

## SPECIALIZED DISCOURSE AND KNOWLEDGE A case study of the discourse of modern genetics

TEUN A. VAN DIJK  
(Universitat Pompeu Fabra, Barcelona)

**RESUMO** Neste artigo, investigamos algumas das relações entre discurso especializado e conhecimento, por meio do estudo discursivo de um livro-texto de genética. Depois da crítica a abordagens convencionais em epistemologia e psicologia em relação ao conhecimento, argumentamos que em primeiro lugar precisamos de uma tipologia detalhada de vários tipos de conhecimento. O conhecimento científico em tal quadro teórico é definido como um conhecimento compartilhado, geral, abstrato e “certificado” de um grupo (profissional). Uma análise da estrutura discursiva de um livro-texto de genética sugere que o conhecimento científico é organizado por meio de esquemas específicos, que englobam categorias como Estrutura, Quantidade, Forma, Composição e Função, entre outras categorias, especialmente para a descrição das “coisas”. Em seguida, assume-se que o conhecimento científico encontra-se necessariamente baseado no conhecimento cotidiano, não especializado de toda a comunidade epistêmica de uma nação ou cultura. O artigo discute muitas propriedades e problemas sobre a natureza da relação entre conhecimento e discurso e conclui que nós ainda sabemos muito pouco sobre tal relação.

### INTRODUCTION

In this paper we examine some properties of the relationships between discourse and knowledge. More specifically, we focus on the role of ‘specialized’ knowledge in the study of scientific discourse, such as the discourse of contemporary genetics<sup>1</sup>.

One of the crucial contributions of modern psychology and cognitive science to the study of discourse has been the recognition of the vital role of knowledge in the production and comprehension of text and talk (Charniak, 1972; Schank & Abelson, 1977; Van Dijk & Kintsch, 1983; Wilkes, 1997). The processing of words, propositions or sequences of propositions presupposes vast amounts of sociocultural knowledge. So much so that the well-known metaphor of the iceberg is often used to explain that only a small part of the meaning of a text is actually ‘visible’, that is, expressed in the text itself, but that a large part remains implicit, while contextually irrelevant or inferable from the explicit propositions of the text --given the world knowledge of the recipient. The very useful notion of a ‘mental model’ was introduced to account (among other

---

<sup>1</sup> This paper is part of a project on discourse and knowledge financed by the Spanish Government (the Ramón y Cajal program).

things) for these implicit meanings of discourse (Johnson-Laird, 1983; van Dijk & Kintsch, 1983; van Oostendorp & Goldman, 1999).

Despite these important theoretical advances, our knowledge about knowledge and about the relations between discourse and knowledge is still quite limited. In this paper, I shall propose some distinctions between different types of knowledge, and show how these also have different impact on discourse processing and structure.

A more sophisticated account of the role of knowledge in discourse processing is especially relevant in the account of specialized (e.g., scientific) discourses, whose production and comprehension crucially depend on various kinds of specialized knowledge of its participants. This is most obvious in the use of technical terminology, but also extends to many other aspects of specialized discourse, such as its preferred topics, overall formats or text-schemas, style, rhetoric (including its typical metaphors), argumentation patterns, methods of proof and demonstration, the use of tables, figures and other non-verbal aspects of discourse, and so on. Indeed, the knowledge about specialized discourse properties is part of the very specialized knowledge of the experts. Geneticists not only know about genetics, but also how to write and talk about genetics for a specialized or a lay audience.

Whereas (specialized or other) discursive structures are traditionally assigned to discourse itself, knowledge is obviously not a property of text or talk itself, but of its users and their communities. This means that knowledge is part of an important category of the context (of use) of a discourse, and that discourse-knowledge relations should be object of research for a theory of text-context relations. In other words, we know that language users needs (lots of) sociocultural knowledge to produce and understand meaningful texts, but we need a much more explicit theory of context to tell us how exactly the management of knowledge in discourse processing takes place (Van Dijk, 2002, 2003).

As may be obvious from the previous paragraphs, the theoretical framework of this paper is multidisciplinary. Knowledge is mainly accounted for in cognitive and social psychology, and more generally in cognitive science. Whereas current philosophical approaches to knowledge are of course the concern of contemporary epistemology, I am afraid that epistemology has little original to offer for a more empirical study of knowledge. Despite a tradition of thousands of years and thousands of books, most work in contemporary epistemology is limited to arcane debates about “justified true beliefs”, absurd invented examples (the so-called Gettier counter-examples) and approaches distributed in various sects (such as Idealism, Realism, Empiricism, Skepticism, Foundationalism, Externalism, Internalism, Coherentism, Contextualism, and many other –isms)(see, e.g., Bernecker & Dretske, 2000; Greco & Sosa, 1999; Pojman, 1999; Indeed, the sociology and anthropology of knowledge probably have more to offer when it comes to spelling out what language users in various cultures (need to) know about the world in order to be able to engage in meaningful practices such as discourse. As for the theoretical account of the many ways knowledge is expressed, implied or presupposed in specialized discourse, we’ll have recourse to several directions in contemporary discourse studies, although these seldom deal with an account of what exactly specialized or scientific knowledge is (see, among

many other recent books, Duszak, 1997; Flowerdew, 2002; Gunnarsson, Linell & Nordberg, 1997; Hyland, 1998; Martin & Veel, 1998; Trosborg, 2000; Ventola & Mauranen, 1996).

## **TYPES OF KNOWLEDGE**

### **Cultural, Personal and Group Knowledge**

The psychology of discourse processing generally distinguishes only one kind of knowledge, which is not further defined in social or cultural terms but which corresponds to some kind of socioculturally shared beliefs (Britton & Graesser, 1996; Kintsch, 1998; Van Dijk & Kintsch, 1983; Van Oostendorp & Zwaan, 1994). The mental representation of this knowledge is commonly described in terms of scripts or other schematic structures stored in Long Term Memory. This kind of sociocultural knowledge is empirically ‘observable’ by the fact that it is the kind of knowledge that is presupposed in public discourse, and because it obviously is needed by human beings in order to be able to understand their surroundings, engage in social practices and in order to produce and understand discourse. I shall come back to this knowledge below.

In discourse understanding this ‘general’ knowledge is activated in order to provide missing information, for instance in the establishment of local or global coherence and the construction, in episodic memory, of mental models of specific events. If we define these mental models as the representation of the personal knowledge of individual language users about such events, we thus obtain another kind of knowledge besides the more general, socially shared knowledge. It may well be true that for others this ‘personal knowledge’ is no more than ‘mere belief’, or an opinion, but for the person who holds these beliefs, they may well be ‘true’ and hence constitute knowledge. The explicit verbal expression of this epistemic status of personal knowledge is that language users in such a case will say or presuppose “I know that...”, but more often than not, such knowledge will simply be presupposed in discourse if the speaker knows that her recipients already know what she knows. That is, personal knowledge may be shared by others, for instance with family members, friends or colleagues, for instance through conversation, but until it has been shared, this knowledge is not generally presupposed in talk or text: Language users know that speech participants usually do not have access to their personal knowledge unless they tell them about it. As is the case for other knowledge, language users may have to account for their knowledge in terms of evaluation (truth) criteria, such as personal experiences: I have seen (heard, etc) it myself, or knowledge obtained from reliable sources.

Besides these two kinds of knowledge, namely sociocultural and personal knowledge, I distinguish a third type of knowledge: Group knowledge. Such knowledge has many of the characteristics of sociocultural knowledge: It is shared by a collectivity or community of people, and it is also presupposed in discourse among the members of this collectivity. However, unlike generally shared, sociocultural knowledge, group

knowledge may not generally be presupposed in discourse directed at members of other groups. Also, as is the case for personal knowledge, group knowledge may well be qualified as 'mere belief' or 'opinion', if not as prejudice, by other groups. Much of this knowledge will be ideologically based.

It is this third kind of knowledge that is paradigmatic for scholarly, academic, scientific or other forms of 'specialized discourse'. Precisely because of the (sociological) implications of the power of science and technology in much of the world today, it is this kind of scientific knowledge that is often considered to be the ultimate criterion of truth and legitimacy, also for the evaluation of the truth of non-scientific discourse (Aronowitz, 1988; Foucault, 1980).

### **Cultural knowledge: Common Ground**

I shall assume that the basis of all social representations shared by all members of society or culture is constituted by the first kind of generally shared, sociocultural knowledge, which I shall call Common Ground (there are many related uses of the notion of Common Ground, such as the work by Clark, 1996).

We suggested that it is this kind of knowledge that is generally presupposed in public discourse, although it may be explicitly expressed and thus taught to new or still incompetent social actors, such as children and members of other cultural groups, for instance immigrants.

Sometimes Common Ground may be co-extensive with the knowledge shared by the users of a specific language, but the social basis of Common Ground beliefs is usually broader, and may include several languages. Thus, in many respects, the shared knowledge of people in Western Europe, North America and South America transcends individual countries and different languages. If necessary, one could make a further distinction between different kinds of "national knowledge" as it is learned and presupposed by public discourse in different countries.

Common Ground knowledge is by definition shared by all members of a culture, and in that respect also defines what is common sense (Kuipers & Mackor, 1995; Purckhardt, 1993). Note though that this inclusion is temporal: There is common ground knowledge today that was considered 'mere belief' of some group (scholars, women, ecologists, pacifists, etc) yesterday, and vice versa, there are now group beliefs (like religions) that were rock-solid part of the Common Ground in the past. In other words, cultural knowledge is constantly, historically changing. This also implies that what is commonly accepted as knowledge today, in one culture, may be considered mere belief tomorrow, or in another culture. The only criterion for beliefs to constitute knowledge in a culture is that such beliefs are generally accepted to be true by all competent members. Each culture, and each group, however, develops 'methods' or criteria that are used to establish knowledge or to account for it. Scientific and specialized discourses are characteristic examples that presuppose or account for such 'methods'.

Common ground beliefs are not limited to knowledge but may also feature evaluative shared beliefs, that is, social opinions or attitudes. Usually there are

differences of opinion about attitudes, and for this reason attitudes usually differ between groups. However, there is no a priori reason why some attitudes may not be shared by a whole culture.

Within the same perspective, finally, Common Ground beliefs feature norms and values, which precisely form the evaluation criteria for what members of a culture hold to be good or bad, and what should (not) be done. Part of these norms and values are also the very knowledge criteria, namely what should be considered good grounds to accept beliefs as knowledge. As is the case for all Common Ground beliefs, also these criteria may change – thus, today they are rather scientific than religious.

We have no idea about the actual size of the Common Ground. One crucial criterion is that members of the culture are able to learn its contents in a reasonable time. Obligatory schooling of some 12+ years is a rough measure, but of course the curriculum is only a first indication of the content of Common Ground knowledge, because children probably learn much more from their parents, from TV, computers, internet and from peers than at school. One might speculate about a size in the order of hundreds of thousands of ‘facts’, of which the lexicon of one language alone constitutes some tens of thousands of word meanings. Another criterion is strategic accessibility – the knowledge must be ordered in such a way that it provides virtually immediate access (measured in milliseconds) for everyday use in discourse processing and other social practices.

I already suggested that despite the vast amount of work on knowledge in epistemology, cognitive science and the social sciences, we have only very little insight into the details of Common Ground knowledge and other cultural beliefs, and how they are being acquired, changed, stored, used, forgotten, and applied. In this paper we are of course unable to even begin to solve this huge task. Relative to the next parts of this paper however, we may say that in the same way as genetics and biology were immensely advanced by major projects such as the Human Genome Project, the psychology of knowledge would be enormously advanced by a similarly ambitious mega-project on the Human Episteme – a large empirical study of the contents and structure of the shared knowledge of a given culture – e.g. by knowledge interviews and systematic analysis of the presuppositions of public discourse. In a much more limited sense, such studies have been undertaken by anthropologists, mostly of non-western cultures, on the one hand (see, e.g. Shore 1996), and by ‘knowledge engineers’ in Artificial Intelligence interested in, for instance, the specialized knowledge of experts (Becker & Selman, 1986; Boose, 1988, 1990; Wielenga, 1990). Typically, these different fields of knowledge study mutually ignore each other.

## **Group knowledge**

Scientific and other forms of specialized knowledge are defined here as types of Group Knowledge, that is, knowledge shared by the (competent) members of a group. The precise sociological or social psychological definition of what constitutes a ‘group’ is a complex problem beyond the scope of this paper, but obviously we are dealing with a collectivity of people who have something in common, such as their origin, language,

education, skills or indeed special knowledge, as well as the social practices based on these shared resources, including shared aims, norms, and values. In this sense, we may usually speak of a *community* if the group has some kind of social continuity. I shall therefore also use the terms ‘knowledge community’ or ‘epistemic community’ for those communities who share the same kind of group knowledge. Some groups may be further organized in very complex ways, and feature leaders, institutions, and procedures to become members of such groups, as is the case for political parties and churches, among others. Indeed, in my theory of ideology, I assume that all social representations shared by the members of a group, including knowledge as well as attitudes have a shared underlying group ideology that represents so to speak the self-schema of a group (Van Dijk, 1998).

Group knowledge is based on, and derived from Common Ground knowledge. This seems a very strong thesis, for it does not seem immediately obvious, for instance, that all scientific knowledge is derived from everyday, common sense knowledge. On the contrary, much scholarship precisely is geared towards the demonstration that Common Ground knowledge is often wrong or misguided. Yet, in an important sense, also specialist knowledge is constrained by the Common Ground beliefs of a culture, as the history of scholarly knowledge amply demonstrates – including by the prejudices that often remain alive in scholarly discourse, for instance about women or Blacks until quite recently (Chase, 1975; Van Dijk, 1993). Often the relation between group knowledge and Common Ground knowledge is indirect, since technical terms of scholarly groups may themselves be defined in other technical terms that are unknown in the Common Ground. However, for newcomers in each group to be able to learn the specialized knowledge contents, as well as the specialized discourse of a group, they have no other knowledge than that of the relevant Common Ground, so that specialized knowledge is ultimately built on Common Ground beliefs – and even if such Common Ground beliefs may later be partly corrected or ‘specialized’ by group knowledge. In other words, it is the acquisition of Group Beliefs that is the most obvious argument for the Common Ground basis of Group Beliefs.

Group knowledge is not necessarily true in at least one sense of the term. As we have seen above, group knowledge may well be false or otherwise misguided, and be called ‘beliefs’, opinions, illusions or prejudices by members of other groups. Thus, even male chauvinist opinions about women may well be considered ‘true’ by male chauvinists. Crucial though for such knowledge to be accepted as such is that the competent members of the group itself accept such beliefs as ‘true’ or otherwise validated by the validation criteria of the group. Obviously, these are different for religious, political and scientific groups.

## **Specialized knowledge**

Specialized knowledge, such as academic, scholarly, scientific, technical and other kinds of ‘expert’ knowledge which require specialized education or training, is a type of group knowledge as defined above. That is, it is acquired, shared and used among members of various kinds of scholarly or other specialized communities, defined not

only by their specialized knowledge, but also by their specialized social practices, including their specialized discourse and communication, as well as by a complex network of organizations and institutions such as universities, laboratories and associations. For this paper, perhaps the most interesting property of these communities is that their principal aim is the production of new knowledge –besides the acquisition of new skills or other specialized practices. At first this knowledge may still be specialized, but at least part of it is assumed to be shared by others in society, such as journalists, teachers and others who distribute and popularize specialized knowledge (Ackermann, Anton & McCourt, 1995; Fayard, 1993; Gregory & Miller, 1998; Myers, 1990; Scanlon, 1999; Stocklmayer, 2001).

The first characterization of specialized knowledge just given has a sociological nature – it is not defined by its own characteristics, but in terms of those who are supposed to possess and use such knowledge: the experts. Of course, much further sociological as well as anthropological analysis of this kind of ‘knowledge communities’ is necessary, continuing the extensive work already done in the sociology of science (Barnes, 1982). Our attempt to characterize specialized discourse may be seen as one of the endeavors to provide such multidisciplinary account of specialized knowledge.

However, before we are able to be more explicit about specialized discourse we need to have more insight into the nature of specialized knowledge itself because it is on the basis of specialized knowledge that we need to define the many varieties of expert text and talk. Obviously, such questions need to be formulated (also, though not exclusively!) in cognitive terms: Specialized knowledge consists of social representations shared by communities as defined above. What do these representations look like? Are they different from the kind of representations that make up our general, sociocultural knowledge of the Common Ground?

Since there does not seem to be a priori replies to these questions, they can only be resolved by empirical study of the expressions and uses of specialized knowledge, as we shall also do below. However, we might formulate some theoretical ideas about the nature of specialized knowledge, so that our empirical research is more than a wild goose chase. We have already assumed that specialized knowledge (ultimately) is based on and derived from everyday knowledge, for the obvious reasons that it can only be acquired on the basis of what people already know: First definitions of expert terminology can only be understood in everyday terms. Of course, expert knowledge may then proceed to show that our original (non-expert) knowledge was misguided, but this again needs to be done in expert terminology which itself is ultimately based on everyday language. The same is true for the acquisition of knowledge based on experiences, skills and expert action and interaction, e.g., in laboratories, which also need to be acquired on the basis of everyday skills of observation or manipulating objects.

In other words, if only for reasons of theoretical parsimony, we may assume that cognitively speaking expert knowledge is probably not fundamentally different from other kinds of socially shared knowledge. Obviously, in terms of the objects about which we have expert knowledge, it is safe to assume also that such knowledge is more

extensive: a biologist simply knows much more about trees, or cells or genes than non-experts. Interesting to know is what this 'much more' usually amounts to, such as components, structures, causes and consequences, and details of many of the processes involved.

For instance, language users are intimately familiar with language: They use it all the time, and in many respects they may be called 'experts' in language use, including extremely refined sensitivity for the most subtle variations of meaning and action. Yet, despite years of language education in schools, lay language users have very little specialized knowledge about language or discourse, beyond the meaning of words, and some implicit knowledge about word order. Specialized linguistic knowledge consists first of all of knowledge about the various levels or dimensions of language or discourse, such as sound, sentence structure, meaning, speech acts, or turn-taking, as accounted for by different sub-theories, such as phonology, syntax, semantics or pragmatics, among others. Then for each of these specialized sub-fields, characteristic 'units' are distinguished, as well as the relations between them, such as phonemes and phonological features, phrase structure, propositions and their internal structure, or speech acts and their appropriateness conditions. Beyond these various forms of increasingly detailed 'structural' analyses, expert linguistic description typically engages in more 'functional' analyses: what do these various structures 'do' in their specific environment, both within the structure itself (e.g., what is the 'function' of pronouns in a text, or the function of headlines in a news article), as well as in terms of its 'context': its users as well as the actions or cognitions accomplished by them. Finally, expert language and discourse knowledge amounts to detailed insight into the psychology of language, such as its acquisition, actual use in production and comprehension, and the relationships of those psychological processes with other cognitive processes, e.g. those of learning and memory. Similar specialized knowledge on language may be provided in a sociological account of the way language is used in social situations, by members of different social groups, and in order to accomplish a host of social aims.

The point of this for linguists all too familiar summary of some of the aspects of the expert linguistic knowledge is to have a first approximation of the nature of expert knowledge as different from everyday lay knowledge. Obviously, at least for language, such expert knowledge consists of a clear form of organization of the field and object of knowledge, e.g. the knowledge that 'meaning' is one aspect of language and that such meaning is studied in semantics. Of course, there are disputes about many borderline cases, for instance, about the study of various forms of action and interaction, which is variously studied as part of meaning, or separately in theories of speech acts, or pragmatics or conversation analysis or sociolinguistics or the ethnography of speaking. The same confusion exists for the separation between meaning and cognition, and hence between semantics and a cognitive account. The point is however (also of the disputes) to get some 'order' in the object of study, and hence some order in the field, e.g., in terms of partial theories or sub-fields, which soon may constitute their own areas of specialization, such as phonology, discourse analysis or psycholinguistics. Specialized knowledge, thus, first of all, consists of specialized knowledge about the conventional



‘parts’ or ‘dimensions’ of the main object of study (language, discourse, cells, genes, etc.) and the fields or sub-disciplines that study them.

Secondly, these various aspects or dimensions are systematically (structurally) analyzed in terms of units, which are themselves related in various ways, for instance how they combine or connect in larger units, and how thus complex hierarchies of units can be distinguished. Rules or laws may then be formulated for the combinatory possibilities, as is the case for all kinds of grammars or the rules of argumentation or storytelling.

Thirdly, once these units and their complex organization have been identified, their many functions are examined, defined either in terms of functions in complex structures, or in terms of various kinds of environments (contexts), for instance the cognitive or social contexts of language, or the body organs of the cells studied in biology.

Although systematic meta-analysis of the discourse and knowledge of all specialized areas would be necessary to get a more detailed picture, it seems warranted to conclude that specialized knowledge has some quite general properties, such as the following:

- a. Overall organization of object => Overall organization of the field
- b. Structural analysis of each aspect, dimension of object: Units, combinatorial possibilities, processes, rules
- c. Functional analysis of units and structures: What are they for?
- d. Contextual analysis: Causes, consequences, explanations of functions, etc.

This is only a very rough first analysis of the kinds of specialized knowledge one may encounter, and how such knowledge is organized. Abstracting even further, we see that there are a number of very general principles involved, namely those of Structure and Function, or those of Structure-Function on the one hand, and those of Function-Process on the other hand, as we also know from the pervasive distinction between text and context in discourse studies, or between organisms and their environment in biology. All these principles themselves suggest forms of order and its explanation or finality: *why* things are organized in that way. In other words, expert knowledge is first of all characterized by extra-ordinary forms of order or organization, using a large number of levels or dimensions of observation, analysis, units and functions not usually distinguished in everyday life –for these same objects of study or practice.

We have reason to assume that this kind of abstract order, as reflected in disciplinary discourse and organization is also the kind of structure associated with the mental representations of specialized knowledge. Indeed, crucial of specialized learning and education is that newcomers learn the fundamental organizational principles of this specialized knowledge and of the discipline. Such knowledge about the structure of specialized knowledge is also crucial for practical reasons such as cooperation, teamwork, referrals, hiring and the organization of departments and laboratories.

Note however that order, organization, functional analysis, contextualization or specialization are not limited to expert knowledge, but very general properties of thought and cognition. Also in everyday life, we make distinctions between different

aspects or sectors of nature or society, and we organize our everyday life accordingly. We know that trees belong to nature, and are able to make a very rough ‘structural’ analysis of trees in terms of their trunk, branches and leaves, and may distinguish between different sizes, forms and colors of leaves, and know that trees often form bigger units like woods, and we have some modest lay knowledge about trees needing light and water and producing oxygen. That is, we know about global organization, local structures, various kinds of functions as well as aspects of the environment. This is even more the case in situations of everyday life, for instance for the organization of the home and its many functions. In other words, specialized knowledge also in this respect builds on very general principles of understanding and explanation of objects, events and processes in every life. We may need to learn the specific principles of the organization of an object or discipline, but the overall ordering principles themselves are already pre-programmed for the organization of all cognition, and hence not specific. In other words, our hypotheses on the organization of specialized knowledge are consistent with what we know about the organization of knowledge and cognition in general.

### **From specialized knowledge to specialized discourse**

There seems to be a straightforward relation between specialized knowledge and specialized discourse if the latter is merely defined as the verbal expression of the former. In this case the only requirement would be that we ‘translate’ the propositions of specialized information or knowledge into those of meanings and that such meanings are given as input to the rules and strategies of ‘formulation’ of grammar.

Things are not that simple, however. First of all, the epistemic organization of the mind is not the same as the organization of discourse, which follows its own rules, and obeys its own functional requirements in processes of communication and interaction. In other words, the fundamental functional differences of knowledge and discourse require very different structures. To wit, discourse –whether spoken or written—is basically linear or sequential, whereas knowledge structures are probably hierarchical and network-like. This means that epistemic structures need to be ‘linearized’ or otherwise transformed when used in discourse. Secondly, both everyday as well as expert knowledge is not always expressed directly in discourse, for instance as general, abstract (generic) propositions, but often instantiated, particularized or ‘applied’ in the representation of specific events, cases or situations. Thus, general principles of grammar may be ‘applied’ in the acquisition or description of grammars of particular languages, and even these particular, but still general grammars may be further applied in the description of specific utterances. Similarly, general genetic knowledge about cells, chromosomes, genes and their internal DNA structure may be applied to more specific, but still general, knowledge about the human genome, and that knowledge may be particularized in the description of (parts of) genome of families or individuals, for instance in the explanation of genetic disease.

We have seen that this step from the general and abstract to the specific and individual is cognitively accounted for in terms of mental models of events or

situations. Such mental models feature a specific selection of particularized general knowledge items, as well as unique descriptions of events, situations and their participants, such as the experiences or interactions of a specific individual. Specialized discourse may thus express both general knowledge as well as its specific application in concrete, particular 'cases' (situations, events, objects, persons, etc.). In the latter case, discourse often takes the form of stories (everyday conversational stories, or news stories), whereas specialized discourse in this case takes the form of case studies and its associated 'stories' and its own 'methods' (life histories, interviews, participant observations, etc.). Characteristic of such stories is not only that general expressions (variables) are replaced by specific ones (constants, names), but also that they feature (unique) combinations of general knowledge that is usually organized or distributed differently. Thus, in a linguistic case study, we may combine (the processing or uses of) phonological, syntactic, semantic, pragmatic, rhetorical or interactional structures and their interrelations in which that is not immediately obvious at a more general level. Similarly, in genetics, case studies of specific diseases may need a combined application of cell biology, genetic analysis, physiology, neurology, and a variety of medical specializations.

### **Specialized contexts**

Discourse however requires more than this kind of model-based specialization or particularization of knowledge in order to represent concrete events or 'cases', and needs to be structured also following the basic principles of functionality mentioned above, and as represented by the relevant constraints of communicative events and situations, that is, by context. It is only through an explicit account of such contexts that we are able to spell out the ways texts are adapted to their environment, for instance in terms of its aims, functions, conditions, consequences or other aspects of the communicative situation (Duranti, 1992; Van Dijk, 1999, 2002). That is, also expert discourse fundamentally requires an account of its many modes of contextualization: Who uses it, when, where, with whom and with what aims. This adaptation requires many specialized variable structures, of which intonation patterns (stress, volume), deictic expressions, word order, politeness forms, hedges, disclaimers, and in general (lexical and other) style and rhetoric are only some examples.

In other words, a theory of specialized discourse requires much more than a cognitive theory of specialized knowledge expressions, either in generic terms or in terms of the combinations of knowledge in mental models. It also requires a cognitive theory of context models that accounts for language users' ongoing interpretation and management of the current communicative event, as well as a sociocultural account of communicative events and the relevance of its variable properties in the accomplishment of discourse. In other words, a theory of specialized discourse combines a theory of specialized texts with a theory of specialized contexts. And since (specialized) participants, or experts, are part of such a context, also their specialized knowledge is theoretically part of this account of specialized contexts. This 'contextualization' of the important knowledge component of specialized discourse has

the considerable advantage that it only focuses on the *relevant* aspects of such knowledge. Also experts need not activate their vast amounts of general knowledge, but only small fragments that are relevant in the understanding or description of specific problems, cases or details.

The theory of specialized contexts is still on the agenda (see however the studies in the sociology of science that might be used as empirical suggestions for such a theory of context, e.g. Latour, 1987; Latour & Woolgar, 1986), but we may briefly mention some of its most obvious features, such as the categories that make up the schemas expert use in the definition of the many context types in which their expert discourses have their specialized function: writing and reading monographs and scholarly articles, observations and analyses, giving lessons and lectures, holding press conferences or giving interviews, and so on. These categories involve for instance the overall domain of the current communicative event (e.g., research, education or health care), the overall action(s) being accomplished (e.g., investigate, teach, etc.), the current Setting (time, location, circumstances), the specific actions involved (hold a seminar, give a lecture, have a research meeting), the participants and their various communicative, social and professional roles, and their aims, interests, and especially knowledge and opinions.

These and other categories are not just a luxury of an extended, contextualized theory of discourse, but crucial in the account of how many specialized and other discourse structures are adapted to the communicative situations of any field of specialization. Thus, in order to be able to write an adequate textbook for their field of specialization, experts need to know that the global action is rather teaching than investigation, that the current action may be to write a chapter or questions for a textbook rather than for a scholarly seminar (involving e.g., didactic questions whose answers are already known), who the recipients are (e.g., students) and what their current specialized or non-specialized knowledge, interests and aims are. The overall organization of the text, its global contents or topics, its lexical and other aspects of style, as well as its speech acts, and other typical characteristics (definitions, tables, illustrations, metaphors and comparisons etc), are all geared towards one overall goal: teach specialized knowledge to those who haven't got such knowledge as yet.

It may be obvious from this brief argument that the analysis of specialized language goes way beyond even an extension from specialized vocabulary or terminology, grammar and style, to the many discourse structures of specialized communication (summaries, abstracts, titles, overall canonical schemata, non-verbal structures –tables, figures, drawing, film, etc.), and also includes a complex theory of context structures and text-context relations.

We are now approaching an overall, multidisciplinary theory of specialized discourse that probably has some future in an adequate description and explanation of this kind of text and talk. This theory has four main components, of which three have been briefly described above:

1. A theory of the (linguistic and other) structures of specialized discourse.
2. A theory of specialized knowledge and its relations to other kinds of knowledge and social cognition.
3. A theory of specialized contexts and text-context relations.

4. A theory of the social and cultural structures that form the broader 'context' of expert discourse and communication.

The social and cultural theories will not further be detailed here, also because they have received ample attention in recent studies of science, such as the constitution of expert communities; the organization of universities and laboratories as well as the everyday interactions within them; the role of international conferences, and the relationships of power and dominance and how specific forms of knowledge in society are more legitimate than others, and so on.

It should however be stressed that large part of the specialized discourse structures are ultimately to be explained in terms of this more global 'context' features as defined by universities, laboratories, and scientific conferences. That is, the communities of experts are not only knowledge communities and special discourse communities, but also social communities, and much of this social dimension required analysis in terms of specialized talk and text: Universities, laboratories, conferences, etc. are largely discursive. They thus help contextually to constitute (and legitimate) specialized discourse and at the same time are constituted by it.

### **Knowledge and discourse in modern genetics**

In the remainder of this paper, we shall apply and further elaborate the overall framework sketched above in a more detailed study of some fragments of the discourse of modern genetics, as exemplified by a widely used textbook, *Genética Moderna*, by Griffiths, Gelbart, Miller and Lewontin (Madrid, etc., McGraw-Hill Interamericana, 2000).

Contemporary genetics not only constitutes an interesting example of the fast developments of expert knowledge in a literally vital area at the boundaries of various life sciences, neurosciences and cognitive sciences. It also has developed a discourse that is especially fascinating for linguists and discourse analysts because it is replete with linguistic metaphors: language, codes, expressions, translation, etc. One might even speculate about the possibility that modern genetics in turn might provide metaphorical suggestions for theory formation on the structure and discursive expression of expert knowledge, an exercise that however is beyond the scope of this paper.

Although genetics is more than one century old, and rooted in the plant experiments of Mendel in the 19<sup>th</sup> century, it is only in the last decades that vast and fast advances have been made in the structural and functional analysis of genes and their role in the reproduction and management of cells and their functions. It is only a few decades ago that the double helix structure of chromosomes was discovered by Crick and Watson, whereas the detailed 'sequencing' of the complex base structure of genes and their ADN has only been taking place in the last decade, finally culminating in the huge Human Genome Project that completed its first major stage on June 26, 2001: the description of all tens of thousands of genes of the human genome (Boon, 2002; Dennis, 2001). That is, the first, 'structural' phase of modern genetics has now been reached, although complete genetic analysis of all living organisms is still on the

agenda for the future. The next major challenge is to spell out the complex functions of these tens of thousands of genes in the control of the construction of the vast amounts of proteins that make up the cells of our body, especially in view of being able to prevent or repair genetic dysfunctions in many diseases.

It is this (very roughly characterized) framework of contemporary genetics that constitutes the global background for contemporary texts on genetics. Also the textbook we shall use as example witnesses this history and current state of the field as well as its breathtaking recent developments within the Human Genome Project.

The aim of the remainder of this paper is to describe in some detail how specialized genetic knowledge is expressed or presupposed in the specialized discourse of a textbook. We use a textbook as example because it is precisely in such types of didactic discourse that we are able to observe in more explicit detail the construction of specialized knowledge. The idea is that students who read this text have little knowledge of genetics as yet, and that throughout the book they slowly build up the canonical minimum knowledge required for students of (among other disciplines) biology and medicine. Of course, this does not yet make them experts in genetics, but the textbook itself is organized and written by experts in genetics, and should show how their expert knowledge is organized, and especially how they think students should organize their knowledge about genetics. That is, textbooks like this also have normative characteristics, and thus show how experts make explicit the norms, values and principles underlying their knowledge and its organization, and how students *should* think about genetic issues.

We shall be relatively brief and informal about –otherwise crucial—issues of discourse genre: Textbooks are a prominent genre by themselves, and in need of detailed formal, semantic, pragmatic and semiotic analysis (Apple & Christian-Smith, 1991; Britton, Woodward & Binkley, 1993; see also Lemke, 1990). Pragmatically speaking, textbooks emphasize the role of the transmission of knowledge, since as global speech acts they count as assertions, which presuppose that the writer(s) assume that the reader(s) do not yet know the content of the text. In a broader contextual analysis, as suggested above, one would thus assume that textbooks belong to the genre of pedagogic discourses, and hence to the overall context domain of education.

Since this textbook we analyze is on genetics, the domain is more specifically that of scientific education, combining science with education. As for the (production) setting, it only is contextually relevant that it is published in 2000, so that we know what ‘modern’ in the title implies. The publisher (McGraw-Hill) will be recognized as well-known publisher of textbooks, and the authors as well-known geneticists and authors of textbooks. The contextual aim of the authors is obvious not only from the genre, but also from the text itself: they aim to give a first introduction to modern genetics. More particularly (p. xv) they want to focus on DNA analysis first, rather than to give a traditional historical account of the development of genetics (typically beginning with Mendel, as they still did in their own earlier book on genetic analysis). The extensive Preface of the book spells out in more detail what the specific aims and characteristics of the book are, and who are addressed as users and readers. Finally, the context model of the authors should feature more or less detailed ideas about the

knowledge of the readers, and we shall see in a moment what these presuppositions amount to. Obviously, the presupposed knowledge is more than just general Common Ground, because there are many technical notions that are supposed to be known already, and in this respect this text on genetics presupposes earlier knowledge on biology or medicine, of at least high school level plus or one or two years of college.

In more formal terms, the textbooks shows the familiar organization in chapters (there are 18 chapters in some 600 pages), and as is the case for modern textbooks, it is richly illustrated by numerous photos and drawings, a multimodal presentation that is essential for education in modern genetics (see also Kress, 2001). Also the superstructure of the book is typical for modern textbooks, especially those in the USA, of which the current one is a translation. The book has two initial tables of contents, one just listing the chapters and their titles, and another one, following it, with more detailed information about the contents of the chapters. A detailed Preface presents the book, its authors and its aims, followed by some literature. Because this kind of major text requires collaboration and evaluation by many people, there is a fairly large list of Acknowledgements following the Preface. A glossary, recommended literature, solution to problems and a detailed index close the book. Each chapter begins with a brief point-by-point overview, and is closed by a summary, a map of concepts and various types of questions and assignments, all designed to optimize the learning process, as we also know from the contemporary psychology of text processing, which confirms the vital role of prior or posterior macro-organization of text for memory.

There is no standard way to do an ‘epistemic’ analysis of discourse. Knowledge may be expressed, signaled or presupposed in many ways. Given the aims of this paper, then, I’ll focus on the interplay between Common Ground knowledge and specialized knowledge at several levels, and especially at the level of semantics and the ‘pragmatic’ expression of context. Such a ‘contextual’ analysis is relevant not so much because it expresses the kinds of specialized knowledge itself, but rather because it signals properties of the communicative situation that – in the context models of the authors – are relevant for the understanding and the uses of this textbook. Indeed, many of these properties define the very genre of textbooks, and are part of the definition of specialized discourse as a class of genres. That is, it may well be the fact that a text is written by experts and for experts-to-be, and with the obvious knowledge these have, rather than any special property of the text itself, that makes it a type of specialized discourse.

## **Introducing a field**

Textbooks introduce disciplines, areas, fields or important scholarly topics, and it is therefore important that such introductions go beyond mere communication of specialized knowledge, but relate the new field to other specialized fields, and indicate why the study of such a field is relevant, for instance because of its practical applications. Because textbooks are generally addressed to students (and authors have such knowledge stored in their context models of this kind of communicative events),

they first need to provide ‘motivation’ by spelling out why the new field is interesting or important for the readers.

This is indeed what the first Chapter (“Genética y ser vivo”) aims to do: Present the field of genetics, why it is important, and what it is all about:

Por qué estudiar Genética? Hay dos razones básicas. En primer lugar, la Genética ocupa hoy la posición de pivote sobre que gira toda la Biología, y, por tanto, resulta esencial para cualquier estudioso serio de la vida animal, vegetal o microbiana. En segundo lugar, la Genética ha llegado a tener un protagonismo central, como no tiene ninguna otra disciplina científica, en numerosos aspectos de la vida del hombre. Afecta a nuestra humanidad de formas muy diversas. Ciertamente los temas genéticos afloran en nuestras vidas casi diariamente, y ninguna persona instruida puede permitirse ignorar sus descubrimientos. En este capítulo haremos una revisión general de la ciencia de la Genética, mostrando como ha llegado a ocupar una posición tan crucial. Ofrecemos, además, un panorama general en el que poder encajar los siguientes capítulos (p. 2).

We see that this opening paragraph of the first chapter of the textbook (after various general remarks about the discipline in the Preface) is especially of a persuasive nature, in the sense of giving arguments why genetics is worthy of study, and providing some first indications of the place of genetics in the sciences as well as in everyday life. These arguments may well be based on facts, but they have a clear evaluative dimension, and hence may also be categorized as opinions, for instance when the authors claim the primacy of genetics in the study of humanity – opinions that are further sustained in the rest of this chapter. Relevant for our own analysis is the fact that within the context-bound (while student-addressed) parts of this first fragment of the first chapter, the specialized knowledge expression is about the overall place and relevance of the discipline in the sciences and human life. That is, students are not only motivated to learn about genetics, and hence to read and use this book (a contextual concern expressed also by references to “estudioso serio” and “ninguna persona instruida”), but also are indoctrinated with some basic attitudes and ideologies about the (superiority, priority of) the field. Knowledge communication is thus not the only aim of a textbook that introduces a new field. Note also that no specialized knowledge is required for the comprehension of this first paragraph. It does not contain technical terms, and the knowledge presupposed here is regular part of readers’ Common Ground knowledge:

*Expression*

Estudiar genética  
Biología  
Study  
Sus descubrimientos  
Etc.

*Common knowledge presupposed*

Genetics is an academic discipline  
Discipline that studies life  
People have reasons to study something  
Science makes discoveries



## Defining a field and its basic concepts

Once mentioned the relevance of the discipline for other disciplines and for everyday life in general, a textbook typically needs to define the field, also in order to distinguish it from other fields. Definitions not only are a prominent and quite typical semantic strategy of specialized discourse, for instance, in the introduction of new terms, but also as a means to delimitate the field and its research, in relation to neighboring fields (Grize, 1981; Moirand, 1999; Calsamiglia & Van Dijk, 2004):

(1) Primero debemos definir la Genética. Algunos la definen como el estudio de la herencia, pero los fenómenos hereditarios han llamado la atención de los seres humanos desde los albores de la civilización. Mucho antes de que existiera la Biología o la Genética como disciplinas científicas que hoy conocemos, los pueblos antiguos mejoraban plantas cultivables y animales domésticos, escogiendo para cruzarlos aquellos individuos de características más deseables. (...) Pero entendido como un conjunto de postulados y de métodos analíticos, la Genética no comenzó hasta la década 1860 cuando un monje agustino llamado Gregor Mendel (figura 1-1) realizó una serie de experimentos que apuntaban a la existencia de unos elementos biológicos llamados **genes**. La palabra Genética deriva de “genes”, y son estos los que constituyen el epicentro del tema. Para los genéticos, los genes son siempre el objeto central de sus estudios, realicen estos al nivel molecular, celular, organismo, familiar, poblacional o evaluativo. Dicho en pocas palabras, la Genética es el estudio de los genes. (p. 2)

This paragraph combines several properties of specialized discourse: it defines and delimits a field of study, it gives a succinct ‘proto-history’ (“los albores de la civilización”), marks the beginning of the ‘modern’ discipline, and at the same time introduces the founding father (Mendel), and provides a first description of what genes are: “elementos biológicos” before providing a more detailed description and definition of genes in the next paragraph. Implicitly it also polemicalizes with other geneticists (“algunos definen...”) who define genetics in terms of heredity, but at the same time shows that questions of heredity need explanation at a more elementary level, namely that of genes.

Again, this paragraph hardly shows technical terminology, and all information can be readily understood in terms of Common Ground knowledge. If any, it introduces the notion of gene, vaguely defined as a ‘biological element’ – presupposing that the readers make a possible association with chemical elements, and hence with some kind of ‘building block’. Any structural analysis, also in genetics, needs to identify its basic building blocks, and genes are thus introduced and defined as the core ‘elements’ studied in the discipline. Of course, this is merely a first, still very primitive, step, because surely geneticists do more than merely study genes.

Another informally introduced aspect of field knowledge is the series of levels or scopes usually applied in genetics and biology, running from the smaller or more elementary to the bigger or macro: cells, organisms, families, populations and evolution. The culminating definition, at the end of the paragraph, is merely trivial in this respect, like defining linguistics as the study of language.

Interesting for our analysis in this paragraph has been first of all the fact that the first paragraphs of technical textbooks begin in non-technical terminology that is readily comprehensible for beginners. Secondly, that first descriptions or definitions of

technical terms (like ‘genes’) may already be given, but only in very informal, non-technical terms. Thirdly, the knowledge conveyed in this case pertains to different kinds of aspects of a discipline (as was the case for the previous paragraph): protohistory, modern history, founding father, misconceptions, and alternative (misguided) definitions, as well as the levels or scopes at which genetics may be at work. In other words and as we have stressed before, specialized knowledge is not limited to complex structural information about specialized objects or phenomena, but may involve many other aspects of a discipline: social and scientific relevance and history.

Of course, academic textbooks of science do not remain at this general, understandable level of terminology. The next paragraph therefore immediately goes over to the more technical definition of genes:

(2) Qué son los genes? Los genes estan hechos de una macromolécula, trenzada en forma de helice doble, llamada **ácido desoxirribonucleico**, abreviadamente **DNA**. El DNA, el material hereditario que se transmite de una generación a la siguiente, dicta las características propias de una especie. La informacion esta codificada en el DNA en forma de una secuencia de subunidades químicas denominadas **nucleotidos**. Cada célula de un organismo contiene, típicamente, uno o dos copias de la dotación completa de DNA, llamada **genoma**. El propio genoma esta constituido por una o mas moléculas de DNA extraordinariamente largas, empaquetada cada una de ellas en una estructura denominada  **cromosoma**. Los genes son, simplemente, las unidades funcionales del DNA cromosómico.

After the first general information about genetics as a discipline, the textbook here proceeds to provide some first technical definitions and terminology – still at an elementary level of introductory biology – thus linking various key notions of genetics. Note that at this first level of definition, the describing part is formulated in structural terms (x is made of y, x consists of y, x has subunits y, x contains y, etc.). The first definition is not entirely formulated in Common Ground terms. For instance, presupposed is the notion of ‘macromolecule’, probably because a genetics course presupposes elementary secondary or college biology or chemistry. On the other hand, molecule may be one of those terms that used to be part of technical vocabulary, but that has entered the general lexicon, and since also ‘macro’ is part of that, even those who do not know about introductory biology or chemistry, would know more or less what a macromolecule is. Note also the standard (famous) metaphor in the description of the form of the macromolecule as a ‘double helix’ (other metaphors, like ‘corkscrew’, could have been used, but sound less poetic-scientific – an issue we shall further ignore here). Third, complex chemical names are usually abbreviated – and in this case the abbreviation (DNA) also has become standardized and famous, and used so often in everyday, media discourse that it also may be considered part of Common Ground knowledge (but only the abbreviation, not the full name). Also a first functional element is introduced – namely a description of what DNA does (transmits hereditary material to the next generation). The relevant parts of this (complex) definition are semantically marked by the metalinguistic expression “are called”, introducing a new term, which is also typographically marked in bold. Note that a sequence of definitions, like this one, also allows a further specification of an earlier definition. Thus, genes are first described as a macromolecule, and after the definition of a chromosome (a

definition built itself on the notions of ‘gene’ and ‘genome’), they can “simply” be defined as functional parts of chromosomes. This form of ‘simplification’ is important in scientific discourse and definitions, because it assigns further order to the system of expert knowledge thus construed, which is essential, both for experts and newcomers, to learn and daily manipulate the complex network of concepts that make up expert knowledge.

Note that the sequence of definitions given in (2) is merely a first step in the expression and didactic production of expert knowledge: the rest of the book provides much more information on the structures and functions of genes, chromosomes and their DNA. Indeed, the definition given here acts like a summary or macrostructure organizing a vast network of further definitions. Yet, even this initial ‘macro-definition’, as we may call it, already shows some basic properties of expert knowledge, such as knowledge about:

- (i) structure, composition (x is composed of y, x consists of y)
- (ii) form (x has the form of y)
- (iii) composition (x consist of y, x is made of y)
- (iv) functions (x has the function y, x ‘does’ y)
- (v) naming, terminology (x is called y)
- (vi) knowledge about abbreviations (x is abbreviated as y)
- (vii) conceptual hierarchies and simplifications (x is simply y)

Each of this elements of knowledge may be based on other elements of expert knowledge (x consists of y, which consists of z, etc.), or is rooted in the knowledge or experiences represented in Common Ground (‘in the form of a double helix’, ‘hereditary material transmitted from generation to the next’). Some knowledge seems to characterize the borderline between general and expert knowledge, such as ‘macromolecules’ and ‘sequence of chemical subunits’. This suggests that distinction between general, sociocultural knowledge and expert knowledge is fuzzy, and that some very general specialized knowledge may already have entered the Common Ground. We might introduce a new term for this kind of ‘semi-technical’ knowledge that is presupposed in textbooks and in popularizing articles in the press, and which usually derives from secondary education or (earlier) popularizing articles in the mass media. The criterion here is that such notions are not defined in the textbook, and probably would not be defined in public discourse, e.g., of the media, either. Note also that the general properties of expert knowledge as listed in (i-vii) are also part of the Common Ground: the concepts of structure, form, composition, naming etc. are not as such technical notions. Sometimes Common Ground notions are used that are difficult to interpret by newcomers. For instance, the definition speaks about “the information is coded...”, but it does not explain here what kind of ‘information’ is involved here, or whether this is merely one of the typical metaphors of genetics (Keller, 1995; Rothbart, 1997).

A sequence of definitions expresses part of expert knowledge, but this does not mean that didactically such a sequence is comprehensible for beginners. It is not unlikely, for instance, that the sequence expressed in (2), needs other didactic means in order to be fully understood and anchored in the new expert knowledge of a beginner,

for instance with a drawing or schema, as is also the case here: A sequence of drawings running from macro (organism: human being) on the left, to the double helix of DNA on the right at the micro level, going through cells, cell nuclei, chromosomes, and genes. Indeed, as is the case for discourse comprehension in general, also experts need general schemata that ‘figuratively’ represent the relations between different concepts.

The textbook continues its first definitions in the same paragraph, by explaining that genes basically control cell (re)production and especially the production of proteins, through DNA-copying (‘transcription’) into messenger-DNA, which is then ‘translated’ into polypeptides which in turn form the proteins, which define the properties of cells, which finally define the properties of the organism. That is, we witness the same sequence of structural and functional descriptions as we have analyzed above, introducing new terms (marked in bold), often of a metaphorical nature. By thus going back and forth between elementary structures (genes, DNA) and their functions and higher level constructs (proteins, cells, organism) we again witness the way expert knowledge organizes and at the same time explains the world, and thus reduces its complexity and mystery. Genetics is thus not just expert knowledge about genes, chromosomes and cells, but also designed, formulated and developed with explanatory aims: Genes explain heredity – why next generations are similar or different from earlier generations.

### **The relevance of a discipline**

The reason why the definitions given above are so brief and still very little technical, and doing no more than introducing a few basic notions, becomes clear in the next sections of the chapter, which deal with the role of genetics in human affairs. As we may recall, this is the overall topic of this chapter, and the technical definitions just provide the minimum technical knowledge needed to explain why genetics is relevant for humankind – which is an essential component of the persuasive (context-dependent) aims of making a discipline attractive for students. The discourse in this case reverts to presupposing general Common Ground knowledge and experience, namely about our clothes, food and their origin:

(3) Eche un vistazo a la ropa que lleva. El algodón de su camisa y de sus pantalones procede de plantas de algodón que difieren de sus ancestros naturales por haber sufrido un intenso programa de mejora, consistente en la aplicación metódica de postulados genéticos estándares. Lo mismo podría decirse de la oveja que produjo la lana de su chaleco o de su abrigo. Piense también en su comida más reciente. Puede estar seguro de que el arroz, el trigo, el pollo, la ternera, el cerdo y el resto de los principales organismos que sirven de alimento a los seres humanos en este planeta y animales, mediante la aplicación de métodos genéticos estándares. (p.3).

Still, although formulated in Common Ground terms, there is still formulation of what might for at least some students constitute new knowledge: That cotton and sheep have been genetically ‘improved’ since earlier generations. Fragments such as (3) are typical of didactic texts, especially because they anchor the aims and reasons of a discipline in everyday life – that is, they define the social relevance of a discipline, and

hence also its scholarly importance. Even the style of this fragment tries to imitate that of ordinary discourse ('eche un vistazo', 'piense tambien', etc.), and for the first time addresses the student(s) directly, an important contextual feature. The rest of the chapter similarly switches between some further technical information (e.g., about variation, phenotypes and genotypes, etc.) and more general statements about its relevance, for instance in the improvement of plants or foods, in medicine as well as in biology. Again, although formulated largely in non-technical terms, these parts of the book do express and teach knowledge, namely about the many applications of the field – including about well-known cloning experiments such as that of lamb Dolly, as well as about such well-known diseases as AIDS. It is only after the general introductory chapter that explains the relevance of genetics in everyday life that we find a chapter that focuses on the much more detailed structures of gene and genome.

### Teaching knowledge about structure

The systematic construction of specialized knowledge in textbooks often proceeds by teaching about basic units and their mutual relationships. As soon as genetics is roughly defined as the science that deals with the structures and functions of genes, the more technical chapters need to detail what genes are made of, how genes are related in the whole genome, and so on. Chapter 2 of our textbook does just that: defining genes in terms of DNA and describing the famous 'double helix' structure of DNA. Note how the textbook does this -- again, often interspersed with generally understandable non-technical language, as in the following passage, which we shall cite at length, because it is very typical:

(4) La mínima unidad viva es la célula (...) Cómo se organiza el DNA en una célula? La dotación completa de DNA de un organismo se denomina **genoma**. Está compuesto por largas moléculas de DNA que son, en definitiva, los principales componentes de los cromosomas. Cada cromosoma está compuesto por una única molécula de DNA, portadora de muchos genes. Cómo ya se trató en el Capítulo 1, los genes son regiones de DNA cromosómico que pueden ser transcritas a RNA. Los genomas de la mayoría de los organismos procarióticos consisten en un en un único cromosoma, mientras que los genomas de los eucariotas están formados por varios o muchos cromosomas. (...)

El DNA tiene tres tipos de componentes químicos: fosfato, un azúcar llamado desoxirribosa, y cuatro bases nitrogenadas: adenina, guanina, citosina, timina. Dos de las bases, adenina y guanina, están formados por un doble anillo típico de compuestos denominados *purinas*. Las otras dos bases, citosina y timina, presentan una estructura con un solo anillo del tipo denominado *pirimidina*. Los componentes químicos del DNA están organizados en grupos llamados **nucleótidos**, cada uno de los cuales está compuesto por un grupo fosfato, una molécula de desoxirribosa y alguna de las cuatro bases nitrogenadas. En adelante, nombraremos a cada nucleótido por la primera letra de su base: A, G, C y T. La figura 2-1 muestra las estructuras de los cuatro nucleótidos presentes en el DNA. (...)

El DNA está compuesto por dos cadenas (hebras) de nucleótidos entrelazadas que forman una doble hélice. Las dos hebras se mantienen juntas mediante la formación de enlaces débiles entre bases de cada una de las hebras, formando una estructura parecida a una estructura de caracol en espiral (Figura 2-2). El esqueleto de cada cadena es un polímero formado por repeticiones de azúcar desoxirribosa y fosfato. Las uniones entre el azúcar y el fosfato se denominan **enlaces fosfodéster**. (p. 24-25).

After the more informal introduction of genes, chromosomes and DNA in the first chapter, it is the aim of this chapter to introduce the students to the internal structures of cells, chromosomes and genes, and it does so in a systematic way, using the following structure-descriptors:

- (i) la mínima unidad es
- (ii) el X se organiza ...
- (iii) X está compuesta de Y
- (iv) X y Y son los (principales) componentes de Z
- (v) Y tiene ... componentes Y
- (vi) X consisten en Y
- (vii) X está formado de Y, Z...
- (viii) X presenta una estructura de Y

Structure descriptors usually are of the format ‘X consists of Y’ and hence are relational descriptions that either link two known notions, or more often, as is the case here, links a known notion with unknown, new, or as yet undiscussed concept. Thus, if X is known, not only is it said that X consists of Y, but also Y needs to be lexically defined or introduced in another way, here for instance which such metalinguistic expressions as “que se denomina...” or “llamados” (for details of the structure of reformulations in popularization discourse, see, e.g., Jacobi, 1987: 64 ff; for paraphrases, see Fuchs, 1982, 1993).

Structural descriptions however, are not limited to the rather general inclusion relations (consists of, has ... units, etc.) but may also introduce more structure, as the “double ring” used in the passage on “purinas”, and especially the use of “cadenas”, “entrelazadas” and comparisons (“estructuras parecidas a”) in the description of the spiraling ring of DNA (compared to a staircase) as a ‘double helix’. It is also at this point that a historical section (printed in a frame and with other printing type) tells the students about the discovery of the structure of DNA by Watson and Crick.

After such a sequence of structure-analytic definitions, the overall structure may be represented graphically, as is signaled at the end of this passage. Note also that structure descriptions vaguely but casually provide some more information than that X consists of Y, such as the further --sometimes rather vague-- description of these units in terms of “largas”, “principales”, “única”, “muchos”, “varios”, “double” or numbers. It is only in a next section that sizes (of genes, chromosomes, etc) will be more precisely defined, for instance in “kb” (kilobases: thousands of bases) o “mb” (megabases, millions of bases).

In sum, the kind of expert knowledge being (didactically) expressed here consists of structural information about units of various levels of generality, relations between the units, and some indication on form, quantity and importance. This knowledge can be graphically expressed, but also by various standard expressions that describe relationships (‘consist of’, ‘composed of’, etc). These expressions are part of the language and discourse of science, but also appear in non-specialized discourse, and their variation is often stylistic rather than an expression of (knowledge about) different relationships. In other words, knowledge about one relationship (part-whole) may be

variously expressed in many ways. What however generally does not vary are the technical terms chosen for the structural units themselves.

As suggested before, both in linguistics and genetics, students need to learn about complex structures and their components, as well as about functions:

(5) Qué relación existe entre la estructura del DNA y su función como molécula hereditaria?  
(p. 29).

This will be the topic of the next chapter of the textbook, but it is here suggested already what kind of structures CDA must have in order to be able to exercise its role in heredity.

## KNOWLEDGE AND IGNORANCE

Scholarly and specialized discourse not only expresses expert knowledge, but may be expected to also express lack of knowledge, that is, ignorance. Note that failing knowledge cannot be expressed as such (there are vast amounts of --future-- knowledge also experts have no clue about), but that only our *knowledge about our lack of knowledge* can be expressed, e.g., We know about the existence of X, but we do not know its structure or function, etc. In general, expressions of ignorance imply at least partial knowledge, because the very fact that we are able to mention or describe a phenomenon already suggests at least some knowledge --or at least, conjectures-- about it. The expression ‘I do not know that p’ is pragmatically inconsistent, and the expression ‘I do not know whether p’ is pragmatically correct if it presupposes that I consider ‘p’ a possible state of affairs. This means that lack of knowledge as expressed in specialized discourse in general and textbooks in particular, should be as carefully managed as knowledge. Only once in a while it makes sense to say what we do *not* know, for instance when readers at a given moment may be expected to want to know about a phenomenon. Let us examine some examples of such displays of lacking knowledge:

(6) El intento de definir con precisión qué es un gen se ve dificultado por el hecho de que (como vimos en Capítulo 1) muchos genes eucarióticos contienen misteriosos segmentos de DNA, llamadas **intrones**, que se encuentran intercalados en la región transcrita del gen (p. 29).

In this example the lack of knowledge is not complete (we do know a lot about genes), but only pertains to the difficulties of exact definition (delimitation) of genes in relation to other (possibly non-functional) parts of DNA in chromosomes. That is, we have rather precise knowledge about where our ignorance lies, and if we had the instruments we probably would be able to resolve that ignorance. The pragmatic and social functions of the expression of (partial) ignorance are obvious. They not only are correct descriptions of the current state of knowledge and ignorance of a discipline, but may imply as many suggestions for future research, that is, a program and indication of the direction of investigation in a discipline. Note though that textbooks need to keep

students interested, and such interest is probably more effectively elicited with ‘positive’ discourse about what we *do* know, than by ‘negative’ discourse about the vast amounts of knowledge we are still lacking. The same is true --for other social or professional reasons (prestige, etc.)-- in scholarly monographs and articles in journals, which seldom focus on all the things we still do not know in a discipline or about a problem.

In our textbook, there are not many lengthy passages that detail the state of our ignorance. Here are some examples:

(7) ¿Qué tipos diferentes de genes se requieren para construir un organismo? Todavía no puede darse una respuesta completa a esta pregunta, pero se tiene cierta idea de las categorías generales de genes y del tamaño relativo de estas categorías. (p. 77)

(8) A partir de tales estimaciones puede empezar a perfilarse vagamente la base genética que sustenta la vida de este planeta, incluyendo la de la especie humana. (p. 78)

(9) Desde tiempos inmemoriales, uno de los misterios de la herencia ha sido la transmisión de las características de un organismo. (p. 86).

(10) (...) aunque actualmente se sabe mucho acerca la estructura y función de genes individuales, se sabe poco acerca los *principios* por los cuales las regiones de DNA funcionales y o no funcionales se organizan en genomas completos. (...) Se sabe que hay conservación en los genes entre grupos taxonómicos relacionados, pero no se ha establecido si esta conservación es significativa. (p. 374).

(11) La existencia de DNA sin función conocida es un dilema para los genéticos. (p. 382).

The ignorance made explicit in these fragments applies to types of genes, the genetic basis of life, the mysteries of heredity and the principles of the organization of genomes, that is, to very general and fundamental issues. However, the formulation of these kinds of ignorance is different:

- todavía no se puede dar una respuesta completa
- estimaciones... vagamente
- misterios
- se sabe poco; se sabe ..pero no se ha establecido...
- es un dilema para los genéticos

These different expressions also suggest different modes of knowledge and ignorance: Giving complete answers presupposes that we already have partial answers and that there are questions; if we know something about a vague profile of the basis of genetics, we at least have vague knowledge; mysteries on the other hand imply total ignorance and hence is associated to the history of science or the experiences of people confronted with natural phenomena. Indeed, notice the difference between “it is not known” and “it has not yet been established”, whereas “dilemma” rather pertains to uncertainty of action rather than to knowledge. In other words, these variations in the description of ignorance are not merely stylistic, but different expressions of different modes of not-knowing.



Based on these few examples, we may go one step further and devise a typology of knowledge and ignorance:

- Complete vs. incomplete (partial) knowledge
- Knowledge vs. belief (conjecture, hypothesis, etc.)
- Precise vs. vague knowledge
- Scholarly vs. lay knowledge and ignorance
- Knowing/ignoring answers to questions (vs. ignoring the questions)
  
- Provisional vs. established knowledge
  - Hypotheses vs. proof
  - etc.
- Knowledge/ignorance about
  - whole vs. parts
  - structures vs. functions
  - phenomena vs. their explanation
  - etc.

This partial (sic!) typology obviously needs to be extended and systematized on the basis of both a theory of knowledge and grounded in an empirical investigation of the (expressions of) knowledge and ignorance both in everyday life, as well as in specialized contexts. Rather obvious however is the general division between the

- (a) nature or quality of the knowledge itself (complete/partial, precise/vague, secure/insecure),
- (b) those who have such knowledge (scholars, lay persons)
- (c) the variations in the referents or extensions of our knowledge, that is, what we have knowledge or ignorance *about*, such as the large number of (meta)properties of things and phenomena: structures, functions, origin, goals, explanations, etc.

## PROCESS KNOWLEDGE

Much of the knowledge and ignorance examined above applies to things (like genes), their structures, their components and functions. Large part of our knowledge, also our specialized knowledge in genetics, however, applies to the *processes* in which such things or structures are involved. Indeed, the very point of genetics, as is also obvious from our textbook, is to explain such general biological phenomena or processes as heredity, mutations or diseases. Thus, as from the second chapter, and once introduced the basic building blocks of the genome, the textbook introduces the students into the ways the genetic information of genes is being transcribed into RNA, and how the transcribed information in RNA finally leads to the construction of the basic proteins of organisms through a process of ‘translation’:

(12) Como hemos visto previamente, los productos iniciales de todos los genes son los ácidos ribonucleicos (RNA). El RNA se sintetiza mediante un proceso que copia la secuencia nucleotídica del DNA. Puesto que este proceso recuerda la transcripción o copia de palabras escritas, el proceso de síntesis de RNA se conoce como **transcripción**. (p. 52).

(13) La secuencia de nucleótidos del mRNA se convierte en la secuencia de aminoácidos de la cadena polipeptídica por un proceso denominado **traducción**. En este contexto, la palabra *traducción* se utiliza con un sentido muy similar al de la traducción de un idioma: la célula dispone de un mecanismo para traducir el lenguaje del RNA al lenguaje de los polipéptidos. Las proteínas se componen de una o más cadenas polipeptídicas. (p. 53).

Process knowledge is generally difficult to describe and explain, and cannot easily be represented in schemas or still pictures. Thus, in order to explain in (12) the synthesis of RNA the authors have recourse to the well-known metaphor of ‘making copies’ and ‘transcription’, of which the students do have at least approximate beliefs that can be inferred from general knowledge. The same is true for the linguistic metaphor of ‘translation’ in (13). Note also that whereas the concepts explained so far are related to a large variety of ‘objects’ (properties of genes and genomas) for which specialized terminology exists, knowledge about processes and their linguistic conceptualization are usually expressed in less technical vocabulary (verbs, nominalizations, and adjectives), such as ‘synthesis’, ‘changes’ etc. We have seen above that much of the didactic language of genetics appears to be based on linguistic metaphors, in which the ‘language’ or ‘code’ of genetics (DNA, RNA) is expressed in ‘words’ (or ‘sentences’?) of genes that need to be translated into the code of another language (that of proteins). It needs no further comment that these metaphors are only very approximate comparisons with what happens in real translation. At this point, we are not even sure whether we are speaking of letters, words, phrases, sentences, paragraphs or whole texts that are to be the linguistic comparisons of genetic units of analysis.

To have another impression of the kinds of discursive expression of genetic processes we find in our textbook, consider the following examples:

(14) Se produce una separación local de las dos cadenas (p. 54)

(15) A continuación, nucleótidos libres sintetizados en otra parte de la célula se emparejan con el DNA molde, “atraídos” por las bases complementarias en el mismo (p. 54)

(16) Este proceso está catalizado por la **polimerasa de RNA** (p. 54)

(17) La polimerasa de RNA se mueve a lo largo del DNA, manteniendo una “burbuja” de transcripción donde expone la cadena molde, y cataliza...

We see that the specialized knowledge about what happens in cells is formulated in rather everyday terms (“separación”, “emparejar”, “se mueve”, “exponer”, etc.) that are used to describe the manipulation of objects. Only sometimes we find more specialized terminology, such as “catalizar” in example (17), probably because there is no everyday term to describe the concept “to promote a chemical or biological process”. Hence it is not surprising that most neologisms printed in black in our text are nouns.

The interesting question is whether this linguistic property of specialized discourse (at least in genetics) also tells us something about specialized knowledge. Indeed, may we conclude that experts conceptualize genetic processes largely in actions and events that are the same or similar to those we know from general knowledge? Or do they have specific knowledge about the actions and events involved but do not invent special

verbs for them? Why indeed would the ‘separation’ of the DNA chains of genes (as in example 14) be called a separation at all, and not by a precise term that applies to the very specific way DNA chains are split up? It may be surmised that such terminological specification of acts, events and processes is seldom necessary, because the epistemic specificity of the events involved can be invested in the object/noun pairs of the specialized discourse: DNA, RNA, polymerase, etc. Only sometimes do we find special terminological terms referring to events (‘meiosis’, ‘mitosis’), but usually only in order to describe a whole complex process, as is the case for cell division. It is thus not the rather general notion of (say) ‘movement’ that needs to be specified with special terminology; it suffices that the experts know *what* moves, and how. Further analyses of specialized discourse need to show whether this is a general feature or specifically applies to the discourse of modern genetics.

### **Meta-knowledge: The divisions of the field**

We have seen earlier that geneticists, like all experts, not only have specialized knowledge about genes and genomes, but also about their own knowledge and discourse *about* the structures and functions of genes. As is the case for any specialized discipline, we could invent a special term for this kind of knowledge, and call it meta-knowledge, or organizational knowledge, because it also relates to the organization or the divisions of the field, in the same way as we know that certain problems belong rather to phonology than to syntax, or to psycholinguistics rather than to sociolinguistics.

Thus, the division of the textbook in various chapters, as we have briefly suggested before, also reflects a division or organization of the field in sub-specializations, as well as knowledge about the various main properties of genes: structures, functions, reproduction, etc. Of course, in textbooks the division may also have a purely didactic function and not reflect any disciplinary division or a conception about the underlying organization of genetic ‘reality’. In general, however, textbooks organization reflects the division of the field or conceptions about different properties of the objects being studied by a (sub)discipline. Let us therefore examine the combined macro- and superstructures of this textbook and examine what kind of knowledge is involved in such higher level discourse structures.

We hardly need to recall that such higher level structures have an important role in the organization of memory, and play a fundamental role in student’s reading and comprehension of the extremely complex meaning of this text. A Table of Contents (in fact, in this book, a short and more detailed one) is an expression of the overall topics being dealt with as well as a list of the schematic, linear order of the textbook as it is organized in chapters and sections. If students learn to memorize the text through these structures, it is plausible that they also learn to organize the fields as such, and it may be assumed that this kind of organizational knowledge will stay with them for a long time. Let us briefly examine the nature of the chapters in this book:

1. **Genetics and life.** Genetic and human affairs. Genetics and other disciplines (biology, medicine). Basic notions.

2. **Structures of genes and genomas.** Cells, chromosomes, genes, DNA, etc. and their structures. The double helix.
3. **The functions of genes.** The transcription of DNA in RNA. Properties of DNA. The translation of RNA into proteins. Structures of proteins. (Dys)functions of proteins in cells.
4. **The heredity of genes.** DNA replication and polymerase. Cell division. Patterns of heredity and crossing. Human pedigrees. Diseases.
5. **Recombining genes.** Crossing and recombining. Genetic maps.
6. **Genetic interaction.** Interactions between alleles of a gene.
7. **Mutations.**
8. **Chromosomal alterations.** Changes in the number and ordering of chromosomes. Diseases.
9. **Genetics of bacteriae and bacteriophages.**
10. **DNA Recombinant technology.** Reconstruction. Cloning. Applications.
11. **Applications of DNA Recombinant Technology.** Plants. Animals. Humans. Gene therapies.
12. **Genomics.** Structural and functional genomics. Mapping. Gene Libraries.
13. **Transposable genetic elements.** Transposons.
14. **Transcription regulation.**
15. **Regulation of cellular numbers: Normal and cancer cells.**
16. **Genetic bases of development.**
17. **Population and evolution genetics.** Variation. Selection.
18. **Quantitative genetics.**

Each discipline has its own traditional divisions, and there are of course differences of opinion among authors about what the most important fields or topics are of a discipline. On the other hand, textbook authors want their books to be widely used, and hence will usually select a division of the field that is most generally accepted. If this is also true for the present textbook, as we believe it is, then the chapter division reflects the knowledge experts have about the overall organization of the field, as well as about the importance of various topics to be assigned in a textbook. Chapter order however not only reflects importance in the discipline, but also a didactic order: Students need to know about the structures and functions of genes before they can apply such knowledge in more specific chapters.

What else do the global topics and order of the chapters of this textbook tell us about the underlying expert knowledge of its authors and about the didactic intention that *this* is what students need to know -- let us say: a moral education. Some divisions of the field are quite general in many disciplines, and therefore probably have more general organizing principles. Thus, applied and quantitative genetics are areas that might be expected also in many other disciplines, and so is developmental genetics, which also has its counterparts in other disciplines, such as development linguistics and psychology. Another area that would be widely recognizable as especially relevant is genetics as applied to various human and humanitarian domains is of course population and evolution genetics, given students and many other people's interest in Darwin, evolution and the application of genetics to population questions. The same is true for the genetic aspects of cancer, as dealt with in chapter 15. Apart from these and other 'applications' of genetics in medicine and other socially relevant areas, the core of the book and hence probably of the field forms the description of various kinds of accidental or intentional (technical) modifications of genes and chromosomes, as in mutation or DNA recombinant research, again with various applications, for instance in

the treatment of genetic diseases, and so on. In other words, the contents of the textbook and the divisions of the field seem to be controlled by expert knowledge that is itself organized by a number of very general concepts:

- (a) **structures** of genes, chromosomes, complete genomes and cells
- (b) **(dys)functions** of (these structures of) genes in cell construction, reproduction of organisms, heredity, development and diseases
- (c) **mutations and manipulations** of structures of genes and chromosomes (DNA recombinant research) and its applications.

There is (intentionally) no separate historical chapter in this book. Unlike an earlier version of this book, and other books on genetics, it does not begin with Mendel for instance (although Mendel gets a separate section in Chapter 1). Rather, each chapter has a useful historical part, in which famous scholars of the field are casually mentioned and integrated. And finally, each chapter has a large didactic section with various kinds of questions and tasks, a glossary of technical terms, etc. Such sections obviously have didactic functions, but they also exhibit expert knowledge about what students need to know (history of the field, famous geneticists, etc), and especially what they should be trained to do, e.g., solve problems of genetics.

Finally, an important part of the book are its numerous illustrations, schemas, photos, and other visuals. Especially in order to understand the complexities of genetic structures and functions such illustrations are particularly useful, and the 'visual' knowledge they presuppose are of course parts of the expert knowledge of experts. Some of this knowledge are based on or abstractions from photographs, but others are only didactically intended visual renderings of genetic structures and functions that have little to do with the 'real' appearance of such structures -- whereas functions cannot be rendered in photographs anyway, and appear with arrows or other ways to represent changes, processes, reproduction or movements of genes, cells or whole organisms (e.g. the role of genes and gene defects in diseases, heredity, etc). I do not know how much the expert knowledge of geneticists is based on or related to visual information or abstractions about genes, chromosomes and cells, but the ample use of illustrations in the textbook strongly suggests that for students the acquisition of basic knowledge about genetics is narrowly related to visual information processing. It is not likely that these illustrations are merely didactic, but (also) an important feature of the expert knowledge of the field.

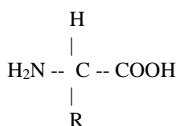
## ANALYSIS OF AN EXAMPLE

After the more general remarks on the structures of a textbook and their relations with expert knowledge in genetics, we finally shall examine in more detail a specific passage and examine how various kinds of knowledge are expressed or presupposed by that passage. The choice of such a passage is difficult and rather arbitrary, because the textbook has many passages that are comprehensible without any or much specialized knowledge, and other passages which are highly 'technical' for non-specialists. We

have selected a passage from Chapter 3 in which new important information is given, namely about the structure of proteins. This passage on the one hand expresses new information, whereas on the other hand it presupposes specialized knowledge that has been conveyed earlier in the book.

(17) ESTRUCTURA PROTEICA

Una proteína es un polímero compuesto por monómeros denominados **aminoácidos**. En otras palabras, es una cadena de aminoácidos a la que en ocasiones nos referimos con el término de polipéptido. Todo los aminoácidos se ajustan a la fórmula general



La cadena lateral, o grupo R (reactivo), puede ser cualquier grupo desde un átomo de hidrógeno (como el aminoácido glicina) hasta un anillo complejo (como en el aminoácido triptófano). Hay 20 aminoácidos que pueden ser constituyentes de las proteínas (Tabla 3-1), cada uno con un grupo R diferente que les confiere sus propiedades específicas. En las proteínas, los aminoácidos se mantienen unido mediante enlaces covalentes denominados **enlaces peptídicos**.

El enlace peptídico se produce por una reacción de condensación durante el cual se elimina una molécula de agua (Figura 3-15). Debido al mecanismo de formación del enlace peptídico, una cadena polipeptídica siempre tiene un **extremo amino** (NH<sub>2</sub>) y un **extremo carboxilo** (COOH) como se muestra en la figura.

Las proteínas presentan una estructura compleja que consta de cuatro niveles de organización (Figura 3-16). La secuencia lineal de aminoácidos de la cadena polipeptídica constituye la estructura **primaria** de la proteína (p. 61)

This passage is a typical example of a definition, given in terms of a chemical description, of a notion (protein) that is generally known, also to the students who use this textbook, but whose chemical characterization may not be known. Interesting is that at least some chemical knowledge is presupposed, because the notion of a ‘polymer’ is not defined here (or earlier in the book). It does not appear either in the glossary. The definition is given in terms of one of the relational terms we have encountered above (‘is composed of’) and another technical notion (monomer) and its (bold printed) equivalent (amino acid), introduced by the standard definitional term “called”. However, even for those who do not know what polymers and monomers are, the next sentence offers (a bit of ) relief with a paraphrase, typically initiated by “in other words”, followed by the everyday term ‘una cadena’ -- from which the reader may infer that a polymer is a chain of a specific kind of molecules (in this case of amino acids). At the same time a special word is introduced to refer to such a chain of amino acids: ‘polypeptides’. Amino acids are then defined by a chemical formula, which also presupposes at least initial chemical knowledge (e.g. about what elements the letters H, N, C and O refer to, what the sub-index 2 means after H, why these letters are sometimes written together and what the lines mean between groups of letters. On the other hand, the letter R is used here not to define a chemical element, but as an abbreviation for a (reactive) group. From the expression “todos los aminoácidos...” the

reader may infer that there are at least several of them, and that they are also different, but have the chemical group in common. The following lines confirm that inference by referring to different amino acids by name and providing a number with the existential verb “hay”. Table 3-1 (not reproduced here) gives the list of the 20 known amino-acids, with their three and one letter abbreviations. Follows the new information about what the different amino acids are different, and how they are linked up in a protein (“el enlace peptídico”).

Note that so far the introduction of new expert knowledge is structural (that is, the structure of proteins is being described) and definitional (that is, new terms are defined by a combination of known terms and new terms). These definitions are sequential, that is, if A is defined by B+C, and B is known, and C is unknown, then C will be defined in terms of D+F, etc. Moreover, structural definitions not only are given in terms of their elements or constituents, as we also know from grammar, but also in terms of their links or relationships, which is the case here for the notion of “enlace peptídico”, which is a new notion and hence also needs to be defined, as is the case here. Figure 3-15 provides a visual account of the information that by condensation a water molecule is formed (combining an H from the NH<sub>2</sub> group and a group OH from the COOH group of the amino acid), which of course presupposes the fairly generally known chemical knowledge that water is H<sub>2</sub>O. The text also calls on the figure to provide an explanation of the fact that a polipeptidic chain has an amino group (NH<sub>2</sub>) on one end and a carboxylic (COOH) group at the other. This is not further explained below the figure, but one can ‘see’ that at the end (where no links between them are formed by the elimination of water) these groups remain intact. The final part of this message returns to the standard structural account, this time of the overall organization of proteins. Again, reference is made to a figure (3-16) to explain these complex structures, whose forms would be difficult to describe without the drawings (a chain being curved into a helix, which in turn may be folded, compacted and combined with others).

We see that this passage provides a number of new notions and knowledge, partly built on more or less technical earlier knowledge (of chemistry), partly on brief, summarized information earlier in the book (on the role of genes, DNA and hence amino acids in the formation of proteins), and partly on general knowledge, such as the general structural relations ‘is composed of’, ‘chain’, ‘general formula’, and probably also the otherwise technical term “hydrogen atom”, as well as ‘condensation’, ‘water molecule’, and ‘complex ring’. Note that besides the structural description of what is composed of what, we also have a process description, namely of how polipeptidic chains are form by eliminating water molecules. Structural definitions however are not merely expressions of the structure of proteins, but also have various linguistic manifestations, of which the classical is ‘is called’ (“denominado”) and here also the expression “en ocasiones nos referimos con el término”. Finally, visual information of the figures helps the student to understand the structures, processes and other aspects involved. In other words, this passage is an expression of fairly low level, beginners’ knowledge of genetics or molecular biology, and presupposes chemical knowledge.

## CONCLUSIONS

In this paper we have examined a number of properties of the interface between specialized language and knowledge, with special application to the acquisition of basic genetic knowledge in textbooks. Although modern cognitive psychology offers many insights into the nature and the representation of knowledge, as well as into the relations between knowledge and discourse processing, there are still vast areas of lacking knowledge about knowledge. This is particularly the case in epistemology, which barely offers insight into the more empirical properties of knowledge. A plea is made for a multidisciplinary approach to knowledge which integrates insights from psychology and the social sciences. Knowledge is here defined as the “certified” shared beliefs of an epistemic community, according to the certification criteria of that community. This means that knowledge is essentially relative – although also this relativity is relative (as it should be in a consistent theory), in the sense that *within* an epistemic community knowledge is not relative but accepted as established. A typology of knowledge is proposed that makes distinctions at different levels, including the scope of the communities that share knowledge: persons, groups, nations, cultures, etc. The relevant relation between discourse and knowledge is established by the fact that a community’s knowledge is typically presupposed in the intragroup discourses of community members.

Against this broader theoretical framework, scientific knowledge is defined as group knowledge, and scientific discourse as the discourses that are produced in contexts that are characterized by (participants who have) such knowledge. In a case study of the discourse and knowledge management of a modern textbook of genetics, we finally examined in some detail the various strategies used in the explanation of new notions. The idea of such an analysis is that these discourse structures not only are functional in the didactic context of the textbooks, but reflect more general principles of knowledge organization and acquisition. Thus, we found that most (new) knowledge is introduced by a limited number of categories, such as Structure, Composition, Quantity, Form, Analogy, and Function. This is especially the case for “things” such as genes, genomes, chromosomes, rather than for “processes”, which tend to be described in everyday terms (such as “move”, “split”, etc.) or in metaphorical terms that conceptualize complex processes (“translate”, etc.). We also found evidence for the obvious assumption that technical knowledge is based on everyday, Common Ground knowledge, something which is of course crucial in initial learning situations. We finally also found that the acquisition of “specialized” knowledge is not limited to technical knowledge about objects (such as genes), but also features knowledge about the history, the organization and the applications of the discipline and its current knowledge, including information about our current ignorance in the field.

The latter is also true for our insight into the nature of the interface of knowledge and discourse. We already suggested that we know surprisingly little about the nature, representation and processes of knowledge, its acquisition, social certification, distribution and expression in text and talk. We know that language users need large amounts of knowledge in order to produce or understand knowledge, and that their



discourses are contextually controlled in such a way that they take into account what they know about the knowledge of their interlocutors. But we do not know in detail what the complex strategies of such knowledge management in discourse are. We do not know how exactly people acquire the various kinds of knowledge, how they acquire specialized knowledge that goes beyond everyday experiences, and how specialized discourse (e.g. of textbooks) is being used to teach and acquire such knowledge. In other words, much of the current research on the interface between discourse and knowledge is still on the program.

## Biographical note

Teun A. Van Dijk is visiting professor at the University Pompeu Fabra, Barcelona, after early retirement as professor of discourse studies at the University of Amsterdam. After earlier work in generative poetics, text grammar, pragmatics and the psychology of text processing, his work in the 1980s and 1990s focused on the study of racism in various types of discourse. He also did research on news, and is currently working mostly on ideology, knowledge and context. He is author and editor of many books in these areas, including the *Handbook of Discourse Analysis* (1985, 4 vols) and the introductory book *Discourse Studies* (1997, 2 vols). His latest books are *Ideology* (1998) and (with Ruth Wodak, Eds.) *Racism at the Top*, and he is now working on a book on context. He is founder of the international journals *Poetics, Text, Discourse & Society*, and *Discourse Studies*, of which he still edits the latter two. Teun A. van Dijk, whose work has been translated into many languages, has lectured in many countries, especially also in Latin America, where he also regularly visited Unicamp.

---

## REFERENCES

- ACKERMAN, D.; ANTON, T. & McCOURT, R. (Eds.). (1995). *The new science journalists*. New York: Ballantine Books.
- APPLE, M.W. & Christian-Smith, L. K. (Eds.). (1991). *The politics of the textbook*. London: Routledge.
- ARONOWITZ, S. (1988). *Science as power: discourse and ideology in modern society*. Houndmills, Basingstoke, Hampshire: Macmillan Press.
- BARNES, B. (Ed.). (1982). *Science in context: readings in the sociology of science*. Cambridge, Mass.: MIT Press.
- BECKER, S. & SELMAN, B. (1986). *An overview of knowledge: acquisition methods for expert systems*. Toronto: Computer Systems Research Institute, University of Toronto.
- BOOSE, J.H. (Ed.). (1988). *Knowledge acquisition for knowledge-based systems*. London San Diego: Academic Press.
- \_\_\_\_\_. (Ed.). (1990). *The foundations of knowledge acquisition*. London San Diego: Academic Press.
- BRITTON, B.K. & GRAESSER, A. C. (Eds.). (1996). *Models of understanding text*. Mahwah, NJ: Erlbaum.

- BRITTON, B.K.; WOODWARD, A. & BINKLEY, M.R. (Eds.). (1993). *Learning from textbooks: theory and practice*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- CALSAMIGLIA, H. & VAN DIJK, T.A. (2004). Popularization discourse and knowledge about the genome. *Discourse Studies*, in press.
- CHARNIAK, E. (1972). *Toward a model of children's story comprehension*. Ph.D. Dissertation. Massachusetts Institute of Technology.
- CHASE, A. (1975). *The legacy of Malthus: the social costs of the new scientific racism*. Urbana, IL: University of Illinois Press.
- CLARK, H.H. (1996). *Using Language*. Cambridge, England: Cambridge University Press.
- DURANTI, A. (Ed.). (1992). *Rethinking context: language as an interactive phenomenon*. Cambridge (England) New York: Cambridge University Press.
- DUSZAK, A. (Ed.). (1997). *Culture and styles of academic discourse*. Berlin, Germany New York: Mouton de Gruyter.
- FAYARD, P. (1993). *Sciences aux quotidiens: l'information scientifique et technique dans les quotidiens nationaux européens*. Nice: Z'éditions.
- FOUCAULT, M. (1980). *Power/Knowledge: selected interviews and other writings, 1972-1977*. New York: Pantheon.
- FUCHS, C. (1982). *La paraphrase*. Paris: Presses universitaires de France.
- \_\_\_\_\_. (1993). *Paraphrase et énonciation*. (Paraphrase and utterance). Paris: Ophrys.
- GRECO, J. & SOSA, E. (Eds.). (1999). *The Blackwell guide to epistemology*. Malden, Mass.: Blackwell Publishers.
- GREGORY, J. & MILLER, S. (1998). *Science in public: communication, culture and credibility*. New York: Plenum.
- GRIZE, J.B. (1981). Discours et connaissances. (Discourse and Knowledge). *Communication and Cognition*, 14(4), 343-357.
- GUNNARSSON, B.L.; LINELL, P. & NORDBERG, B. (Eds.). (1997). *The construction of professional discourse*. London New York: Longman.
- HYLAND, K. (1998). *Hedging in scientific research articles*. Amsterdam: Benjamins.
- JOHNSON-LAIRD, P.N. (1983). *Mental models: towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridgeshire New York: Cambridge University Press.
- KELLER, E.F. (1995) *Refiguring life: metaphors of twentieth century biology*. New York: Columbia University Press.
- KINTSCH, W. (1998). *Comprehension: a paradigm for cognition*. Cambridge: Cambridge University Press.
- KUIPERS, T.A.F. & MACKOR, A. R. (Eds.). (1995). *Cognitive patterns in science and common sense*. Amsterdam, The Netherlands: Rodopi.

- LATOUR, B. (1987). *Science in action: how to follow scientists and engineers through society*. Cambridge: Harvard University Press.
- LATOUR, B. & WOOLGAR, S. (1986). *Laboratory life: the construction of scientific facts*. Princeton, NJ: Princeton University Press.
- LEMKE, J.L. (1990). *Talking science: language, learning, and values*. Norwood, N.J.: Ablex Corp.
- MARTIN, J.R. & VEEL, R. (1998). *Reading science: critical and functional perspectives on discourses of science*. New York: Routledge.
- MYERS, G. (1990). *Writing biology: texts in the social construction of scientific knowledge*. Madison, WI: University of Wisconsin Press.
- POJMAN, L.P. (1999). *The theory of knowledge: classical and contemporary readings*. Belmont, CA: Wadsworth.
- PURKHARDT, S. C. (1993). *Transforming social representations: a social psychology of common sense and science*. London: Routledge.
- ROTHBART, D. (1997). *Explaining the growth of scientific knowledge: metaphors, models, and meanings*. Lewiston, N.Y.: E. Mellen Press.
- SCANLON, E. (1999). *Communicating science: contexts and channels*. London New York: Routledge.
- SCHANK, R.C. & ABELSON, R. P. (1977). *Scripts, plans, goals, and understanding: an inquiry into human knowledge structures*. Hillsdale, N.J. New York: L. Erlbaum Associates distributed by the Halsted Press Division of John Wiley and Sons.
- SHORE, B. (1996). *Culture in mind: cognition, culture, and the problem of meaning*. New York, NY: Oxford University Press.
- VAN DIJK, T.A. (1993). *Elite discourse and racism*. Newbury Park, CA, USA: Sage Publications.
- \_\_\_\_\_. (1998). *Ideology: a multidisciplinary approach*. London, England UK: Sage Publications.
- \_\_\_\_\_. (1999). Context models in discourse processing. In: van Oostendorp, Herre, & Goldman, Susan R. (Eds.). *The construction of mental representations during reading*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, pp. 123-148.
- \_\_\_\_\_. (2003). The discourse-knowledge interface. In Ruth Wodak and Gilbert Weiss (Eds.) *Multidisciplinary CDA*. London: Palgrave.
- \_\_\_\_\_. (2004). Discourse, knowledge and ideology. In: M. Pütz, J. Neff & T. A. van Dijk (Eds.), *Language and ideology*. Berlin: Mouton (in press).
- VAN DIJK, T.A. & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- VAN OOSTENDORP, H. & GOLDMAN, S. R. (Eds.). (1999). *The construction of mental representations during reading*. Mahwah, NJ: Erlbaum.
- VAN OOSTENDORP, H. & ZWAAN, R. A. (Eds.). (1994). *Naturalistic text comprehension*. Norwood, N.J.: Ablex.
- VENTOLA, E. & MAURANEN, A. (Eds.). (1996). *Academic writing: intercultural and textual issues*. Amsterdam Philadelphia: John Benjamins.
- WIELINGA, B. (Ed.). (1990). *Current trends in knowledge acquisition*. Amsterdam: IOS Press.
- WILKES, A.L. (1997). *Knowledge in minds: individual and collective processes in cognition*. Hove: Psychology Press.