the Scientists*, London: Verso, 69-165.
- Althusser, L., and E. Balibar (1970), *Reading Capital*. London: New Left Books.
- Austin, J.L. (1962), *How to Do Things with Words*. Oxford: Clarendon Press.
- Authier-Revuz, J. (1982), 'Hétérogénéité montrée et hétérogénéité constitutive. Éléments pour une approche de l'autre dans le discours', *DRLAV* 26, 91-151.
- Bachelard, G. (1949), *Le rationalisme appliqué*. Paris: Presses Universitaires de France.
- Bachelard, G. (1984), *The New Scientific Spirit*. Boston: Beacon Press.
- Badiou, A. (1969), *Le concept de modèle. Introduction à une épistémologie matérialiste des mathématiques*. Paris: Maspéro.
- Barthes, R. (1967), *Elements of Semiology*. London: Cape.
- Bègue, D. (1983), 'À propos d'un emprunt historique de la grammaire générative à la théorie des langages formels', *Archives et Documents de la Société d'Histoire et d'Épistémologie des Sciences du Langage* 3.
- Bellert, I. (1970), 'On a condition of the coherence of texts', *Semiotica* 2, 335-363.
- Benveniste, E. (1966), *Problèmes de linguistique générale*. Paris: Gallimard.
- Berge, C. (1958), *La théorie des graphes et ses applications*, Paris: Dunod.
- Borel, M.J. (1970), 'Pour définir l'argumentation', *Travaux du Centre de recherches sémiologiques de l'Université de Neuchâtel* 3, 1-52.
- Borillo, M., and J. Virbel (1973), 'Remarques méthodologiques à propos des "Recherches sur le discours illuministe au XVIII<sup>e</sup> siècle" de G. Gayot et M. Pêcheux' (unpublished).
- Bourdieu, P. (1990), *Photography. A Middle-brow Art*. Cambridge, Polity Press.
- Calvet, L.J. (1975), *Pour et contre Saussure. Vers une linguistique sociale*. Paris: Payot.
- Canguilhem, G. (1977), *Idéologie de rationalité dans l'histoire des sciences de la vie*. Paris: Vrin.

- Cartwright, D.P. (1953), 'Analysis of qualitative material'. In: L. Festinger and D. Katz, *Research Methods in the Behavioral Sciences*. New York: Holt, Rinehart and Winston, 421-470.
- Chomsky, N. (1957), *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1965), *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1969), *Structures syntaxiques*. Paris: Seuil.
- Chomsky, N., and G.A. Miller (1963), 'Introduction to the formal analysis of natural languages'. In: R.D. Luce, R.R. Bush and E. Galanter, eds., *Handbook of Mathematical Psychology* 2. New York and London: John Wiley and Sons, 269-321.
- Cipolli, C. (1972), 'Considerazioni teorico-metodologiche sull'analisi del discorso. A proposito del methodo AAD di M. Pêcheux', *Lingua e stile* 1, 199-229.
- Compagnon, A. (1979), *La seconde main ou Le travail de la citation*. Paris: Seuil.
- Conein, B., J.-J. Courtine, F. Gadet, J.-M. Marandin and M. Pêcheux (1981), *Matérialités discursives*. Lille: Presses Universitaires de Lille.
- Courtine, J.-J. (1981), 'Quelques problèmes théoriques et méthodologiques an analyse du discours. À propos du discours communiste adressé aux chrétiens', *Langages* 62, 9-128.
- Cousins, M. (1985), 'Jokes and their relation to the mode of production', *Economy and Society* 14, 94-112.
- Cros, R.-C., J.-C. Gardin and F. Lévy (1964), *L'automatisation des recherches documentaires. Un modèle général 'LE SYNTOL'*. Paris: Gauthier-Villars.
- Culioli, A. (1968), 'La formalisation en linguistique', *Cahiers pour l'analyse* 9, 106-117. - Also in: A. Culioli, C. Fuchs and M. Pêcheux (1970), *Considérations théoriques à propos du traitement formel du langage*. Paris: Dunod, 1-13.
- Culioli, A., C. Fuchs and M. Pêcheux (1970), *Considérations théoriques à propos du traitement formel du langage*. Paris: Dunod.
- Del Vigna, C., and M. Dupraz (1974), 'Recherches liées à l'analyse automatique du discours', *Travaux du Centre de recherches sémiologiques de l'Université de Neuchâtel* 19, 23 ff.
- Derrida, J. (1970), 'Structure, sign, and play in the discourse of the human sciences'. In: R. Macksey and E. Donato, eds., *The Structuralist Controversy. The Languages of Criticism and the Sciences of Man*. Baltimore and London: Johns Hopkins, 247-265.
- Derrida, J., and J. Mosconi (1975), *Lévi-Strauss dans le dix-huitième siècle. Cahiers pour l'analyse* 4.
- Doležel, L. (1964), 'Un modèle statistique de codage linguistique', *Études de linguistique appliquée* 3, 51-63.
- Dubois, J. (1962), *Le vocabulaire politique et social en France de 1869 à 1872*. Paris: Larousse.
- Dubois, J. (1969), 'Énoncé et énonciation', *Langages* 13, 100-110.
- Dubois, J. (1970), 'Compte rendu de M. Pêcheux, Analyse automatique du discours', *Journal de psychologie normale et pathologique* 3, 359-360.
- Ducrot, O. (1966), 'Logique et linguistique', *Langages* 2, 3-30.
- Dupraz, M. (1972), 'Caractéristiques du programme AADP d'analyse automatique du discours écrit en ALGOL W d'après la méthode de M. Pêcheux'. Université des sciences sociales de Grenoble (unpublished).

- Engler, R. (1967-1974), *Édition critique du Cours de linguistique générale de Ferdinand de Saussure*, Wiesbaden: Harrassowitz.
- Fauconnet, P., and M. Mauss (1901), 'Sociologie'. In: *Grande Encyclopédie* 30, Paris: Société anonyme de la Grande Encyclopédie, 165-176. Reprinted in: M. Mauss (1969), *Oeuvres* 3, Paris: Minuit, 139-177.
- Fiala, P., and C. Ridoux (1973), 'Essai de pratique sémiotique', *Travaux du Centre de recherches sémiologiques de l'Université de Neuchâtel* 17.
- Fichant, M., and M. Pêcheux, eds. (1969), *Sur l'histoire des sciences*, Paris: Maspero.
- Fisher, S., and E. Veron (1973), 'Baranne est une crème', *Communications* 20, 162-181.
- Foucault, M. (1970), *The Order of Things. An Archaeology of the Human Sciences*, London: Tavistock.
- Foucault, M. (1971), 'Orders of discourse', *Social Science Information* 10 (2), 7-30.
- Foucault, M. (1972), *The Archaeology of Knowledge*, London: Tavistock.
- Foucault, M. (1977), *Discipline and Punish. The Birth of the Prison*, London: Lane, and New York: Pantheon.
- Foucault, M. (1980), *The History of Sexuality* I, New York: Vintage.
- Freud, S. (1971), *The Complete Introductory Lectures on Psychoanalysis*, London: Allen and Unwin.
- Fuchs, C. (1979), 'Référentiation et paraphrase. Variation sur une valeur aspectuelle', *DRLAV* 21.
- Fuchs, C., C. Haroche, P. Henry, J. Léon and M. Pêcheux (1972), 'Premiers éléments d'un analyseur morpho-syntaxique du français', Paris: C.N.R.S., E.P.H.E. (unpublished).
- Fuchs, C., and J. Milner (in collaboration with P. le Goffic) (1974), 'À propos des relatives' (unpublished).
- Gadet, F. (1989), *Saussure and Contemporary Culture*, London: Hutchinson.
- Gadet, F., and T. Hak, eds. (1990), *Por uma análise automática do discurso. Uma introdução à obra de Michel Pêcheux*, Campinas: Unicamp.
- Gadet, F., and M. Pêcheux (1981), *La langue introuvable*, Paris: Maspero.
- Gadet, F., J. Léon and M. Pêcheux (1984), 'Remarques sur la stabilité d'une construction linguistique: la complétive', *LINX* 10, 23-50.
- Gayot, G. (1973), 'Discours fraternel et discours polémique'. In: R. Robin, ed., *Histoire et linguistique*, Paris: Colin, 229-244.
- Gayot, G., and M. Pêcheux (1971), 'Recherches sur le discours illuministe au XVIII<sup>e</sup> siècle. Louis-Clause de Saint-Martin et les "circonstances"', *Annales* 3-4, 681-704.
- Geffroy, A., P. Lafon, G. Seidel and M. Tourmier (1973), 'Lexicomatic analysis of co-occurrences'. In: J. Aitken, R.W. Bailey, and N. Hamilton-Smith, eds., *The Computer and Literary Studies*, Edinburgh: Edinburgh University Press, 113-133.
- Ghiglione, R., and A. Blanchet (1991), *Analyse de contenu et contenus d'analyses*, Paris: Dunod.
- Gochel, P. (1972), *Esquisse d'une théorie nominaliste de la proposition*, Paris: Colin.
- Godel, R. (1957), *Les sources manuscrites du Cours de linguistique générale de Ferdinand de Saussure*, Geneva: Droz.
- Greimas, A.J., ed. (1968), 'Pratiques et langages gestuels', *Langages* 10.
- Grésillon, A. (1975), 'Les relatives dans l'analyse linguistique de la surface textuelle. Un cas de région-frontière', *Langages* 37, 99-118.
- Guespin, L. (1971), 'Problématique des travaux sur le discours politique', *Langages* 23, 3-24.

- Guilhaumou, J. and D. Mالدیدیر (1979), 'Courte critique pour une longue histoire', *Dialectiques* 26: 7-23.
- Hak, T., and N. Helsloot, eds. (1991), *De taal kan barsten. Spanning tussen taalkunde en maatschappijwetenschap*, Amsterdam: Krisis.
- Halliday, M.A.K. (1967), 'Notes on transitivity and theme in English (2)', *Journal of Linguistics* 3, 177-274.
- Haroche, C., and M. Pêcheux (1972a), 'Manuel pour l'utilisation de la méthode d'analyse automatique du discours', *T.A. Informations* 13 (1), 13-55.
- Haroche, C., and M. Pêcheux (1972b), 'Facteurs socio-économiques et résolution des problèmes', *Bulletin du Centre d'Études et Recherches Psychotechniques (C.E.R.P.)* 21 (2-3), 67-83.
- Haroche, C., P. Henry and M. Pêcheux (1971), 'La sémantique et la coupure saussurienne. Langue, langage, discours', *Langages* 24, 93-106.
- Harris, Z.S. (1952), 'Discourse analysis', *Language* 28, 1-30. Reprinted in: Z.S. Harris (1970), *Papers in Structural and Transformational Linguistics*. Dordrecht: Reidel, 313-348.
- Harris, Z.S. (1957), 'Co-occurrence and transformation in linguistic structure'. In: Z.S. Harris (1970), *Papers in Structural and Transformational Linguistics*. Dordrecht: Reidel, 390-457.
- Harris, Z.S. (1970), *Papers in Structural and Transformational Linguistics*. Dordrecht: Reidel.
- Helsloot, N., and T. Hak (1992), Review of D. Mالدیدیر, ed., *L'inquiétude du discours*, *Journal of Pragmatics* 16, 372-376.
- Henry, P. (1971), 'On processing of message referents in contexts'. In: E.A. Carswell and R. Rommetveit, eds., *Social Contexts of Messages*. London: Academic Press, 77-95.
- Henry, P. (1974), 'De l'énoncé au discours: présupposition et processus discursifs'. Paris: C.N.R.S., E.P.H.E. (unpublished).
- Henry, P. (1975), 'Constructions relatives et articulations discursives', *Langages* 37, 81-98.
- Henry, P. (1977), *Le mauvais outil*. Paris, Klincksieck.
- Herbert, T. (1966), 'Réflexions sur la situation théorique des sciences sociales et, spécialement, de la psychologie sociale', *Cahiers pour l'analyse* 2, 141-167.
- Herbert, T. (1968), 'Remarques pour une théorie générale des idéologies'. *Cahiers pour l'analyse* 9, 74-92.
- Hirsbrunner, M., and P. Fiala (1972), 'Les limites d'une théorie saussurienne du discours et leurs effets dans la recherche sur l'argumentation', *Travaux du Centre des recherches sémiologiques de l'Université de Neuchâtel* 13.
- Houdebine, Jean-Louis (1976), 'Les Vérités de La Palice ou les erreurs de la police (d'une question obstinément forclosée)', *Tel Quel* 67: 87-97.
- Irigaray, L. (1985), 'Négation et transformation négative dans le langage des schizophrènes'. In: L. Irigaray, *Parler n'est jamais neutre*. Paris: Minuit, 35-53.
- Jakobson, R. (1971), *Selected Writings* 2. The Hague and Paris: Mouton.
- Jakobson, R. (1981), *Selected Writings* 3. The Hague and Paris: Mouton.
- Katz, J.J., and J.A. Fodor (1964), 'The structure of a semantic theory'. In: J.A. Fodor and J.J. Katz, eds., *The Structure of Language. Readings in the Philosophy of Language*. Englewood Cliffs NJ: Prentice-Hall, 479-518.



- Krippendorff, K. (1969), 'Introduction to part I'. In: G. Gerbner et al., eds., *The Analysis of Communication Content. Developments in Scientific Theories and Computer Techniques*. New York and London: John Wiley and Sons, 3-16.
- Kundera, M. (1986), *Life is Elsewhere*. London and Boston: Faber and Faber.
- Lacan, J. (1966), *Écrits*. Paris: Seuil.
- Lacan, J. (1977), *Écrits. A Selection*. New York: Norton.
- Lakoff, G. (1971), 'Presupposition and relative well-formedness'. In: D.D. Steinberg and L.A. Jakobovits, eds., *Semantics*. Cambridge, MA: Cambridge University Press, 329-340.
- Lecomte, A. (1981), 'Comment Einstein raconte comment Newton expliquait la lumière', *Revue européenne des sciences sociales* 56, 69-93.
- Lecomte, A., and J.-M. Marandin (1986), 'Analyse du discours et morphologie discursive'. In: M. Charolles, J.S. Petőfi and E. Sözer, eds., *Research in Text Connexity and Text Coherence*. Hamburg: Buske, 61-100.
- Lecourt, D. (1972), *Pour une critique de l'épistémologie (Bachelard, Canguilhem, Foucault)*. Paris: Maspéro.
- Leeman, D. (1973a), 'Distributionnalisme et structuralisme', *Langages* 29, 6-42.
- Leeman, D. (1973b), 'Les paraphrases', *Langages* 29, 43-54.
- Leeman, D. (1973c), 'La paraphrase dans l'oeuvre de Hiz', *Langages* 29, 82-89.
- Lepschy, G. (1966), *La linguistica strutturale*. Turin: Einaudi.
- Lerman, I. (1970), *Les bases de la classification automatique*. Paris: Gauthier-Villars.
- Lévi-Strauss, C. (1966), *The Savage Mind*. Chicago: University of Chicago Press.
- Lévi-Strauss, C. (1963), 'The structural study of myth'. In: C. Lévi-Strauss, *Structural Anthropology*. Garden City NY: Anchor, 202-228.
- Lévi-Strauss, C. (1969), *The Raw and The Cooked*. New York: Harper and Row.
- MacCabe, C. (1979), 'On discourse', *Economy and Society* 8, 279-307.
- Macdonell, D. (1986), *Theories of Discourse*. Oxford: Blackwell.
- Macherey, P. (1966), *A Theory of Literary Production*. London: New Left Books.
- Maingueneau, D. (1991), *L'analyse du discours*. Paris: Hachette.
- Maldidier, D. (1984), 'Michel Pêcheux. Une tension passionnée entre la langue et l'histoire'. In: P. Achard, M.-P. Gruenais and D. Jaulin, eds., *Histoire et linguistique*. Paris: Maison des Sciences de l'Homme, xi-xiv.
- Maldidier, D., ed. (1990), *L'inquiétude du discours. Textes de Michel Pêcheux*. Paris: Cendres.
- Marandin, J.-M. (1983), 'Approche morphologique du discours' (unpublished).
- Marinetti, A., and M. Meli, eds. (1986), *Ferdinand de Saussure. Le leggende germaniche*. Este: Zielo.
- Martinet, A. (1962), *Éléments de linguistique générale*. Paris: Colin.
- Mauro, T. de, (1972), 'Notes'. In: F. de Saussure, *Cours de linguistique générale*. Paris: Payot, 405-477.
- Morin, M. (1973), *Actes, discours et formation*. Université d'Aix-en-Provence (dissertation).
- Moscovici, S., and M. Plon (1966), 'Les situations-colloques. Observations théoriques et expérimentales', *Bulletin de psychologie* 19, 702-722.
- Mounin, G. (1963), *Les problèmes théoriques de la traduction*. Paris: Gallimard.
- Mounin, G. (1968), *Saussure ou le structuraliste sans le savoir*. Paris: Seghers.
- Normand, C. (1970), 'Propositions et notes en vue d'une lecture de F. de Saussure', *La Pensée* 154, 34-51.

- Normand, C. (1973), 'L'arbitraire du signe comme phénomène de déplacement', *Dialectiques* 1-2, 109-126.
- Normand, C. (1978), 'Langue/parole: constitution et enjeu d'une opposition', *Langages* 49, 66-90.
- Normand, C. (1985), 'Le sujet dans la langue', *Langages* 77, 7-19.
- Ogilvie, B. (1987), *Lacan. La formation du concept de sujet (1932-1949)*. Paris: Presses Universitaires de France.
- Pagès, R. (1955), 'Images de l'émetteur et du récepteur dans la communication', *Bulletin de psychologie* 8 (7-8), 452-460.
- Pêcheux, M. (1967), 'Analyse de contenu et théorie du discours', *Bulletin du Centre d'Études et Recherches Psychotechniques (C.E.R.P.)* 16 (3), 211-227.
- Pêcheux, M. (1968), 'Vers une technique d'analyse du discours', *Psychologie française* 13 (1), 113-117.
- Pêcheux, M. (1969a), *Analyse automatique du discours*. Paris: Dunod.
- Pêcheux, M. (1969b), 'Sur la conjoncture théorique de la psychologie sociale', *Bulletin de psychologie* 23 (4-5): 290-297.
- Pêcheux, M. (1969c), 'Idéologie et histoire des sciences'. In: M. Fichant and M. Pêcheux, eds., *Sur l'histoire des sciences*. Paris: Maspero, 13-47.
- Pêcheux, M. (1969d), 'Les sciences humaines et le "moment actuel"', *La Pensée* 143, 62-79.
- Pêcheux, M. (1973), 'L'application des concepts de la linguistique à l'amélioration des techniques d'analyse de contenu', *Ethnies* 3, 101-118.
- Pêcheux, M. (1975), *Les Vérités de La Police*. Paris: Maspero.
- Pêcheux, M. (1978), *Hacia el análisis automático del discurso*. Madrid: Gredos.
- Pêcheux, M. (1981), 'Ouverture du colloque'. In: B. Conein, J.-J. Cour-tine, F. Gadet, J.-M. Marandin and M. Pêcheux, eds., *Matérialités discursives*. Lille, Presses Universitaires de Lille, 15-18.
- Pêcheux, M. (1982a), *Language, Semantics and Ideology*. London and Basingstoke: Mac-Millan.
- Pêcheux, M. (1982b), 'Lire l'archive aujourd'hui', *Archives et Documents de la Société d'Histoire et d'Épistémologie des Sciences du Langage* 2, 35-45.
- Pêcheux, M. (1984), 'Matériel en vue de l'article "Complétives/Infinitifs/Infinitives"', *LINX* 10, 7-22.
- Pêcheux, M. and E. Balibar (1969), 'Définitions'. In: M. Fichant and M. Pêcheux, eds., *Sur l'histoire des sciences*. Paris: Maspero, 8-12.
- Pêcheux, M., and C. Fuchs (1975), 'Mises au point et perspectives à propos de l'analyse du discours', *Langages* 37, 7-80.
- Pêcheux, M., and J. Wesselius (1973), 'À propos du mouvement étudiant et des luttes de la classe ouvrière: 3 organisations étudiantes en 1968'. In: R. Robin, ed., *Histoire et linguistique*. Paris: Colin, 245-260.
- Pêcheux, M., C. Haroche, P. Henry and J.P. Poitou (1979), 'Le rapport Mansholt: un cas d'ambiguïté idéologique', *Technologies, Idéologies, Pratiques* 2, 1-83.
- Plante, P. (1981), *DÉREDEC. Manuel de l'usager*. Montréal: Université de Québec à Montréal.
- Plante, P. (1983), 'Le Système de programmation DÉREDEC', *Mots* 6, 101-134.
- Plante, P. (1984), *Manuel de programmation*. Montréal: Université de Québec à Montréal.
- Politzer, G. (1928), *Critique des fondements de la psychologie*. Paris: Rieder.
- Propp, V. (1968), *Morphology of the Folktale*. Austin: University of Texas Press.

- Provost-Chauveau, G. (1970), *Compte rendu de M. Pêcheux, Analyse automatique du discours*, *La Pensée* 151, 135-138.
- Robin, R., ed. (1973), *Histoire et linguistique*. Paris: Colin.
- Roudinesco, E. (1990), *Jacques Lacan & Co. A History of Psychoanalysis in France, 1925-1985*. Chicago: University of Chicago Press.
- Ruwet, N., (1968), *Introduction à la grammaire générative*. Paris: Plon.
- Safouan, M. (1973), *Le structuralisme en psychanalyse*. Paris: Seuil.
- Saussure, F. de (1983), *Course in General Linguistics*. London: Duckworth.
- Slakta, D. (1974), 'Essai pour Austin', *Langue Française* 21, 90-105.
- Sola-Pool, I. de (1959), *Trends in Content Analysis*. Urbana: University of Illinois Press.
- Starobinski, J. (1979), *Words upon Words. The Anagrams of Ferdinand de Saussure*. New Haven, CT: Yale University Press.
- Stone, P.J., D.C. Dunphy, M.S. Smith, and D.M. Ogilvie (1966), *The General Inquirer. A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.
- Thompson, J.B. (1983), 'Ideology and the analysis of discourse. A critical introduction to the work of Michel Pêcheux', *Sociological Review* 31, 212-236.
- Todorov, T., ed. (1965), *Théorie de la littérature. Textes des formalistes russes réunis*. Paris: Seuil.
- Tedorov, T., ed. (1966), *Recherches sémantiques. Languages* 1.
- Todorov, T. (1984), *Mikhail Bakhtin. The Dialogical Principle*. Manchester: Manchester University Press.
- Trognon, A. (1972), *Analyse de contenu et théorie de la signification*. Université Paris VII (dissertation).
- Ullmann, S. (1951), *The Principles of Semantics*. Glasgow: Jackson, and Oxford: Blackwell.
- Ullmann, S. (1952), *Précis de sémantique française*. Bern: Francke.
- Wahl, F., ed. (1968), *Qu'est-ce que le structuralisme?*. Paris: Seuil.
- Williams, G. (1992), *Sociolinguistics. A sociological critique*. London: Routledge.
- Woetzel, H., and M. Geier (1982), 'Sprachtheorie und Diskursanalyse in Frankreich', *Das Argument* 133, 386-399.
- Wolf, F.O. (1985), 'Nachwort'. In: L. Althusser, *Philosophie und spontane Philosophie der Wissenschaftler*. Berlin: Argument, 151-171.
- Woods, W.A. (1970), 'Transition network grammars for natural language analysis', *Communications of the ACM* 13, 591-606.
- Woods, R. (1977), 'Discourse analysis: the work of Michel Pêcheux', *Ideology and Consciousness* 2, 57-79.
- Zolkovskij, A.K., and I.A. Mel'cuk (1971), 'Construction d'un modèle actif de langue: sens-texte', *La Sémantique en USSR. Documents de linguistique quantitative* 10, 11-50.



# Name index

- Achard 3  
 Ackermann 71  
 Althusser 4-9, 13, 17, 20, 22, 26,  
 32-33, 35-39, 41-44, 47, 55,  
 125-127, 143, 189, 190, 238  
 Aristotle 120  
 Austin 80  
 Authier-Revuz 240  
 Bachelard 4-9, 16, 18, 23, 26  
 Badiou 7  
 Balibar 5-7, 35  
 Bally 10, 137  
 Barthes 50, 69, 189  
 Bègue 216  
 Bellert 230  
 Benveniste 19, 49-50, 54, 80, 102,  
 106, 111, 136-137, 184-185  
 Berge 57  
 Binet 28  
 Bizard 58  
 Blanchet 3  
 Bonnaïfous 192  
 Borel 132  
 Borillo 123, 145, 157, 158, 167  
 Bourdieu 79  
 Bresson 42  
 Calvet 50  
 Canguilhem 4-6, 9, 23, 26, 28, 120  
 Cartwright 66-67  
 Chomsky 8, 19, 49, 51-52, 57, 96,  
 100, 131, 142, 153, 183-184,  
 225  
 Cipolli 3, 123  
 Clérambault 41  
 Compagnon 213, 215  
 Concein 53, 241  
 Courtine 209  
 Cousins 4, 18  
 Cros 58, 71  
 Culioli 19, 49, 54-55, 135, 140  
 Del Vigna 58, 123, 162  
 Derrida 4, 11, 33-35, 37-39  
 Dubois 55, 77, 123, 190, 210  
 Ducrot 47, 80, 88-89, 230  
 Duhem 23  
 Dulong 220  
 Dupraz 58, 123, 162, 174  
 Durkheim 79  
 Duval 58, 196  
 Engler 10, 50  
 Fauconnet 79  
 Fiala 136-137, 145  
 Fichant 7  
 Fisher 16, 123, 131, 138, 157-159,  
 166  
 Fodor 183  
 Foucault 4-6, 11, 23, 32-35, 37-40,  
 189, 237  
 Frege 18  
 Freud 22, 32, 36, 41-44, 46, 139-140,  
 190  
 Fuchs 17, 54-55, 58, 134, 137, 140,  
 211  
 Gadet 4, 18-19, 40, 49, 52  
 Galileo 6, 24  
 Gardin 58, 70-71  
 Gayot 16, 123, 145-147  
 Geffroy 146  
 Geier 14, 52  
 Gochet 176  
 Godel 10, 50-51  
 Greimas 81  
 Grésillon 134  
 Gruenais 3  
 Guespin 132  
 Guilhaumou 238  
 Guiraud 77  
 Hak 3-4  
 Halliday 225  
 Haroche 16, 49, 76, 96, 123, 125-127,  
 133, 139, 153, 165  
 Harris 14, 19, 49, 52-53, 55, 98, 153,  
 161, 169, 175-176, 178-180,  
 190-191, 209, 229

- Harvey 120  
 Hegel 36-37  
 Helsloot 3  
 Henry 19, 47, 49, 51, 57, 76, 89, 96,  
 125-127, 129, 134, 181, 238  
 Herbert 8-10, 17, 21-27, 29, 30,  
 35-37, 39, 43-45, 118, 139  
 Hirsbrunner 136-137  
 Irigaray 81  
 Jakobson 10, 12-13, 19, 32, 49-51,  
 53-54, 74, 76, 84, 101-102,  
 137  
 Jaulin 3  
 Kant 26  
 Katz 183  
 Krippendorff 12  
 Kristeva 50  
 Kundera 40  
 Lacan 4, 10-11, 17, 19-20, 22, 32-39,  
 41-47, 139-140  
 Lakoff 76  
 Lecomte 222, 232-233  
 Lecourt 23  
 Le Goffic 134, 137  
 Leeman 176, 178-180  
 Léon 18-19  
 Lepschy 50  
 Lerman 170  
 Lévi-Strauss 10-11, 33-34, 36, 69,  
 189, 190  
 Lévy 58, 71  
 MacCabe 4, 18  
 Macdonell 4, 18  
 Macherey 7, 43  
 Maingueneau 3  
 Maldidier 3, 19, 238  
 Marandin 222  
 Marinetti 10  
 Martinet 50  
 Marx 5, 35-37, 190  
 Matoré 77  
 Mauro 10  
 Mauss 79  
 Mel'cuk 176  
 Meli 10  
 Merton 79  
 Miller 57  
 Milner 134, 137  
 Morin 168  
 Mosconi 33  
 Moscovici 46-47, 82-83  
 Mounin 50, 71, 77  
 Nietzsche 18, 34, 190  
 Normand 50, 54  
 Ogilvie 45  
 Pagès 42, 81  
 Pizarro 58  
 Planie 59, 216-217, 219, 223-225,  
 232, 239, 241  
 Plon 19, 46-47, 82-83  
 Politzer 41  
 Propp 36, 69, 190  
 Provost-Chauveau 16, 55, 123, 146,  
 158, 168, 170, 175  
 Regnault 6-7  
 Ridoux 136, 145  
 Robin 67, 123, 137, 139, 238  
 Roudinesco 42  
 Ruwet 51, 53  
 Safouan 140  
 Saint Martin 147  
 Saussure 10-11, 13, 15, 19, 32, 39,  
 49-52, 54, 63-65, 72-76,  
 78-79, 98, 133, 136, 140, 190  
 Sechehaye 10  
 Skinner 83  
 Slakta 80  
 Sola-Pool 66  
 Spinoza 36-37, 238  
 Starobinski 10, 50-51  
 Stone 12, 58, 68  
 Thompson 4, 14  
 Todorov 50, 69, 77, 95, 223  
 Trier 77  
 Trognon 16, 123, 130, 132, 157-158,  
 167, 170  
 Ullmann 73  
 Veron 16, 123, 131, 138, 157-159,  
 166  
 Virbel 123, 145, 157-158, 167  
 Wahl 33  
 Wesseliuss 16, 123  
 Williams 4  
 Wittgenstein 18

Woetzel 14, 52  
Wolf 7  
Woods, R. 4, 18  
Woods, W.A. 216  
Zolkovskij 176  
Zygouris 71





# Subject index

- acceptability 75-76, 239
- adjunction 109, 194, 204
- adverb (ADV) 112-113, 115, 156, 194-195, 198, 222
- algebra
  - Boolean (vector) 57, 89, 112, 162-163, 168, 198
  - discursive 236
- algorithm 16, 19, 57-58, 117, 149-150, 152-153, 155, 170-172, 191-192, 196-203, 211-213, 215, 223-224, 227, 229, 239
- ambiguity
  - linguistic and discursive 130, 138, 208, 211-212, 220-222, 227, 238-239
  - theoretical 40, 44, 123, 126-127, 132, 141, 143, 148, 179
- anaphora 109, 111, 159, 193, 233-234
- antagonism 37-38, 79, 126, 141
- anthropology 11, 28-29, 32
- anticipation (expectation) 80-81, 87, 92, 101, 120, 138
- apparatus
  - ideological state 37, 47, 55, 78, 126-128, 132, 212, 238
  - conceptual 43, 57-58, 136-137, 191
  - experimental 83, 132, 143, 179
- application
  - technical vs. theoretical 7-8, 25, 30, 46, 73, 95, 149-150, 152-153
  - of ADA 117-121, 123, 147, 173, 191-203
- apposition 110, 159, 225-226
- arbitrariness 96, 121, 145-146, 160, 163, 167-168
- archaeology 37, 40, 131
- archival method 67, 147-148, 192
- arithmetic 163, 167, 170
- atomic expression (EXFAD, EXFAL) 217-219, 228-230, 231, 239
- attitude 126, 132, 185, 213
- automatic discourse analysis (ADA) 14-18, 20, 22-25, 39-40, 57, *passim*, *see also*: phase 3
- automatic translation 54, 57, 150-152
- autonomous discursive sequence 193-196, 199, 202-203
- autonomy
  - of discursive processes 147, 162, 182-183
  - of language (esp. syntax) 135, 152, 211, 225
  - of linguistics 135, 236
  - of sciences 6
  - of the subject 126
- base
  - economic 125
  - linguistic 76, 140-141, 235
- behaviour 46, 79, 81-83, 92, 126
- binary relations (BRs) 109-110, 113, 115, 155-157, 160-163, 166-168, 170, 173-174, 194-199, 202, 210
- biology 6-7, 32-33
- break
  - epistemological 5-7, 9, 11, 19, 24, 26, 32, 36, 120, 240
  - Galilean 6-7
  - Saussurean 11, 76
- class relations 23, 29-30, 37, 125-128, 147
- categories
  - basic and syntagmatic 218-221
  - morpho-syntactic (eight-position schema) 55, 109, 112-115, 156, 194, 210
- categorization 145, 158
  - sub- 217, 227-228
- chain 116, 163-164, 173, 199-200
- chemistry 27
- code 74, 76, 84, 91-93, 98, 102

- sub- 53, 76
- coding 12, 58, 66-67, 71, 107-113, 121, 158, 210-211, 228
- combination 53, 64, 74-75, 90, 96, 101-103, 107, 113-114, 146-147, 155, 173-174, 182, 235-237
- communication 12, 30-32, 47, 53, 82-84, 89, 141, 184
- community 15, 74, 76, 184
- commutation 101, 107, 172, 180-181, 186, 191, 193
- conflict 6-7, 126, 137, 227, 238, 241
- connotation 100, 178
- conscious 137-138, 189, 230, *see also*: unconscious, preconscious
- co-occurrence 66, 146, 169-170, 179, 223, 229, 231
- compaction 146, 166, 234
- comparison 27, 29, 36, 49, 65, 100, 106, 114-115, 117-119, 134, 151, 153, 161-163, 167-170, 172-174, 176, 182-183, 193, 198-199, 204, 208-211, 224, 229, 232, 236
- competence 140, 180
- compounding 90-92
- conditions of production of discourse (f')
  - 5, 12-15, 17-18, 43, 46, 52, 81-82, 85, 89-91, 78-95, 99, 105-108, 113-114, 118, 127, 128, 132, 141-144, 160, 176, 191, 212, 224, 236
  - state of 82, 87, 89-94, 96-100, 105-107
- connectors 115, 155, 160, 162, 166-167, 169, 173-174, 169, 193-194, 200, 210
- content 11, 71, 94, 98, 110, 114-115, 136, 146, 162, 168, 175-176, 203-204, 211
  - analysis 12, 19, 57, 64-67, 124, 189
- context 50, 73, 74, 76, 78, 84-85, 97-103, 108, 120, 193, 204, *see also*: dependency
  - equivalent 165, 176, 209, 229
  - maximum 196-170, 172
- coordination 109, 134, 153, 159, 160, 204, 210, 220
- co-reference 233-234
- corpus (C) 14-15, 20, 66, 99, 104-105, 107, 113-114, 133, 144, 160-161, 163-166, 172-174, 176, 181-183, 191-193, 203, 208, 224, 236, 239
  - complex 146, 148
  - closure 70, 146, 181, 224, 236-238
  - construction, *see*: phase I
- co-text 229, 231
- creativity 51, 54, 74, 106, 127, 129, 184-185
- data 16, 27-28, 66, 70, 158, 161, 163, 169, 172, 174, 186, 191-196, 198, 219, 224
- deconstruction 238
- deixis 233-234
- delinearization 134, 152, 160, 164, 191, 194, 208
- denotation 100, 177-178, 229
- dependency 91, 94, 102-103, 106, 108-110, 112, 114-117, 173, 200, 202-204, 209-210
  - contextual (RCDs) 217-222, 227
- DÉREDEC 18, 20, 49, 59, 212, 216-219, 222-224, 226, 230-234, 239
- description 10, 38, 66, 80, 82, 90, 132, 149, 190, 209-210, 213, 215-234
- desire 69-70, 135, 160
- desyntagmatization 134, 142-143, 155-156, 158, 160-167
- determination
  - (DET) 111, 113, 156, 168-169, 186, 194-195, 198, 218, 220-222, 227
  - (DT) 134, 153, 155-157, 160, 193, 202, 210
  - social/discursive 28, 52, 96, 106, 123, 125-127, 129, 132,

- 137, 139, 142-143, 184, 235, 238
- determiner 54, 185, 194, 218, 226
- dialogue 27, 89, 94, 139, 241
- dictionary 12, 143, 161, 165, 169
- discourse analysis 10, 13-14, 20, 30, 33, 39, 55, *passim*, *see also*: phase 3
- discourse theory 3, 8-11, 15-16, 18-21, 30, 36, 39, 58, 72, 95, 102, 118, 123-125, 131, 133, 136, 140, 149, 157, 181, 186, 190
- discursive process, *see*: process
- displacement
- discursive/linguistic 89, 101, 139, 239
  - Freudian 32
  - theoretical 7, 51, 64-66, 68-69, 76, 124, 141, 190, 237, 238
- dissymmetry 17, 30-31, 46, 130, 154, 165, 178, 206
- distance 87, 131, 137, 162, 166-170, 172-173, *see also*: paradigmatic proximity
- documentary analysis 58, 66, 70, 150-153
- domain, semantic 15, 52, 77, 91, 94, 106, 114-118, 130-131, 163-164, 166, 170, 172-175, 177, 181-182, 199-210
- economy 43, 85, 119, 125-128
- effect 190, 238
- dominance- 100, 105-106
  - meaning- 13, 18, 66, 84, 98, 118-120, 129, 131, 133-134, 170-171, 193
  - metaphoric 13, 17, 51, 96-101, 103, 105-106, 178
  - produce an 127, 150
  - reality- 10, 233
  - society- 45-46
  - subject- 35, 39, 124, 130-132, 134-135, 138, 182, 237, 241
  - surface 52, 95-96, 108
  - of forgetting 138, 143-144
  - without cause 37
- eight-position schema, *see*: categories
- empiricism 5, 7-8, 9-10, 19, 23-24, 27-29, 91, 136, 144, 167, 190
- empirical side of discourse analysis 4, 11, 16-17, 32, 39-40, 57, 80, 94, 99, 106-107, 132, 139, 142-143, 146, 149, 168, 184, 198, 235, 237, 239
- enunciation 53-54, 106, 111, 123, 132-139, 141, 143, 148, 159-160, 180, 182, 184-185, 210-213, 223-224, 230-231, 235-236, 238, 240
- epistemology 4-9, 19, 23, 26, 57, 125, 127
- equivalence 99, 104, 117, 154, 164-165, 169-170, 176, 180-182, 191, 193, 209, 211, 229
- classes 67, 70, 75, 175, 200, 202
- ethnography 34, 68
- experience 5, 7, 8, 11, 136, 142-143, 184
- experiment 5-8, 11, 13, 18-19, 24-25, 27-29, 83, 123, 132, 144, 147-148, 192
- expression 22, 41, 44, 46, 63, 64, 66, 67, 71-73, 86, 100, 125, 127, 129, 131, 135, 140-143, 148-149, 161, 166, 170, 178, 182, 184, 208, 217-218
- exterior (extra-discursive) 129, 139, 143-144
- specific 119, 139, 165, 224, 237-238
- extra-linguistic 75, 80, 91, 103, 113, 150, 152-153, 158, 176
- findings 14, 16, 57, 144, 158, 175-176, 183, 185, 189, 199-200, 203, 209-210, 212, 219, 224
- forgetting 23, 54, 84, 89, 124, 129, 132, 138-140, 142-143, 152, 170, 181-182
- form 231
- of utterance (F) 112-113, 115, 156, 169, 173, 194-195, 198
- formalism 8, 11-12, 14, 16, 20, 47, 57-59, 90, 107, 131, 135-136,

- 167, 170, 183, 236
- formation**
- discursive 33, 76, 88-89, 127-131, 133, 139, 141-142, 182, 210, 212-213, 237-238
  - ideological 76, 89, 126-129, 141, 143, 145
  - imaginary 85-86, 88, 94, 132
  - social 43, 79, 85, 123-126, 129
- frequency** 65, 173
- function**
- and functioning 9-10, 51, 64, 66-69, 71, 79-83, 133-137
  - functional form 58, 220
  - functional relations 66, 68, 103, 107, 110, 114
  - functionalism 78, 100, 125, 225, 240
- grammar** 135
- formal 57, 216
  - generative 8, 18, 51-52, 68, 74-76, 95-96, 153-154, 184, 222, 225, 236
  - recognition 18, 57, 150-152, 157, 219
  - surface 216, 219-222, 224-228, 234, 239
  - text-descriptive 216-217, 219
  - traditional 49, 63, 69, 148, 194
  - transformational 8, 18, 153
- grammaticality** 75, 95
- hermeneutics** 14, 130, 189, 213
- heterogeneity**
- discursive 77, 174, 182, 213, 235-236, 240-241
  - theoretical 49, 158-159, 223
- hierarchical analysis** 220-221
- historical materialism** 21-22, 38, 43-44, 123, 125, 137, 144-145, 148
- history** 5, 11, 34, 37, 123, 127-128, 134, 144, 147-149, 191-192, 206, 212, 215, 224, 239-240
- of ideas 38-39, 224
  - of science 4, 6-7, 23-24, 29, 64, 120-121
- homogeneity**
- discursive/linguistic 73, 75, 142, 144, 146, 154, 176, 181, 210, 233-234, 236, 239
  - theoretical 24, 64, 68
- hyperdomain** 164, 173-174, 200-204
- idealism** 5, 22, 36-37, 125, 127, 131, 134-135, 137, 145, 148
- identity**
- functional 98-99, 103-105, 113-115, 159, 162-163, 181, 198-199
  - lexical 115, 198, 211, 229
  - in ADA 212, 236-238
  - of meaning 129, 168, 175-176, 203, 209
  - of origin (IRG) 202
  - of the subject 45-46, 139, 237-239
- ideology** 4, 6-7, 9-12, 24, 26-32, 44-45, 76, 79-80, 118-119, 124-128, 134, 138-139, 143, 145, 189-190, 235, 238
- theory of 9, 17, 20-21, 35-41, 118, 123, 145
- illusion**
- necessary 54, 79, 137, 142
  - subjective 125, 130-132, 135-137, 139, 142, 146, 175, 182, 185, 238
- imaginary** 12-13, 45-46, 85, 88, 126, 131-132, 139, 143-144, 147
- inclusion (ICL)** 164, 173, 202
- information** 31, 46, 82, 84, 91, 164, 169, 172, 177, 185, 216-217, 227-229
- theory 84, 189
- institution** 70-73, 78-81, 132, 143, 179
- instrument**
- communicative 31-32, 111, 184
  - scientific 5, 8-9, 11, 13-16, 18-20, 23-31, 40, 44, 57, 68-69, 74, 118, 125, 216, 222
- intention** 73, 137, 146, 176, 235, 240
- interactive analysis** 58, 212, 219, 225, 239, 241
- interdiscourse** 18, 89, 139, 142, 166,

- 172, 181, 230, 237, 240
- interpellation 35, 125-128, 139, 143
- interpretation 34, 223, 230-231
  - discourse analytic, *see*: phase of interpretation
  - semantic 96, 103-105, 114, 170, 175, 183, 209
  - text 12, 63, 146
- intersection 98-99, 106, 117, 164, 173, 202, 210, 240-241
- intervention
  - linguistic 144, 149, 157-158
  - interpreter's 81, 85-86, 213, 241
  - political 8, 29, 126, 128, 139, 165, 181-182, 241
  - theoretical 8, 14, 21-22, 30-31, 118, 150, 213, 241
- intradiscourse 89, 166, 231-233, 239
- invariance 30, 53, 77, 82, 99-100, 108, 114, 118-119, 176-178, 190, 200, 235
- juridical 30, 72, 126, 128
- key word 71, 145-146, 169, 236
- knowledge
  - common 5, 31, 89, 121, 138, 224
  - scientific 5, 9, 33, 75, 102, 106, 123
  - theory of 29, 183
- language (*langage*) 8, 10, 17, 22, 30-32, 34-35, 39, 68, 72-73, 75-76, 79, 82-83, 92, 95, 111, 127, 129, 141, 182-184, 223
  - theory of 31, 52, 63-64, 113, 183
- language, the (*langue*) 10, 13, 15, 32, 38, 49-51, 64, 69, 72-79, 90-91, 93, 95-96, 113-114, 124, 133-141, 142-143, 151, 179, 184
  - artificial 98, 216
  - natural 57, 98, 133, 216-217, 235
  - universal 152, 235
- lexical analysis 96, 102-103, 107, 112-115, 133, 135, 138, 159, 173, 181, 198, 209, 211, 216-218, 223, 227-231
  - lexicology 77
  - lexicometry 12, 14, 58, 190
  - lexis 55, 112-113, 174
- linearity and non-linearity 100, 102, 111, 133, 142, 146, 152, 174, 215, 218
- linguistic analysis, *see*: phase 2
- linguistics 8, 10-11, 15, 25, 40, 49-55, 63-69, 73-75, 78, 102, 123, 133-136, 148-151, 158, 183-184, 190-191
  - discursive 137, 148-149, 184
  - structuralist 10, 32, 49-51, 95, 190, 235
  - text 149, 159
  - of speech 140, 149
- LISP 216, 234
- literary analysis 11, 50, 68, 146, 182, 189-190
- logic 25, 57, 70, 103, 106, 118, 131, 152, 165, 170, 177
- machine, discursive 25, 212, 235-238
- Marxism 5, 7-8, 17, 19, 22-23, 38, 41, 44, 47, 50, 55, 80, 125, 128
- materiality
  - discursive 80, 107, 127, 129-130, 145, 240
  - ideological 125, 127, 145
  - linguistic 127, 133, 135, 141-142, 182, 223
- mathematics 8, 25, 117, 150, 161, 167-168, 170
- meaning 11-14, 35, 39, 64-68, 77, 129-130, 135, 146, 176, 178, 182-184, 189-191
  - effect, *see*: effect
  - matrix of (metaphoric matrix) 14-17, 91, 129, 131, 142
  - relations of 80, 89, 119
  - sliding of 98, 101
- mechanism, discursive 76, 79-80, 82, 96, 132, 147, 149-150, 212, 223, 232, 237
- memory 212-213, 240-241

- metalanguage 98, 159, 170, 235
- meta-term 159, 169, 172, 194, 198
- metaphor 10, 13-14, 32, 40, 51, 99-101, 120, 165, 171, 178
- method 23-25, 39, 52, 65, 68, 95, 113, 154, 192-193, 235
- metonymy 10, 32, 51, 165, 171
- mind 33-35, 40, 152, 235
- monologue 94, 117, 139
- morpho-syntactic analysis 55, 58, 107, 109, 113-115, 134, 152, 156-157, 173, 182, 186, 194, 198
- morphology 158, 183, 213, 215, 222-223, 231
  - morphemes, analysis of 77, 95-96, 99, 138, 180, 233
- myth 36, 68-69, 70, 141, 146, 235
- nominalization 161, 185, 210, 232
- noun ( $N_1$ ,  $N_2$ ) 112, 113, 115, 156, 168, 173, 194, 195, 198, 218, 221, 226, 228, 230
- phrase ( $NP_1$ ,  $NP_2$ ) 110-112, 194, 210, 220-221, 226-228, 230
- object
  - discursive 80, 82, 142-144, 151-152, 173, 209, 229, 233, 237, 239
  - linguistic 39, 64, 68-73, 75, 107, 131, 142, 149, 159, 178
  - scientific 9-10, 24-25
- objectivity 27, 66-67, 79, 85, 94, 179, 189, 191
- obviousness 35, 39, 78, 130, 151, 182, 190-191, 237, 239
- oral language 119, 141-142, 211, 223, 240-241
- other, the 13, 17, 45-46, 81, 85, 87-89, 92-93, 126, 138-139, 181, 211, 237-238, 240
  - discourse of the 14, 18, 36, 80, 120, 138, 166, 182, 191, 213, 240
  - the Other 45, 139, 182
- overdetermination 182
- paradigmatic
  - relations 68, 161, 163-164, 200, 202, 209, 217, 229, 231-232, 239
- proximity 103-105, 114-116, 153, 162-163, 170, 173, 183, 192
- paraphrase 53, 129-131, 138, 159, 170-171, 176-183, 186, 191-192, 196, 198, 200, 202, 209, 211, 236
  - paraphrastic family 129-131, 138, 160, 176-177, 183
- parser 15, 18, 49, 212, 216, 219-220, 224-225, 228
- pertinence 67, 100, 106
- phases of discourse analysis
  - 1 (corpus construction) 14, 141-148, 160-161, 192, 212, 235, 238-239
  - 2 (linguistic analysis) 14-15, 16, 18-20, 39, 52-53, 103, 131, 133-135, 138, 142, 144, 148-153, 157-159, 164, 174, 180, 184, 186, 193, 236, 239, *see also*: syntactic analysis
  - 3 (discourse analysis) 14-16, 108, 149-150, 155, 160, 174, 186, 209, 211, 213, 222-224
  - phase of interpretation 14, 16-17, 171, 175, 181, 200, 203-204, 210, 212, 236, 239, 241
- phenomenology 88, 189, 240
- philosophy 4-8, 11, 17, 22, 25-26, 33-34, 36-38, 42-43, 49, 73, 76, 80, 131, 135-137, 189-190, 237
  - of knowledge 29
  - of science 4, 23, 25-26, 33
- phonology 133, 152, 182-183
  - phonemes, analysis of 53, 74, 102
- physics 4, 6-7, 14, 23-25, 150
- play (freeplay, game) 34, 37-38, 64, 182, 191, 215
- plurality 4, 130, 140-141, 178, 223
- politics 7-9, 27, 29-31, 35, 40, 72, 79-81, 126, 140, 192, 212, 236, 241
- power 39, 79-80, 89, 119, 237

- practice**  
 linguistic 64, 67, 69, 129, 150, 209  
 philosophical 25-26  
 reading 118, 189  
 technical 24, 27-29, 150  
 scientific 5, 7-9, 13, 15, 20, 23-27, 29, 70-71, 120, 124-125, 133, 135, 144, 149, 213  
 social/political 7, 27-28, 30-31, 35, 120, 126, 129, 148
- pragmatics** 77, 91, 230
- preconscious** 138-140, 143, 181
- presupposition** 11, 52-53, 68, 88-89, 106-107, 113, 121, 146, 149, 152, 180, 184
- process**  
 discursive 46, 76-77, 79-81, 85-88, 90, 106, 108, 119-120, 131, 133-135, 139, 141-147, 151, 158, 160, 162, 164-166, 177-178, 181-183, 191, 193, 235-237  
 dominant 106-107, 118, 120, 147  
 semantic 123, 127, 158, 175, 184, 193  
 of enunciation 106, 123, 132-134, 137-138, 141  
 of production of discourse ( $\Delta$ ) 24, 29, 77-78, 82, 90-96, 99-101, 103, 105-108, 114, 117-121, 129, 177, 235
- production, see: conditions, process**
- pronoun** 111, 159, 220-221, 226
- proposition** 108-109, 118, 131, 159, 165, 178, 184, 194, 209-211, 220-221, 229-231  
 basic 176-177, 203-204, 112, 236
- proximity, see: paradigmatic**
- psychoanalysis** 7, 17, 19-22, 32, 38, 41-45, 47, 70, 123, 137
- psycholinguistics** 127, 143
- psychology** 7, 31-32, 38, 40-42, 45-47, 67-68, 72, 82-84, 95-96, 137, 144, 183
- social** 6-9, 21-22, 31, 44-47, 72, 82-83, 148
- quadruplets** 163-164, 169, 172-173, 196, 198-200
- quantitative (vs. qualitative)** 66, 189-190, 198-199
- reading** 67-68, 118, 120, 124, 134, 147, 186, 189-191, 213, 215, 230, 240-241
- non-subjective** 131-132, 134, 144, 176
- plural** 130, 178, 190
- subjective** 12, 130
- symptomatic** 36, 118, 190
- reality** 9-10, 27, 38, 85, 92, 111, 119, 158, 160, 176, 233
- real** 45, 132
- recording** 52-53, 83, 107-113, 117, 121, 158
- reference** 13, 47, 84-88, 107-108, 111, 129, 136, 178, 184, 234
- reformulation** 36, 138-139, 186, 231-232, 234
- region** 37, 44, 72-73, 76, 123-125, 128
- relative clause** 109, 134, 153, 220
- repetition** 99, 101, 125, 161, 191, 196, 224, 234, 237
- replacement** 124
- repression** 44, 130, 139-140
- reproduction** 9, 23-25, 44-45, 49, 67, 125-128, 130, 136-137, 146-147, 157, 161, 166, 182, 185, 224, 237
- revolution, scientific** 40, 51-52, 95-96, 124
- rewriting** 15, 215
- rhetoric** 49, 77, 82, 88, 91, 101, 103, 106, 178, 212-213
- said** 38, 92, 127, 137-138  
 and unsaid 13, 119, 137, 140, 160, 181  
 earlier 13, 80, 88, 111-112, 120, 156, 213  
 elsewhere 13, 89, 119-120, 230, 237, 240
- sciences** 4-11, 16, 18, 20, 22-29, 38,

## 262 *Subject index*

- 40, 64, 70-71, 119-121, 141, 189-190
- human 8, 31-33, 37, 58, 66, 189
- life 6, 28
- natural 6, 31
- neuro- 42
- social 6, 8-11, 21-23, 25-27, 30-32, 40
- segmentation 101-102, 146-147, 160, 166, 193, 203, 209, 213, 228, 233-234
- selection
  - rules 13, 90, 95, 113-114, 137, 147, 182
  - of the corpus, *see*: phase I
- semantics 9-14, 16-18, 63, 71, 75-77, 82, 91, 95-96, 98-99, 101, 103-105, 118-119, 123, 131, 133-136, 141, 150-152, 157-158, 166, 170-171, 183-184, 229
  - discursive and formal 135, 185
  - strong and weak 176-179
- semes 96
- semiology 32, 69-70, 72-73, 78
- semiotics 69, 136
- sentence 18, 51-52, 54, 74-76, 91, 95, 102, 108, 145-146, 148, 154, 159-160, 184-185, 221, 227-228, 231-233
  - inter- 14-15, 81, 141-142, 146, 148-149, 159, 170, 180, 186, 193, 210-211, 240
- sequence
  - autonomous discursive, *see*: autonomous
  - discursive 14, 77, 84, 91-95, 99-101, 105-106, 129-131, 133, 140-142, 146, 151, 161, 164, 167-168, 172-175, 209, 213, 227-240
  - linguistic 58, 65, 82, 107-108, 113-114, 116-117, 153-157, 159-160, 182
- sign
  - 34, 38-39, 52, 68, 71-72, 81, 101-103, 106, 109, 111, 136, 184-185
  - signification 9, 31, 66, 81, 87, 96, 127, 130, 138, 152, 184, 211
  - signified 9-10, 65
  - signifier 9-10, 44-46, 65, 72, 81, 178, 182, 215
- similitude 116, 180
- situation 73, 78, 80, 85, 89, 92-94, 106, 117, 119, 132, 136, 139, 143, 145, 147-148, 182, 184, 209, 224
- social 7, 9, 23, 29-30, 35, 44, 50, 55, 64, 72-73, 78-79, 117, 119, 125-127, 136, 140-141
- sociolinguistics 50-51
- sociology 19, 31-32, 43, 58, 68, 78-79, 81-82, 85, 89, 94, 113, 117, 119-120, 125-126, 144
- speech (*parole*) 11, 32, 50-51, 54, 64, 72-76, 79, 82, 84, 106, 135, 137, 140, 149, 179, 184-185, 190, 212
  - acts 73, 80
- spontaneity 31, 69, 125, 129-131, 189-190, 213
- stability 13, 17-18, 70, 82, 113-114, 133, 136, 142, 144, 148, 151, 157, 170, 174, 178, 211, 224, 233, 235-236, 239-240
- statistics 65-66, 133, 146, 168-169, 172-173, 189-190, 198, 223
- stimulus-response 46, 82-84, 215
- strategy
  - discursive 87, 138, 238, 240
  - theoretical 4, 22, 40, 42, 190, 210, 215
- string 209, 212-213, 229, 231, 234
- structuralism 10-11, 22, 32-34, 36-37, 49-51, 53-54, 183, 189-190, 235
  - post- 4, 10-11
- structure 34, 70, 82, 85, 190-191, 235-241
  - deep 18, 52, 96, 99-100, 151-152, 177, 210
  - discursive 52, 109-110, 132



- linguistic 32, 35, 38, 67, 107, 142, 148, 174, 211, 223
- social/ideological 30, 43, 79-80, 85, 117, 119, 125, 134, 143
- surface 52, 100, 177, 190
- tree- 153, 216-219, 228
- underlying 100, 142, 151, 176, 184, 191, 203, 236
- struggle, ideological 5, 17, 29, 37, 120, 126, 128
- style 101, 119, 178, 182
- subject
  - discursive 12, 19, 38, 54, 85-86, 95, 124, 129-130, 132, 135, 137-138, 142, 161, 185, 213, 237-238, 240, *see also*: effect
  - free 53, 73-74, 126, 130, 136, 140
  - human 33-35, 183
  - ideological 7, 35-40, 45-46, 125-130, 137, 139, 235
  - linguistic 39-40, 45-46, 95, 106, 135-137, 139
  - psychoanalytic 32, 38, 44-46, 123, 137, 139, 240
  - psychological 38, 45-46, 54, 83, 95-96, 137
  - scientific 71
  - social 119-120, 136
  - speaking 12, 34, 38, 67, 69, 73-76, 79, 89, 95, 111, 129, 136-137, 139, 166, 176, 181-182, 185
  - reading 12, 124, 189, 241, *see also*: reading
  - structural 235
  - transcendental 34-35
  - universal 95-96, 183-184, 230
  - as origin 36-37, 130-131, 135, 235
  - position 12-13, 34, 38, 43, 46, 79-83, 85-86, 94, 113, 117, 119, 126-127, 132, 140, 147, 240
- subjection 125, 130-131, 134, 139, 183, 235, 237-238, 240
- subject/object noun phrase, *see*: noun phrase (NP<sub>1</sub>, NP<sub>2</sub>)
- substitution 13, 15, 53, 64, 74, 96-99, 101-105, 111, 164-166, 169-172, 174-183, 193, 211, 222, 234
- oriented 17, 165, 170-172, 178-183, 218, 239
- symmetrical 130, 165, 179
- transformation 180-181
- zone 172, 181, 186, 222, 238
- superstructure 125, 134
- surface
  - discursive 52-53, 96, 100, 105-108, 112, 117-119, 121, 142, 210, 233, 236
  - empirical 190
  - linguistic 142-144, 150, 152, 158, 160, 166, 174, 178
  - textual 66, 99, 152, 191, 194
  - effect, *see*: effect
  - grammar, *see*: grammar
  - structure, *see*: structure
- symmetry 92, 130, 165, 179
- synonymy 17, 53, 171, 176, 178, 186, 209, 229
  - contextual 98, 165, 203
  - linguistic 55, 180-181
  - non-contextual 98
- syntactic analysis 135, 153, 158-159, 170, 191, 193-194, 203, 210-212, 218-219, 222, 224-225, 228, *see also*: morpho-syntactic analysis
- syntagmatic
  - contiguity 103
  - relations 53, 68, 75, 109, 161, 163-165, 168-169, 171, 200-202, 215-222, 225, 229, 231-232, 239
- syntagmatization 165, 171-172, 180-181, 186
- syntax 77, 114, 135, 138, 152, 158, 183-184, 211-212, 217, 223, 225, 229, 231, 236
- system
  - discours as 114, 145, 154, 181, 209, 213

- social 9, 30-31, 35, 44-45, 67, 72, 79, 179
- textual 68, 190
- the language as 10-11, 15, 51, 53, 64, 66, 72-74, 76, 90, 98, 106, 114, 133, 135-136, 140, 179, 184-185
- of analysis 23, 25, 33, 36, 39-40, 71, 131, 212, 216
- of non-linguistic signs 81
- of representations 99, 121
- of rules 95, 133, 177
- technology 23, 25, 28, 71
- tense 53, 112, 136, 193-194
- text 11-15, 36, 52, 63-71, 82, 99, 111, 120, 131, 133, 145-147, 151-152, 161, 175-176, 182, 189, 193-194, 204, 213, 215-219, 222-234, 236, 240-241
- description 217-219, 222-225, 228, 231-232, 234
- exploration 216-219, 225-226, 229-231, 233
- interpretation 63, 146
- manipulation 157, 215
- textual analysis 12, 25, 58, 63-71, 149, 159
- theology 33-34
- theme 67, 189, 236, 241
  - issue (TI) 220-221, 225-226, 230-231
- transformation 7, 9, 24-25, 30, 44, 77, 85, 87, 90, 92, 108-110, 118, 120, 123-125, 127, 130, 142, 159, 174, 227, 231, 236
  - incremental 53, 179-180
  - paraphrastic 52-53, 68, 178-181
- translation 25, 57, 125, 150-152, 219
- transverse discourse 181-182, 237
- truth 34, 38, 111, 177, 230, 237
- unconscious 11, 19, 21, 32, 38, 42-45, 118, 139-140, 190, 213
- unity 53, 69, 76, 107, 146, 213
  - thematic 193
- universality 37, 74-77, 79, 95-96, 107, 110, 114, 126, 130-131, 134, 152, 178, 211, 235
- univocality 67, 88, 223
- use 7-8, 11, 26-29, 40, 50, 73-74, 136, 179, 235
- utterance (U) 13, 15, 38, 53, 101-105, 107-117, 135, 137, 140, 147-148, 154-155, 159-160, 162-164, 167-169, 173, 175-177, 194-195, 199-202, 210, 212-213, 215, 229-232
  - anomalous 75
  - canonic (stereotyped) 74-75
  - elementary (EU) 49, 52, 57, 91, 102, 108-110, 113, 153, 156, 160-162, 173-174, 191, 194
  - form of 111-112
  - matrix- 154
- value
  - cultural 6, 67, 120
  - mathematical 55, 112, 115-116, 156, 161, 163, 167-169, 172-173, 198, 209
  - Saussurean 13, 51, 89, 98
  - threshold 116, 163, 168-169, 173, 183, 198
- vector, *see*: algebra
- verb (V) 108, 112-113, 115, 150, 156, 169, 194-195, 198, 217, 220-221, 225-226, 228, 232
  - phrase (VP) 53, 110, 112, 210, 216, 218, 221-222, 226, 230
- VSL algorithm 227, 232
- weight 28, 115, 163, 168, 170, 198, 227
- writing 12, 72, 92, 119, 141-142, 215, 223, 240-241



This volume offers the long-awaited overview of the work of the French philosopher and discourse analyst Michel Pêcheux, who was the leading figure in French discourse analysis until his death in 1983. The volume presents the first English publication of the work of Pêcheux and his coworkers on automatic discourse analysis. It is presented with extensive annotations and introductions, written by former colleagues such as Françoise Gadet, Paul Henry and Denise Maldidier.

Outside France, French discourse analysis is almost exclusively known as the form of philosophical discourse presented by such authors as Michel Foucault and Jacques Derrida. The contemporary empirical forms of French discourse analysis have not reached a wider public to the degree they deserve. Through its combination of original texts, annotations, and several introductory texts, this volume facilitates an evaluation of both results and weaknesses of French discourse analysis in general and of the work of Michel Pêcheux and his coworkers in particular.