

## Popularization discourse and knowledge about the genome



Discourse & Society  
Copyright © 2004  
SAGE Publications  
(London, Thousand Oaks,  
CA and New Delhi)  
www.sagepublications.com  
Vol 15(4): 369–389  
10.1177/  
0957926504043705

HELENA CALSAMIGLIA AND TEUN A. VAN DIJK  
UNIVERSIDAD POMPEU FABRA

**ABSTRACT.** In this article, we examine some properties of the interface between meaning and knowledge in popularization discourse in the Spanish press about the sequencing of the human genome. After a multidisciplinary account of popularization in terms of text, context and knowledge, we analyze some semantic aspects of 42 texts in *El País*, focusing especially on denominations, explanations and the description of new objects. Besides the usual metaphors conceptualizing the genome in terms of a code or a book, and sequencing as decodification, we especially found that descriptions of new objects tend to be organized using a limited number of fundamental categories, such as Location, Composition, Size, Quantity, Appearance and Functions. We surmise that these meaning categories correspond to underlying cognitive categories that organize the schematic structure of knowledge about things. Both in the discursive and the epistemic analysis of these texts, we are specifically also interested in the strategies of specialized journalists for the management of knowledge: what knowledge is being presupposed, what knowledge is being 'reminded' or actualized and what knowledge is expressed and newly constructed.

**KEY WORDS:** *context, El País, explanation strategies, human genome, knowledge, popularization discourse, press, science communication*

### *Introduction: Aims*

The news about the (nearly) completed sequencing of the human genome in June 2000 appeared in many newspapers around the world, and also in Spain. In this article, we examine some properties of such popularization discourse in the Spanish press. Our analysis will focus on two main dimensions of press articles about the completed sequencing of the human genome: their textual structures and strategies, on the one hand, and their underlying knowledge management on the other hand.

Popularization is a vast class of various types of communicative events or genres that involve the transformation of specialized knowledge into 'everyday' or 'lay' knowledge, as well as a recontextualization of scientific discourse, for instance, in the realm of the public discourses of the mass media or other institutions. This means that popularization discourse needs to be formulated in such a way that non-specialized readers are able to construct lay versions of specialized knowledge and integrate these with their existing knowledge. Thus, various strategies of explanation, such as definitions, examples, or metaphors, among many others, are the semantic means that allow language users to relate new knowledge to old knowledge. This article will focus on the interface of these discursive and epistemic structures and strategies of popularization.

Our analysis is based on a corpus of 42 articles that were published in the Spanish daily *El País*, the leading elite newspaper in Spain, between 24 June and 2 July 2000, that is, the week that the news on the (nearly) completed sequencing of the human genome was published, namely on 26 June 2000.

Most of the news on the soon to be accomplished sequencing of the human genome is not so much about the biological or chemical properties of genes or chromosomes themselves, but rather on various aspects of the 'social context' of this eminent scientific feat. Thus, we read about the participating scholars, about laboratories, historical developments in genetics, about the competition and conflicts between the 'private' and the 'public' teams that were involved in this 'race', and especially about the possible applications of genome research in the fields of health, food and human reproduction (Boon, 2002; Dennis, 2001; Kay, 2000; Sulston and Ferry, 2002). Also these 'social' aspects are part of science communication: People thus learn about science as a social and human activity, about scholars, conflicts, problems, developments as well as the uses or abuses of science in society. Indeed, it is likely that the public will memorize much better these aspects of science than the more technical knowledge involved, especially if such information is more relevant in their everyday lives.

In this article, however, we ignore most of these 'social' aspects of the communication of this scientific event, and focus on those (parts of) texts that are about the biological and chemical properties of the genome itself, or about the technical procedures of sequencing. Whereas most of the social aspects of genome research are more or less understandable by the public at large, this is not the case for the specialized knowledge about the genome, and it is therefore that aspect of the press articles that deserves special attention here. That is, it is the formulation of these 'technical' or 'specialized' aspects of the news that is usually considered the core of the global communicative act of popularization.

### *Popularization*

Before we deal with the discursive and epistemic structures of popularizing discourse in the press, we need to briefly comment on our concept of science popularization (*divulgación* in Spanish, *vulgarisation* in French; see also, e.g. Bucchi,

1998; Fayard, 1993; Gregory and Miller, 1998; Jeanneret, 1994; Scanlon, 1999; Stocklmayer, 2001).

In this framework, this article aims to study the linguistic–discursive structures of the ways in which science is reported in the written media of mass communication, that is, outside the realm of science itself. In this sense, our investigation is in line with research on the discourses of science popularization by several other authors (Beacco, 1999; Calsamiglia, 1997, 2000, 2003; Ciapuscio, 1993, 2001; Jacobi, 1999; Moirand, 1997). Important in this work is not only the study of popularization discourse in general, or popularization in the press in particular, but also the special attention paid to the detailed structures of text and talk that play a role in the presentation of knowledge that is difficult to access by non specialized readers.

Instead of developing here at length our theory of popularization, we merely summarize some of its main tenets:

1. Popularization is a social process consisting of a large class of *discursive–semiotic practices*, involving many types of mass media, books, the Internet, exhibitions and other genres of communicative events, *aiming to communicate lay versions of scientific knowledge*, as well as opinions and ideologies of scholars, *among the public at large*.
2. Popularization in general, and in the press in particular, is not primarily characterized by specific textual structures, but rather by the properties of the *communicative context*: participants and *participant roles* such as scientific sources, specialized journalists, a lay public; their respective *purposes*, beliefs and *knowledge*; as well as the *relevance* of such knowledge in the everyday lives of the citizens.
3. Given our concept of context, these context properties of popularization discourse are relevant for the linguistic analysis of the ‘textual’ (verbal) structures of such discourse.
4. Popularization involves not only a *reformulation*, but in particular also a *recontextualization* of scientific knowledge and discourse that is originally produced in specialized contexts to which the lay public has limited access. This means that popularization discourse must always adapt to the appropriateness conditions and other constraints of the media and communicative events, e.g. those of the daily press or specialized magazines, in which they appear.
5. The mass media are not passive mediators of scientific knowledge, but *actively contribute in the production of new, common knowledge* and opinions about science and scientists – including information and views that do not derive from scientific sources. That is, despite their dependence on other institutions and organizations for most of their information and advertising, media managers and journalists ultimately decide *what* and especially also *how* to publish (or not to publish) about science, scientists and scientific knowledge, as part of complex processes of news production that will not further be analyzed here (see Bell and Garrett, 1998; Tristani-Potteaux, 1997).

6. The role of new knowledge production by the mass media needs to be further contextualized in relation to the other, especially the entertainment, functions of the media. Thus, the reporting about the human genome project should also be interpreted as a science 'show'.

### *Theoretical framework*

Against the background of these general assumptions on popularization, this article requires a multidisciplinary, theoretical background that combines insights into context, text and knowledge. We may summarize the main points of this theoretical framework as follows:

1. The insight is growing that adequate theories of discourse also need to deal more explicitly with the structures of *context* and text–context relations (Duranti and Goodwin, 1992).
2. Similarly, popularization discourse in the press is not merely characterized by special textual structures, such as those of news (Bell and Garrett, 1998; van Dijk, 1988), but also by the *relevant properties of the social situation* – that is by the context defined as a mental model of participants (van Dijk, 1999) – such as a specific domain (media instead of science), institutional settings (e.g. newspaper), special professional and communicative roles and goals of participants (science journalists, readers of *El País*, etc.), as well as various kinds of specialized and lay knowledge of these participants.
3. Within the framework of such a contextual analysis, the analysis of the *structures of popularization texts* needs to be geared to the appropriate accomplishment of the communicative acts of introduction, explanation, or clarification, among others, all of them with the particular function of filling a gap of knowledge through verbal means of paraphrasing.
4. The passages of popularization discourse that specifically interest us in this article, exhibit various textual instances of *explanation*, a discursive activity that has a theoretical status that is similar to – but different from – that of narration, argumentation and description (Adam, 1992; Borel, 1981; Calsamiglia and Tusón, 1999; Fuchs, 1993; Grize, 1981, 1990; Moirand, 1999, 2003; Zamudio and Atorressi, 2000).
5. There are many *types of explanation*, such as denomination, definition/description (Candel, 1994), reformulation or paraphrase (Ciapuscio, 2003; Gülich and Kotschi, 1987, 1995; Loffler Laurian, 1983, 1984), exemplification, generalization and analogies such as comparisons and metaphors, which are quite typical of popularization discourse on the human genome (Emmeche and Hoffmeyer, 1991; Keller, 1995; Rothbart, 1997).
6. Part of a contextual analysis is also a detailed *cognitive account of the structures of different types of knowledge* as well as of the *strategies of knowledge management* of the participants (Bernecker and Dretske, 2000; van Dijk, 2003; Wilkes, 1997), which presupposes a theoretical component usually ignored in studies of popularization.

7. Contrary to traditional approaches to knowledge in epistemology (Bernecker and Dretske, 2000), and psychology (Wilkes, 1997), we define knowledge in pragmatic terms, namely as the beliefs that are shared by epistemic communities and certified by criteria that are historically and culturally variable (van Dijk, 2003). Thus, what is knowledge for one community can be mere opinion, superstition or ideology for others. This knowledge is legitimized by – possibly biased – criteria established by special epistemic institutions in each community, such as universities, laboratories, schools, or the media, and their professional participants in our current K-community (Aronowitz, 1988).
8. There are several kinds of knowledge, such as episodic knowledge about specific events or more abstract, general, socio-culturally shared knowledge about the structure of the world (such as the structure of the human genome) (van Dijk and Kintsch, 1983). Similarly, knowledge may be shared by different communities: between individuals, among groups (such as scientists or newspaper readers), among the citizens of a country, or all the competent members of a culture.
9. Each kind of knowledge, defined in terms of mental representations in memory, has its own format. Thus, knowledge about specific events, such as the presentation of the near-completion of the human genome, is assumed to be represented in mental models, which are typically expressed in everyday stories and news reports. More general, abstract knowledge about the human genome is assumed to be represented by schemata of which the categories (such as location, types, components, or size) typically show up in explanations in didactic and popularization discourse.

### *Corpus*

The popularization texts to be analyzed here are taken from *El País*, the prestigious elite newspaper of Spain, founded after the demise of Franco's dictatorship in 1976, and politically close to the socialist party, the PSOE. Its readers typically belong to the better-educated upper-middle class, especially those who are politically left of center, although the newspaper also functions more generally as the newspaper of reference, as is the case for *Le Monde* in France or *The New York Times* in the USA. For an elite newspaper it is remarkably widely distributed – it has nearly two million readers, more than the other national newspapers in Spain, such as *El Mundo* and *ABC*. It features a weekly section, *Futuro*, which has regularly paid attention to the news on the genome since 1985. Articles on the genome also appear in other sections, such as *Sociedad*. The electronic version of the newspaper is the most visited Internet site in Spain (Informe Quiral, 1997–2002).

In 2000 the average number of articles on the genome suddenly jumped from a few dozen to more than 200, chiefly because of the announcement of the (near) completion of the sequencing of the human genome on 26 June 2000. This frequency has decreased somewhat but has remained high since 2000. There are three major 'highs' of human genome news:

- June 2000 (42 articles in *El País*): The first announcement of the soon to be completed sequencing of the human genome. We analyze only the 'technical' fragments of these articles, ignoring the social aspects of the news about the event.
- February 2001 (65 articles in *El País*): the nearly completed sequencing of the human genome, and the news about the unexpectedly low number of human genes.
- Mid-April 2003 (8 articles in *El País*): the recent announcement of the fully completed sequencing.

We have selected the first of these moments for special attention, given that this marks the first broad coverage of the human genome in the daily press, and hence is an ideal moment to study the first stages of popularization on this topic. Also, this event had important political dimensions, as emphasized by the accompanying presentations by the then US President Bill Clinton and the UK Prime Minister Tony Blair.

News of this major scientific event was preceded and followed by many articles on all possible aspects of the genome and the possible applications and dangers of our insight into its structures and functions.

The articles we analyzed are of different genres, such as proper news articles, background articles, interviews and an extract from a popular book on genetics. The authors are general journalists, science journalists, and other specialists in the general field of science and culture.

Half of the 42 texts comprising our corpus appeared in the *Society* section of the newspaper, which carries both news and background articles, whereas two news articles appeared on the front page, and two as international news. Our corpus also features two editorials, four cartoons, four opinion articles, one letter to the editor, and a translation of a chapter from an introductory book on the genome in the (special Sunday supplement). Five articles only appeared in the electronic version of the newspaper. Two articles were accompanied by a graphical explanation of the structure of the genome.

### *The semantics of popularization discourse about the genome*

In this section we analyze the most prominent semantic properties in the press articles about the sequencing of the human genome. Our analysis focuses on the following explanatory structures of these passages: denominations, metaphors, definition and description and related devices used to discursively represent and link old and new knowledge about the genome.

#### DENOMINATION

##### *'Sequencing' as description*

The news in the last week of June 2000 was not primarily on the human genome in general, but – as is normally the case for most news articles – rather about a

prominent event, namely the (nearly) completed *sequencing* of the human genome. 'Sequencing', however, is not the kind of act or procedure ordinary language users and newspaper readers are familiar with. So, from the start, journalists who want to write about this remarkable scientific feat, need to introduce a concept that is new to many people. Thus, before we examine some of these fragments let us for comparison briefly cite the webpage of the Human Genome Project (HGP; [www.ornl.gov/TechResources/Human Genome](http://www.ornl.gov/TechResources/Human%20Genome)) in order to get the 'official definition' of what sequencing is and how it is done. Note that given the functions of the website of the HGP, also this explanation is a form of popularization, as we can see in the first definition of what sequencing is, which is followed by a more technical explanation of *how* sequencing is done:

DNA sequencing, the process of determining the exact order of the 3 billion chemical building blocks (called bases and abbreviated A, T, C, and G) that make up the DNA of the 24 different human chromosomes (. . .)

- Chromosomes, which range in size from 50 million to 250 million bases, must first be broken into much shorter pieces (*subcloning step*).
- Each short piece is used as a template to generate a set of fragments that differ in length from each other by a single base that will be identified in a later step (*template preparation and sequencing reaction steps*).
- The fragments in a set are separated by gel electrophoresis (*separation step*).
- New fluorescent dyes allow separation of all four fragments in a single lane on the gel.
- The final base at the end of each fragment is identified (*base-calling step*). This process recreates the original sequence of As, Ts, Cs, and Gs for each short piece generated in the first step.
- Current electrophoresis averages about 500 to 700 bases sequenced per read. Automated sequencers analyze the resulting electropherograms, and the output is a four-color chromatogram showing peaks that represent each of the four DNA bases.
- After the bases are 'read,' computers are used to assemble the short sequences (in blocks of about 500 bases each, called the read length) into long continuous stretches that are analyzed for errors, gene-coding regions, and other characteristics.

Now let us examine how such sequencing is described in the press (the numbers after the examples refer to the texts from which they were taken):

(1,1) una descripción precisa y esencialmente completa del libro de instrucciones que lo ha hecho posible: el genoma humano. (5)

a precise and essentially complete description of the book of instructions that *made this possible: the human genome*

(1,2) descripción completa del material genético humano (38)

a complete description of human genetic material

In order to be able to explain (and ultimately to change) the world, one of the major tasks of science is to *describe* the world. It is not surprising, therefore, that description plays such a prominent role in the denomination of the scientific

procedures of sequencing. This term might be considered as a hyperonym of 'sequencing' if we take 'sequencing' to mean (roughly) 'describing the nature and the order of the components of a genome'. In Extract (1,2) we also find expressed the current news, namely that the sequencing is no longer partial, like before, but complete. Note that besides insight into the sequence of genes, a really 'complete' description would also require information about the packaging of DNA in chromosomes. The object of the scientific action of sequencing is variously described in terms of genes or genome or, as in these examples, as genetic material, or in terms of metaphors ('the book of life') which have been around in science since the 1960s and later became used more generally in the media (Kay, 2000). We shall come back to these and other descriptions of the human genome itself.

Second, the characterization of the scientific accomplishment may remain quite close to the notion of 'sequencing' by using this very term, or expressions such as:

(1,3) determinar el orden de las bases en el ADN (para secuenciar el ADN) (11)  
to determine the order of the bases in the DNA (in order to sequence the DNA)

(1,4) la determinación del número y lugar que ocupan los genes en la inmensa cadena química desentrañada (2)  
the annotation of the genome, that is, in determining the number and the place the genes occupy in the immense chemical chain that was discovered

In this case, the popularization description focuses on a core property of sequencing, namely establishing the precise order of the genes. Less common is reference to the scientific act of 'assembly' in order to get to know the order of genes in the genome. In Extract (1,3), the description is more precise, and semi-technical, while referring to bases and DNA.

#### METAPHOR

One of the main semantic means of establishing links between two domains of experience, meaning or knowledge is metaphor, and it is therefore not surprising that metaphors play a prominent role in popularization discourse. This is especially true in the case of genetics, where scholars have, from the start, themselves defined and explained genes and the genome in terms of pervasive linguistic or semiotic metaphors of the genome in terms of a 'text', 'code' or 'language'.

Thus in our data we find two sets of metaphor, namely those applying to the scientific acts or procedures, such as that of sequencing, as examined earlier, on the one hand, and those applying to the properties of the genome itself, on the other hand.

#### *Sequencing as 'decoding'*

If the human genome is described in terms of a code (see later), then the complex scientific procedure of sequencing, initially denominated as a form of description, may further be *explained* in terms of the metaphorical verbs 'decoding' and 'deciphering', as in the following examples:

(2,1) el desciframiento del genoma humano (2, 6, 38)  
the deciphering of the human genome

(2,2) al descifrar el genoma humano se entreabrió el libro de la vida (5)  
by deciphering the human genome, the book of life was opened

(2,3) el genoma humano, considerado el auténtico libro de la vida, se dio ayer por descifrado en sus partes esenciales (6)  
the human genome, considered to be the authentic book of life, was finally deciphered yesterday in its essential parts

(2,4) descodificar todo el genoma humano (10)  
the decoding of the whole human genome

(2,5) desentrañar el 'libro de la vida' (1)  
discovering the 'book of life'

These verbs have meanings that imply that the information in the genome can be compared to some kind of text, and that reading the text is not obvious or straightforward, but difficult: the 'text' is apparently not easy to understand, and therefore not immediately accessible, but needs to be analyzed by scholars. It is this philological and linguistic area of concepts that organize much of our commonsense thinking and discourse about the genome. This is even the case within genetics itself, where complex molecular relationships or processes have been described in terms of codes, coding or codification. Note that once the metaphor of 'decoding' has been used, the object may also become metaphorized, so that the genome is conceptualized as 'the book of life', as we shall see in more detail later.

*The metaphorical construction of the genome*

Whereas the nearly completed sequencing and their actors constituted the main event for the news, the object of that act, namely the human genome itself, also played a central role. However, the human genome is a biological or chemical construct and in order to understand the very point of the scientific news, readers need to understand at least something of what genomes are. Describing the genome is more complex than the act of sequencing, which can roughly be understood, as we have seen earlier, in terms of such everyday notions as 'establishing the order of' or in terms of metaphorical notions such as 'decoding'. It is not surprising therefore that the commonsense description of the human genome in the press is largely formulated in terms of familiar linguistic or semiotic metaphors, already referred to earlier:

- A. THE GENOME IS A CODE
- B. THE GENOME IS THE CODE OF LIFE
- C. THE GENOME IS A TEXT
- D. THE GENOME IS THE BOOK OF LIFE

as well as other, related metaphors from the same field, as is obvious in the following examples from our corpus:

- (3,1) primer borrador de la secuencia del código genético humano (6)  
the first draft of the sequence of the human genetic code
- (3,2) 'el lenguaje con el que Dios creó la vida' [Clinton] (6)  
the language with which God created life
- (3,3) el orden de todos los elementos del código genético humano, el llamado libro de la vida (6)  
the order of all the elements of the genetic human code, the so called book of life
- (3,4) El libro con las instrucciones del *Homo sapiens* se acabaría en 2003, dos años antes de lo previsto (13)  
the book with the instructions of *Homo sapiens* would be finished in 2003, two years before it was foreseen.
- (3,5) base de datos independiente, aislada, estable a lo largo de las generaciones (11)  
an independent, isolated database that is stable during the generations
- (3,6) el genoma humano, el manual de órdenes genéticas que definen al hombre (1)  
the human genome, the handbook of genetic instructions that define man
- (3,7) El conjunto de genes alojados en 23 pares de cromosomas constituye una autobiografía de la especie humana. (32)  
the set of genes stored in 23 chromosomes constitute an autobiography of the human species.
- (3,8) el plano general genético, el genoma de cada ser vivo en cuestión (38)  
the general genetic plan, the genome of each living being in question

As we already suggested in our analysis of the decoding metaphors for the procedure of sequencing, the use of the linguistic, philological or semiotic metaphors to describe the genome itself implies that the very complex structures of the genome are not easy to understand. Hence the recourse to the concepts of a large metaphorical domain as targets to explain some aspects of the genome, as is more generally the case in both science and popularization (Rothbart, 1997). Thus, we find that the genome may be conceptualized, as in these examples, as a code, a language, a book, or as a database, or even an autobiography of the human species, to which second-order also metaphors apply. If the genome is a text or book, we may also expect terms such as 'borrador' ('draft'), as in Extract (3,1). We hardly need to recall here that, according to contemporary theories of metaphor, these metaphors are cognitively based and organized in such a way that an area of knowledge or perception that is unknown to the common readers is being reconceptualized in terms of a more concrete, familiar area, such as that of language (see, among many other studies, Hallyn, 2000; Lakoff and Johnson, 1980). Combined with the conceptualization of the *functions* of the human genome in the life of people, the metaphor may then be extended towards the standard metaphor of the book of life, as we also see in these examples. Somewhat more specific in this case is the use of metaphorical expressions such as 'instructions' or 'plan', which focus first of all on elements of the text or book that is the entire genome, namely the genes, and second, upon their functions in

the formation of proteins and the (indirect) control of bodily development and processes. Finally, in Extract (3,2) we find that in the quote from Clinton's speech a more religious perspective associates the book of life with God, maybe so that the reader would not limit the description of life to mere biological or chemical programming. Another text, not cited here, metaphysically associates the genome with the human 'soul'. Closely related to the standard metaphors for the human genome, there are many others, such as the earlier (not quite appropriate, and two-dimensional instead of sequential) uses of the metaphor of 'map of life', which emphasize the visual and representational nature of the code.

#### DEFINITION AND DESCRIPTION

Other fundamental forms of explanation are definitions and descriptions, where definitions are used to explain unknown words, and descriptions to explain unknown things. In actual practice these different kinds of explanation are often confounded. Indeed, also theoretically word meaning and world knowledge are hard to distinguish. In our analysis, then, we focus on the explanation of some technical genetic notions by examining the descriptions of the phenomena they refer to: cells, genes, genome, chromosomes, DNA, etc.

Interestingly, what we have found is that such definitions and descriptions are not arbitrary, but organized by a number of very fundamental conceptual categories that probably reflect the structure of knowledge. That is, our semantic discourse analysis of explanations offers new insights into the interface between discourse and knowledge, an interface that is crucial in the account of the social and cognitive processes of popularization.

Definitions and descriptions may have widely varying scopes. They may be limited to one or a few words or a clause, but also extend throughout paragraphs and large sections of discourse.

Let us begin with a few examples of the *discursive* organization of explanatory descriptions. Consider the following example explaining the nature of life:

(4,1) Los seres vivos están compuestos por miles de máquinas microscópicas (las proteínas) que les permiten moverse, respirar, obtener energía e intercambiar materiales con el entorno. Hasta ahí todo bien. Pero lo que nadie podía imaginarse era que la información necesaria para construir todo eso se guardara en una base de datos independiente, aislada, estable a lo largo de las generaciones, transmitida de padres a hijos de forma ajena a los avatares de la existencia de cada individuo. Ese fue el descubrimiento del monje austriaco. (11)

Living beings are composed of thousands of microscopic machines (the proteins) that allow them to move, breathe, get energy, and exchange material with the environment. So far, so good. But what nobody could imagine was that the information necessary to construe all that was stored in an independent, isolated database that is stable from generation to generation, transmitted from parents to children in a form that is impervious to the vicissitudes of each individual existence. That was the discovery of the Austrian monk.

This description features the following analytical categories:

- 
- a. Composition: Living things consist of proteins.
  - b. Quantity: thousands of proteins.
  - c. Size: microscopical.
  - d. Localization: all this is stored in a database.
  - e. Time, development: over generations.
  - f. Properties: independent, isolated, stable.
  - g. Process: transmitted from parents to children.
  - h. Context, circumstances: in a way that is impervious to the vicissitudes of the existence of each individual.
  - i. Functions: Proteins allow living things to move, etc.; the information necessary to construe all this (proteins).

We see that descriptions appear to be ordered by abstract categories characterizing the structural properties and functions of biological objects and phenomena. We have found that most of the descriptions of objects as well as the definitions of technical terms feature several of these schematic meaning categories. Later we see that such *semantic* structures also reflect underlying *cognitive* structures.

Descriptions not only provide 'structural' information in terms of components or properties of things, but often also have a more dynamic, functional or causal dimension, in which a well-known fact B is related to an often unknown condition or cause A. Thus, in the example just quoted, we see that phenomena at the 'macro'-level of living things, bodies, and reproduction are explained in terms of 'micro' phenomena such as proteins and the information that is stored in the genome. It is this double level of description as a form of explanation that is contextually relevant for popularization discourse, because it addresses the readers' interest in knowing things about themselves and other human beings.

As we have seen before for other examples, this fragment features *metaphors* as part of the description, namely the use of 'microscopic machines' in order to explain the notion of proteins, or when it compares genes and their information with a database. In both cases these metaphors link biological phenomena with contemporary technical phenomena that are better known to the educated readers. We encounter information about the history of science, namely a reference to Mendel, described as the 'Austrian monk' (more correct would probably be the 'Czech' monk if we speak of contemporary countries), a rhetorical enhancement of the originality of his findings ('what nobody could imagine'), and a semantic-pragmatic marker ('so far, so good') organizing the paragraph in two parts, a summary of old information, and emphasizing what is new. Finally, we see that the concept of proteins is introduced by a reformulation, using parentheses for the semi-technical notion. The rest of the explanation is formulated in terms of everyday discourse.

Somewhat more 'technical' is the following description of the structure of DNA:

(4.2) El gigantesco ADN está compuesto por miles de millones de compuestos

pequeños llamados *bases*. Las bases son sólo de cuatro tipos: A (adenina), C (citocina), G (guanina) y T (timina), cuyas cantidades relativas varían de una especie a otra de animal o planta. Pero, para cualquier especie, la cantidad de A siempre es igual a la de T, y la cantidad de G siempre es igual a la de C. (11)

The gigantic DNA is composed of millions of small compounds called bases. Bases are only of four types: A (adenine), C (cytosine), G (guanine) and T (thymine) whose relative quantities vary from one species of animals or plants to another. But for each species the amount of A is always equal to that of T, and the quantity of G is identical to that of C.

Again, in this fragment we find the expression of information in the following categories:

Composition: DNA is composed of (. . .) bases.

Size: Gigantic DNA, small components.

Number: Billions.

Type: Bases are of only four types: A, etc.

Variation: Whose quantities vary from one to the other species.

Generalization: For any species, the quantity . . .

Comparison/pairing: Quantity of A always equal to that of T, etc.

We see that descriptions appear to be organized in terms of a number of general (if not universal) categories that language users know for any other object in their surroundings. In addition to the frequent categories of Composition, Size and Number, we here find categories of variation, such as Type and Variation, as well as Generalization, which extends the description from individual (structural) properties to properties of whole sets of objects or phenomena. Note the use of metalinguistic expressions such as 'called' and of abbreviations 'A (adenine)', both introducing technical or specialized (biological) terms. The overall explanation, however, is limited to an account of structure that is not much different from any kind of everyday description of (new) objects or things.

Finally, let us examine the following more extended example:

(4,3) El ADN, según acababan de descubrir Watson y Crick, es una larguísima doble hilera de bases (la célebre doble hélice). Cada hilera tiene este aspecto: '. . . AATCC-TAGGCT. . .' y así millones de bases, en cualquier orden posible. Pero entre una hilera y la otra, las bases se aparean siempre siguiendo las reglas de Chargaff: A siempre con T; G siempre con C. Esta complementariedad es el secreto de la vida, porque si las dos hileras se separan, cada una puede reconstruir a la otra: por eso los genes pueden sacar copias de sí mismos, las células pueden replicarse y los seres humanos pueden reproducirse. Otra consecuencia crucial de la estructura de la doble hélice es que la información genética (lo único que es distinto entre dos genes distintos) debe necesariamente estar contenida en el orden de las bases en la hilera al igual que la información literaria está contenida en el orden de las letras en un texto. (11)

DNA, as Watson and Crick had just discovered, is a very long double string of bases (the famous double helix). Each string has the following appearance: '. . .AATCCTAGGCT. . .' and so forth for millions of bases, in any possible order. But between each string and the next, the bases always come pairwise following the rules of Chargaff: A always with T, G always with C. This complementarity is the secret of life, because if

the two strings separate, each one can replicate itself and human beings can reproduce themselves. Another crucial consequence of the structure of the double helix is that the genetic information (the only thing that is different between two different genes) must necessarily be contained in the order of the bases of the string, in the same way as literary information is contained in the order of the letters of a text.

In addition to the categories already postulated earlier (Composition, Quantity, Size, Localization, Cause, Consequence), we find the following: Form (the double helix) and Order (in any possible order), as well as a Comparison (order of genetic information is comparable to the letters in a literary text), as one of the well-known didactic ways to explain abstract or difficult to visualize phenomena in terms of everyday experiences. And finally, again the explanation of the structures of genes is related to real-life experiences and questions ('the secret of life').

Explanatory descriptions are organized by semantic categories that project the 'order' of the world of things, or rather of the world of knowledge, into the meaning of the text. Another way to do this, as we have seen, is by tropes of various kinds, such as comparisons, metaphors, analogies, and so on. Such expressions are not limited to isolated words or notions, but may also organize large parts of a text, for instance, when a whole series of comparisons or metaphors are used. Note in this text the various forms of personification of the micro life of genes in terms of the macro world of human society:

(4,4) Somos repúblicas de células. El genoma – el mismo en todas las células – dice a cada una lo que tiene que hacer y cómo hacerlo. Es el director de la orquesta celular y la partitura de la música de la vida. La célula sabe cómo leer los genes, pero nosotros, hasta hace muy poco, no. Cada triplete de bases codifica un aminoácido. El gen entero, una proteína. El genoma es la biblioteca de la célula.

Cada vez que la célula tiene que hacer algo consulta la biblioteca y copia (en ARN mensajero) el libro o capítulo que le interesa, poniendo luego en práctica sus instrucciones mediante el ensamblaje de las correspondientes proteínas en los ribosomas. En efecto, las instrucciones genéticas conciernen directamente sólo a la fabricación de proteínas. Pero estas proteínas pueden ser muy distintas (enzimas, hormonas, anticuerpos, etc.) y producir todo tipo de efectos, desde uñas, hasta enfados, pasando por enfermedades y curaciones. (19)

We are a republic of cells. The genome – the same in all cells – tells each of them what it has to do and how to do it. It is the director of the cellular orchestra and the score of the music of life. The cell knows how to read the genes, but we, until only a short while ago, did not. Each base-triplet codifies an amino acid. The entire gene, a protein. The genome is the library of the cell.

Each time the cell has to do something it consults the library and copies (in messenger RNA) the book or chapter it is interested in, and then puts into practice the instructions by means of the assembly of corresponding proteins and ribosomes. Actually, the genetic instructions are only directly relevant for the fabrication of proteins. But these proteins may be very different (enzymes, hormones, antibodies, etc.) and they may produce all kinds of effects, from nails to anger, through illnesses and cures.

The metaphors here are quite heterogeneous: the quote opens with a comparison of people with a republic and of genes as the director of an orchestra, and then

continues with the full sequence of metaphors based on some form of symbolic code: genes are score of music and the genome is a library. These need no further analysis here apart from saying that in popularization discourse almost any metaphor will do to explain the structures and functions of genes. Note also the GENES ARE INSTRUCTIONS metaphor, which conveys the meaning that genes exercise control, as is also the case when the genome is described as the 'director' of the cell.

In the majority of the semantic procedures used in the popularization of the genome we witness a general tendency to relate the units of biological 'micro' analysis with objects and actions in everyday human life at the 'macro' level, not only by relating these two levels by means of the functions of genes, but also by constant metaphorical links that make the micro-world more comprehensible. In another perspective this is also the case in the frequent media stories about the development of modern genetics, from Mendel to the present, via the discovery of the double helix by Crick and Watson. Also in these stories the relevance of the micro-world of the genome is routinely related to both the world of everyday life as well as to the relevance of scientific knowledge of the genome for humankind.

#### OTHER EXPLANATORY STRUCTURES

We shall be very brief about the other explanatory structures we have found in our data. Thus, quite common in popularization discourse, and also in our corpus, is the use of *reformulations and paraphrases*, formally marked by relative clauses, appositions, parentheses, dashes, quotes and metalinguistic expressions ('are called', etc.). Interesting for our analysis is that such explanatory moves establish a link between old and new knowledge, where usually a new notion is introduced first, followed by an explanatory reformulation or paraphrase (for details see Ciapuscio, 2003; Fuchs, 1993; Gülich, 2003; Jacobi, 1986). In the following example, the new term follows the (metaphorical) paraphrase:

(5.1) Los seres vivos están compuestos por miles de máquinas microscópicas (las proteínas) (11)

Living beings are composed of thousands of microscopic machines (the proteins)

Another explanatory device, not quite frequent in our data, is *exemplification*, which means providing specific examples of general phenomena, such as mentioning Alzheimer's as one of the diseases that might be better understood now that the human genome has been sequenced. Similarly, describing special cases, for instance, in a story about a patient, expresses mental models that are mental 'instantiations' of general knowledge of such a disease. Such stories and mental models are often easier to remember than general knowledge and hence are quite useful as an explanatory device in didactic and popularization discourse. The explanatory move of *generalization* does the opposite by drawing general conclusions from specific examples or cases, and thus is a discursive manifestation of learning. Here is an example of such a generalization in our data:

(7.1) Reducidos al principio de las cosas, todos los seres vivos, desde los gusanos a los

humanos, compartimos la misma organización elemental, la misma geometría, una armonía exacta y diminuta. (22)

Reduced to the principle of things, all living beings, from worms to humans, share the same elementary organization, the same geometry, an exact tiny harmony.

### *The knowledge structures of popularization*

After our analysis of some textual structures of popularization discourse in the Spanish press, let us finally make some brief remarks on the knowledge structures and strategies that are the basis of science communication of this event focusing especially on the underlying knowledge structures that may be related to the structures of explanations in definitions and descriptions, as discussed earlier.

Our analysis of the structures of the newspaper texts suggests the following forms of contextually controlled knowledge management:

1. Presupposition of epistemic universals, such as those of order, degree and quantity (sequence, very, numbers, etc.).
2. Activation of a metaphorical knowledge source-domain (language, code) in order to describe unknown properties of DNA: a sequence of letters, etc. This source domain need not always be more 'concrete', but it should be more familiar to the readers (more accessible to everyday experiences, as is the case for the language metaphor).
3. New terms, concepts or knowledge are related to existing knowledge by operations such as part-whole, composition, constituency, and expression.
4. There are pragmatic markers that keep interest alive, facilitate comprehension, dramatize events (so that the text becomes more interesting), stimulate the establishment of relations between target and source domain, etc., thus facilitating the formation of new knowledge by stimulating interest and motivation.
5. The kind of knowledge presupposed among readers is general socio-cultural knowledge 'of the world', e.g. about bodies, cells, texts and letters, some semi-technical (educated) lay knowledge about elementary biology, chemistry and genetics, for instance about DNA, chromosomes, genes and molecules, and occasionally some technical knowledge explained earlier in the same or related articles, for instance about what a nucleotide is or the names of the four types of bases that constitute it.
6. Our analysis of the popularization texts has suggested that knowledge about the genome is organized by a schema consisting of such categories as Localization (Genes are found in the chromosomes), Composition (DNA consists of bases), Form, structure (DNA has the form of a double helix), Appearance, size, color, etc. (DNA is a gigantic string), Quantity (DNA consists of 3 billion bases), Relations (The bases A and T and G and C always go together), Types (The bases are of 4 different types), and Functions (Genes control the production of proteins). Knowledge about more 'dynamic' aspects of genetics, such as acts and procedures (sequencing, etc.) or biochemical

processes (cell division) is usually expressed by well-known, more general concepts (such as description) or metaphors (decoding, annotation).

We may conclude more generally that the new knowledge expressed and conveyed in our data, and probably by most science communication in the press, is largely social, namely about the world of science, about scientists and their competition, and about possible applications of scientific results in everyday life. The more 'technical' knowledge conveyed is very limited and vague, organized by a general schema of knowledge categories, and often conceptualized in metaphorical notions. This more technical knowledge presupposes fuzzy semi-technical ('high school') knowledge, e.g. about cells, molecules, DNA and genes, and in our data is limited to quantity (the number of bases in our DNA), composition (the four different bases of DNA), functions (relations between genome, genes, proteins and properties of cells and bodies), and processes (copying or translation of genetic information). These general aspects of knowledge communication and construction in the press probably hold more generally for popularization and education: Further analysis of a popularization book (Ridley, 1999) as well as of the first chapter of a textbook of genetics (Griffiths et al., 2000) show the same strategies.

### *Conclusions*

In this article, we provided an analysis of the news about the announcement of the nearly completed sequencing of the human genome in June 2000 as it was reported in the Spanish press. Most of this news is about the social aspects of this finding, about the scientists, conflicts, and on the possible application of the findings in the areas of health, food or reproduction. Also relevant is the political dimension of this event, as is obvious from the joint declaration of Clinton and Blair. Interested in the properties of popularization discourse in the press, we focused on the few passages that tell us about the human genome itself. Popularization was conceptualized in a multidisciplinary framework around three fundamental concepts: text, context and knowledge, which articulate the discursive, social and cognitive aspects of popularization and of communication in general.

Succinctly defined as the communication of science, popularization needs to be examined, in particular, in terms of its context properties: the participants (specialized journalists, newspaper readers), media (newspaper), objectives (to explain or learn something) and the knowledge gap between writers and readers. Analysis of the 42 texts in our corpus focused on semantic strategies of explanation. That is, we explored the interface between such discourse and the underlying structures and management of knowledge: how do people learn about new things, and how does popularizing discourse manage its means so as to enable or improve such understanding and learning.

Our main findings pertain to the interesting parallelism between the

structures of meaning and of knowledge in the discursive act of explanation. New objects are being described in terms of a limited number of basic categories, such as Localization, Composition, Size, Number, Appearance and Functions. We assumed that these are not merely semantic categories, but probably categories of a basic knowledge schema for the representation of things. In order to facilitate understanding of complex biochemical processes or laboratory procedures, these schematic categories are made more accessible by metaphorization, essentially around the conceptual equivalence GENOME IS A TEXT, which allows the rich series of metaphors about decoding and deciphering for laboratory actions, and copying for molecular processes. That is, whereas objects such as cells, chromosomes, DNA and molecules tended to be described in terms of schematically organized object-dimensions (fundamentally as structures and functions), molecular processes tend to be represented in terms of corresponding human actions (copying, dividing, reading, and so on). The initial denominations and first definitions follow these principles of schematic knowledge categories, on the one hand, and anthropocentric metaphorization, on the other hand. We stressed that the majority of these metaphors are not introduced by popularizing discourse, but within genetics itself.

Finally it is interesting to observe that the recontextualization of scientific knowledge in popularization discourse and its transformation into everyday, commonsense knowledge, combines precise knowledge with fuzzy or approximate knowledge, which, however, will form the basis for further 'learning': either in future news and popularization discourse or motivating more systematic approaches through other media. It would therefore be very interesting to compare the use of old and new concepts of genetics during the respective highs of genome coverage of the press in 2000, 2001 and 2003. One may wonder, however, whether the more or less vague new knowledge thus being created among the public at large by popularization discourse is adequate as a basis for further learning about the genome – a question which requires a very different kind of (psychological) study of the relations between popularization discourse and actual understanding by the readers.

Although we have developed some ideas about the meaning–knowledge interface of popularization discourse, and hence more generally about the discursive strategies of explication, exposition and learning, the discursive, social and mental processes of new knowledge construction in society, are as yet only superficially understood. More multidisciplinary research will be necessary to obtain the kind of understanding about how exactly such and other discourse structures are involved in the acquisition of new knowledge.

#### ACKNOWLEDGEMENTS

We are indebted to Brigitte Nerlich, Robert Dingwall and an anonymous reviewer for their useful commentary on an earlier version of this article.

## REFERENCES

- Adam, J.M. (1992) *Les textes: types et prototypes*. Paris: Nathan.
- Aronowitz, S. (1988) *Science as Power: Discourse and Ideology in Modern Society*. Minneapolis: University of Minnesota Press.
- Beacco, J.C. (1999) *L'astronomie dans les médias. Analyses linguistiques de discours de vulgarisation*. Paris: Presses de la Sorbonne Nouvelle.
- Bell, A. and Garrett, P. (eds) (1998) *Approaches to Media Discourse*. Oxford: Blackwell.
- Bernecker, S. and Dretske, F.L. (eds) (2000) *Knowledge: Readings in Contemporary Epistemology*. Oxford: Oxford University Press.
- Boon, K.A. (2002) *The Human Genome Project: What does Decoding DNA Mean for Us?* Berkeley Heights, NJ: Enslow.
- Borel, M.J. (1981) 'L'explication dans l'argumentation: approche sémiologique', *Langue française* 50: 20–38.
- Bucchi, M. (1998) *Science and the Media Alternative Routes in Scientific Communication*. London: Routledge.
- Calsamiglia, H. (1997) 'Divulgar: itinerarios discursivos del saber', *Quark* 7: 9–18.
- Calsamiglia, H. (ed.) (2000) 'Decir la ciencia: las prácticas divulgativas en el punto de mira', *Discurso y Sociedad* 2 (2).
- Calsamiglia, H. (2003) 'Popularization Discourse', *Discourse Studies* 5 (2).
- Calsamiglia, H. and Tuson, A. (1999) *Las cosas del decir. Manual de análisis del discurso*. Barcelona: Ariel.
- Candel, D. (1994) 'Le discours définitoire: variations discursives chez les scientifiques', in S. Moirand, A.A. Ali Bouacha, J.C. Beacco and A. Collinot (eds) *Parcours linguistiques de discours spécialisés: Colloque en Sorbonne les 23–25 septembre*, pp. 33–4. Berne: P. Lang.
- Ciapuscio, G.E. (1993) *Wissenschaft für den Laien: Untersuchungen zu populärwissenschaftlichen Nachrichten aus Argentinien*. Bonn: Romanistischer Verlag.
- Ciapuscio, G.E. (2001) 'Procesos y recursos de producción textual en la divulgación de la ciencia', in J. Brumme (ed.) *La historia de los lenguajes iberorrománicos de especialidad. La divulgación de ciencia Institut de Lingüística Aplicada*, pp. 17–42. Barcelona: Universidad Pompeu Fabra.
- Ciapuscio, G.E. (2003) 'Formulation and Reformulation Procedures in Verbal Interactions Between Experts and (Semi) Laypersons', *Discourse Studies* 5 (2): 207–33.
- Dennis, C. (ed.) (2001) *The Human Genome*. Houndmills: Nature/Palgrave.
- Duranti, A. and Goodwin, C. (eds) (1992) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press.
- Emmeche, C. and Hoffmeyer, J. (1991) 'From Language to Nature: The Semiotic Metaphor in Biology', *Semiotica* 84 (1/2): 1–42.
- Fayard, P. (1993) *Sciences aux quotidiens. L'information scientifique et technique dans les quotidiens nationaux européens*. Nice: Z Éditions.
- Fuchs, C. (1993) *Paraphrase et énonciation*. Paris: Ophrys.
- Gregory, J. and Miller, S. (1998) *Science in Public: Communication, Culture and Credibility*. New York: Plenum Press.
- Griffiths, A.J.E., Gelbart, W.M., Miller J.H. and Lewontin, R.C. (2000) *Genética moderna*. Madrid: McGraw-Hill.
- Grize, J.B. (ed.) (1981) 'L'explication, Revue Européenne des sciences sociales et Cahiers Vilfredo Pareto', *Tome XIX*: 56.
- Grize, J.B. (1990) *Logique et langage*. Paris: Ophrys.
- Gülich, E. (2003) 'Conversational Techniques Used in Transferring Knowledge Between Medical Experts and Non-Experts', *Discourse Studies* 5(2): 235–63.

- Gülich, E. and Kotschi, T. (1987) 'Les actes de reformulation dans la consultation "La dame de Caluire"', in P. Bange (ed.) *L'analyse des interactions verbales*, pp. 15–81. Berne: Lang.
- Gülich, E. and Kotschi, T. (1995) 'Discourse Production in Oral Communication', in U.M. Quasthoff (ed.) *Aspects of Oral Communication*, pp. 30–66. Berlin: de Gruyter.
- Hallyn, F. (ed.) (2000) *Metaphor and Analogy in the Sciences*. Dordrecht: Kluwer.
- Informe Quiral (1997–2002) *Observatorio de la comunicación científica*. Barcelona: Universidad Pompeu Fabra.
- Jacobi, D. (1986) *Diffusion et vulgarisation: itinéraires du texte scientifique*. Paris: Belles Lettres.
- Jacobi, D. (1999) *La communication scientifique. Discours, figures, modèles*. Grenoble: Presses Universitaires de Grenoble.
- Jeanneret, Y. (1994) *Écrire la science: Formes et enjeux de la vulgarisation*. Paris: Presses Universitaires de France.
- Kay, L.E. (2000) *Who Wrote the Book of Life? A History of the Genetic Code*. Palo Alto, CA: Stanford University Press.
- Keller, E.F. (1995) *Refiguring Life: Metaphors of Twentieth Century Biology*. New York: Columbia University Press.
- Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Loffler Laurian, A.M. (1983) 'Typologie des discours scientifiques: deux approches', *Études de linguistique appliquée* 51: 8–20.
- Loffler Laurian, A.M. (1984) 'Vulgarisation scientifique: formulation, reformulation, traduction', *Langue Française* 64: 109–25.
- Moirand, S. (1997) 'Formes discursives de la diffusion des savoirs dans les médias', *Hermès, Communication, Politique, Cognition* 21: 33–44.
- Moirand, S. (1999) 'L'explication', in J.C. Beacco (ed.), *L'astronomie dans les médias. Analyses linguistiques de discours de vulgarisation*, pp. 141–66. Paris: Presses de la Sorbonne Nouvelle.
- Moirand, S. (2003) 'Communicative and Cognitive Dimensions on Science in the French Media', *Discourse Studies* 5(2): 175–206.
- Ridley, M. (1999) *Genome the Autobiography of a Species in 23 Chapters*. New York: HarperCollins.
- Rothbart, D. (1997) *Explaining the Growth of Scientific Knowledge: Metaphors, Models, and Meanings*. Lewiston, NY: E. Mellen Press.
- Scanlon, E. (1999) *Communicating Science. Contexts and Channels*. London: Routledge.
- Stocklmayer, S. (ed.) (2001) *Science Communication in Theory and Practice*. Dordrecht: Kluwer.
- Sulston, J. and Ferry, G. (2002) *The Common Thread. A Story of Science, Politics, Ethics, and the Human Genome*. London: Bantam.
- Tristani-Potteaux, F. (1997) *Les journalistes scientifiques médiateurs des savoirs*. Paris: Economica.
- Van Dijk, T.A. (1988) *News as Discourse*. Hillsdale, NJ: Erlbaum.
- Van Dijk, T.A. (1999) 'Context Models in Discourse Processing', in H. van Oostendorp and S.R. Goldman (eds) *The Construction of Mental Representations During Reading*, pp. 123–48. Mahwah, NJ: Erlbaum.
- Van Dijk, T.A. (2003) 'The Discourse–Knowledge Interface', in R. Wodak and G. Weiss (eds) *Critical Discourse Analysis. Theory and Interdisciplinarity*, pp. 85–109. London: Palgrave.
- Van Dijk, T.A. and Kintsch, W. (1983) *Strategies of Discourse Comprehension*. New York: Academic Press.

Wilkes, A.L. (1997) *Knowledge in Minds. Individual and Collective Processes in Cognition*. Hove: Psychology Press.

Zamudio, B. and Atorressi, A. (2000) *La explicación*. Buenos Aires: Eudeba.



HELENA CALSAMIGLIA has been teaching at the Universitat Autònoma de Barcelona since 1973 and is now Professor in the Department of Translation and Philology of the Pompeu Fabra University, Barcelona. She coordinates the Section of Discourse Studies of the department, as well as the international Xarxa dels Estudis del Discurs (Network of Discourse Studies) that connects scholars from various universities in Spain, France and Argentina. Her research has been in the fields of pragmatics, sociolinguistics, and discourse analysis, and she is currently focusing on the popularization of science. With Amparo Tusón she is the author of *Las cosas del decir. Manual de análisis del discurso* (Ariel, 1999). ADDRESS: Department de Traducció i Filologia, Universidad Pompeu Fabra, Rambla 30, Barcelona 08002, Spain. [email: helenacalsamiglia@upf.edu]

TEUN A. VAN DIJK was Professor of Discourse Studies at the University of Amsterdam until 2004 and is currently teaching in the Department of Translation and Philology of the Pompeu Fabra University, Barcelona. After earlier work on generative poetics, text grammar and the psychology of text comprehension, his work in the 1980s and 1990s focused on more social and political topics, such as racism, the press and ideology. His current research focuses on context and knowledge. He is founder editor of *Discourse and Society* and of *Discourse Studies*. His latest books in English are *Discourse Studies* (ed., 2 vols, 1997), *Ideology* (1998) and *Racism at the Top* (ed. with Ruth Wodak, 2000). He is now preparing a new book on context.