PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF ORAN
FACULTY OF LETTERS, LANGUAGES AND ARTS
DEPARTMENT OF ANGLO-SAXON LANGUAGES
SECTION OF ENGLISH

ACCESS TO THE DISCOURSE COMMUNITY OF BIOLOGISTS IN ALGERIA: IMPROVEMENT IN ENGLISH LANGUAGE PEDAGOGY USING A CORPUS-BASED APPROACH

This thesis is submitted to the University of Oran
in fulfilment of the requirements of the Degree of 'DOCTORAT D'ETAT'
in Applied Linguistics and TEFL

Presented by
BELKHENCHIR Khadoudja

Supervised by
Professor MILIANI Mohamed

Members of the Jury:

President :	Dr. Benmoussat Smain	Université de Tlemcen
Supervisor :	Pr. Miliani Mohamed	Université d'Oran
Examiner:	Dr. Simpson James	Université de Leeds
Examiner:	Pr. Kaid-Harche Meriem	U.S.T.Oran
Examiner:	Dr. Lakhdar Barka Sidi Mohamed	Université d'Oran

December 2006

DECLARATION

I, hereby, declare that this work has not already been accepted in substance for any degree, and is not concurrently being submitted in candidature for any other degree.

Khadoudja BELKHENCHIR

The researching, preparation and presentation of the thesis have been undertaken entirely by the author.

Khadoudja BELKHENCHIR

Dedicated to the memory of my father and my uncle Professor Belkhenchir Djilali.

May peace be upon them !

ABSTRACT

This thesis is a pedagogically-motivated study. It deals with the teaching of English (E.S.P) at university level, and more particularly English for Science and Technology (E.S.T).
The present study was conducted at the Department of Biology of the University of Oran. It aims to investigate the context in which the language course takes place, and to explore the scientific discourse used by this community using a corpus-based approach.

An investigation of the discourse community under study is explored through the use of questionnaires. The data were analysed both for explicit information and implicit perspectives. Based on the evidence gathered, this research offers useful implications to English language teaching, focusing mainly on the reading skill, and proposes a task-based teaching to training programmes for effective participation of new members in the discourse community.

A detailed analysis (based mainly on 'collocation' and 'colligation') of a corpus of biology texts which the learners were reading during their academic courses, is done using two methods: One manual and the other computational. The findings show occurrences of language patterns through concordancing .

Language teachers are provided with useful guidelines as to how they can provide prospective students of a particular academic community with an experience in acquiring fundamental concepts of their discipline before they become members of their target community.

The present research makes a potential contribution to the study of the English language (E.S.P.) by examining closely both the lexis and the grammar of one specific genre i.e. scientific journal articles utilising computerised technology.

The thesis consists of six chapters. 'Chapter One' establishes the research area by introducing the research questions with reference to a specific scientific community. 'Chapter Two' deals with the review of the literature of the two concepts, namely 'discourse community' and 'discourse analysis', related to this research. 'Chapter Three' describes the English teaching/learning context through questionnaires study. 'Chapter Four' specifies the corpus for a manual analysis. In 'Chapter Five' a computational analysis of the same corpus is done using concordancing. 'Chapter Six' covers the pedagogical implications.

Keywords: Discourse community; scientific discourse; genre; concordancing, collocation, colligation; computational analysis; corpus-based approach; task-based teaching/learning.

ACKNOWLEDGEMENTS

My first thanks go to my supervisor Professor Miliani Mohamed for all his guidance, learned comments, support, and positive encouragement throughout the whole of my research work. Thanks to him I have been able to develop both professionally and intellectually.

I am very honoured and very much indebted to the members of the jury for having accepted so kindly to provide me with a critical reading of this work: Dr Smain Benmoussat, Dr James Simpson, Professor Merièm Kaid-Harche and Dr Sidi Mohamed Lakhdar-Barka .

Many thanks also to Dr Peter Roe and Professor Denis Ager (at the Language Studies Unit (L.S.U) / Department of Languages and European Studies - University of Aston in Birmingham, U.K) for their advice and guidance at the preliminary stages of this research, and to all the L.S.U. staff who provided us with seeds for thought during the PhD summer courses.

Thanks also to Dr Peter Roe and Professor Frank Knowles for granting me permission to use A.T.A (Aston Text Analyser) programme.

Special thanks to all the authors and editors of the scientific journals (in Appendix C).

My thanks also go to Dr Lakhdar Barka Sidi Mohamed for his advice and assistance, and all my gratitude to Mr Nait-Brahim Abdelghani for having sorted out some computer problems.

I am grateful to the teachers and students at the Biology Department of the University of Oran for their cooperation, as well as to the Chief-Librarian Mrs Touhami.

Many thanks and gratitude to my family: my mother, my brother Abdelhakim and his wife, my sisters and brothers in law, for their love, support and encouragement, as well as to all my friends and colleagues who have supported me during all these years !

TABLE OF CONTENTS

Chapter Two: Review of the Literature and Research Methods

Chapter Three: The teaching and learning context

Chapter Four: Framing the biology corpus for manual analysis

Chapter Five: Computational analysis of the biology corpus

Chapter Six: Pedagogical Implications

List of Tables

List of Figures

Glossary of terms used in concordancing literature

ASCII (American Standard Code for Information Interchange): A common character set supported by all computers, representing 128 English letters, numbers, and symbols.

Collocate: place side by side; a collocation of a word shows words appearing frequently near it.

Collocation: The co-occurrence of lexical items.

Colligation: The co-occurrence of grammatical classes.

Concordance: It gives access to many important language patterns in text.

Concordancer: A computer programme to do the task of concordances.

Corpus: a collection of texts, from a specific source or field (e.g., medical, literary, etc.) in the form of a single computer-readable file.

Corpora: plural of corpus.

Corpus-based Approach: Approach using a corpus in a single computer-readable file.

KWIC (Key Word In Context): a common option for displaying search results for text databases. For a concordance, the search term appears in the centre of each line, in coloured or bold text, with the surrounding sentence context to the right and left.

Node: The 'node' or the 'node word' is the word that appears in the middle of the screen in a list of concordances. It is the keyword.

Span: The measurement, in words, of the co-text of the node (- 4 to the left / + 4 to the right)

Synoptic Profile: It calculates collocations for all items to the left and the right of the node for a span of 4 x 4.

Tokens: The running words in the corpus.

GENERAL INTRODUCTION

> *"Learning is not finding out what other people already know, but is solving our own problems for our own purposes, by questioning, thinking and testing until the solution is a new part of our life."*
>
> (Handy, 1990: 50)

The primary concern of this research is the present and future development of the Algerian society. The particular focus it takes is the role of language in that development, more particularly that of English as the language for international communication.

In Algeria, considered as an "emerging" nation, there is much hope invested in science and technology to answer issues facing this country to move into "modernity." Therefore, scientific communities feel more concerned by the acquisition of a language that will enhance this mutation, and more and more Algerian researchers believe that the "normalisation" of the teaching of English would be a fundamental contribution to these objectives.

Since its independence in 1962, a number of decisions related to language teaching policy have been taken in Algeria. Therefore, Arabic is used as the language for national identity and integrity. But, like many other newly independent and developing countries in Africa, Latin America and Asia, Algeria needs a world language to be in contact with the world's latest technological and scientific information which is essential for economic development.

Algeria has reckoned that the goals of economic and commercial development cannot be achieved through the national language alone. Therefore, for Algeria to gain membership in the world community and compete effectively on the international scene because of globalisation (as for example joining the World Trade Organisation in 2007, and other organisations), students should be well-initiated in English to meet their own needs and the needs of the national education which aims at economic and commercial development.

In Algeria, the teaching of the different components is done through English in few institutions only e.g. I.A.P.(Institut Algérien du Pétrole in Arzew and Boumerdès) and I.N.E.L.E.C. (Institut National d'Electronique in Boumerdès). In other institutions, English is taught as a foreign language from primary to university level. At university level, it is taught as a main subject to students majoring in the 'English language and literature' at the Department of English. It is also taught as a compulsory subject in the form of E.S.P (English for Specific Purposes) in all science and social science institutes at undergraduate and postgraduate levels.

Throughout the world, on a macro-scale, there has been a growing demand for English courses to provide both undergraduate and postgraduate students with the kind of skills they require in order to pursue their studies (English for Academic Purposes) to reach a more advanced level of performance.

On a micro-scale, scientists (professionals) and science students (students still getting initiated in their field of study) need to achieve a reasonable level of reading competence in English as much of the material they have to read is published only in this language.

Therefore, it is important for Algeria's future development (both socially and economically) that scientists (who already master Arabic and/or Berber, and French) be properly educated in English (E.S.P) to become accepted members of the international scientific community. To meet these needs, appropriate courses should be designed for students at university level.

There is, of course, an increased number of courses for teaching E.S.P. commercially designed for different groups of learners. These courses might be suitable to some groups of learners and unsuitable to others, because either the audience aimed at is not the same or the purposes to reach are different. This means that an E.S.P. course designed for French, Spanish or even for other Arab students could not be suitable for Algerian students for different reasons: language and cultural backgrounds, schemata, and so on.

In Algeria, the teaching of English (E.S.P) in science and engineering faculties is often done by either language teachers or sometimes by subject-specialists. The teaching of English to science students by either a language teacher only or a subject-specialist only (instead of the collaboration of the two) may cause learning problems.

A long practice in the field i.e. teaching engineers and technicians in the oil industry, doctors, etc., enriched me professionally (through contact with different colleagues, finding out about teaching methods and techniques and applying them), and prompted me to undertake this research as it urged all of us (as teachers) to be up to date with what is happening in the world as far as teaching E.S.P. is concerned.

However, this same experience made us notice some pedagogical mismatch in terms of teaching / learning practices of the English component: the teaching syllabus, or the teaching methodology, and so forth, leading to an inadequacy between the stated objectives and the actual practice.

In Algeria, designing a syllabus for science students has long been based on teachers' intuition and common sense, sometimes even without assessing the real needs of the students. While designing E.S.P syllabuses, teachers used mainly 'E.S.P concocted texts' (i.e. texts generated for English teaching purposes, instead of 'authentic texts'). Also, some syllabuses have focused on grammar only and others on lexis only. Therefore, focusing on one area of language teaching at the expense of the other may lead to an inefficient English language training and a growing sense of dissatisfaction. In this research, we aim to avoid taking such a risk.

Theme of the research

The principal objective of this research is to dig into the nature of the problem of inefficiency of language training programmes in a particular Algerian community in tertiary education, namely that of biologists, with the purpose of undertaking a remedial action.

Our aim is to gain insights that may help in the revision of the pedagogical methods and curricula so that English should not remain for long as just a subject on the timetable, but become a resource which the student can use.

Therefore, the research framework will deal with two concepts :

a) the notion of 'discourse community' in reference to the teaching/learning practices of the English language, and how this latter affects the development of the discourse community under study, and the Algerian society as a whole; and

b) 'Discourse analysis', the description of how the English language functions in extremely specialised scientific texts, in order to consequently develop an appropriate teaching methodology.

In order to achieve this purpose, we have selected one specific scientific community (undergraduate biology students) studying at both Arabophone and Francophone Departments at the University of Oran, in their third and fourth years of study (reasons for such a choice will be given in 'Chapter One'). These students studied all their subjects in the primary and secondary schools in Arabic. They studied French as a first foreign language and English as a second foreign language, with a few exceptions, those who studied German or Spanish. They then had three semesters (two semesters in the third year and one semester in the fourth year) of English study at the department of biology.

The aim and objectives of the English course at the biology department (as stated by the institution 'the university and departments', and the Ministry of Higher Education) is to train students to become able to read scientific texts written in English in their own discipline. This specialised reading objective is necessary at the writing-up stage of their

projects, i.e. mémoire. The end-of- course project in biology (i.e. the mémoire) is seen to be performing a social function because it "*provides evidence to examiners of a candidate's suitability for the award of a degree"* as stated by Thompson about the doctoral theses (2001:4), and to climb up the social ladder afterwards.

This research aims to develop a methodology to help the students read quickly and efficiently. To reach such an objective, the research follows a rational progression.

We first started by the description of the community under study (the biologists) and showed how this latter could form a discourse community by applying the twelve criteria (Swales' six criteria, and six management science elements summarised in the acronym 'CATWOE'), as we are primarily interested in managing English language teaching / learning in the classroom. For that, two questionnaires were designed.

One questionnaire addressed the teachers of English at the biology department to find out what teaching methodology is used by both kinds of teachers i.e. language teachers and subject-specialists teaching English.

As one of the characteristics of E.S.P. courses is the importance placed upon meeting the specific needs of the learners to establish the target tasks they need to be able to perform, the other questionnaire addressed students to find out what their real needs (not those assumed by their teachers and/or institution) for learning English are, what difficulties they have met and what suggestions they could make to improve the English course, to be in a better position for the preparation of appropriate teaching materials.

The results of the preliminary needs analysis showed the students' real needs and unveiled the problems met by students while learning English. They also provided the motivation for formulating the research questions which will be introduced in 'Chapter One' and investigated throughout this study. It seems from this needs assessment that a revision of the English course at the Department of biology is necessary in order to locate where the attendant inefficiencies lie in order to suggest improvements.

Motivation for the study

The primary motivation for the study is pedagogic. The aim is to reach a richer understanding of the nature of the scientific texts read by Algerian biology students, in order to improve the quality of the English course.

As the English course at the biology Department consists mainly of developing the reading skill of the students, the results of the needs assessment specified the type of scientific literature the students need to read to get the necessary information in their subject-discipline. According to the students, they mainly need to read scientific journal articles. These were selected to form a corpus which will be analysed to find out how the English language functions in these extremely specialised texts. Therefore, two kinds of language analyses of this corpus of texts are done in 'Chapters Four and Five'. The first analysis is quantitative and it is done manually to show what lexical and grammatical structures are used in the whole corpus.

The second approach undertaken in this research is the use of computer technology in order to make the texts tractable for analysis. The texts are indexed on a computer to identify particular features, and then to quantify and analyse the different instances with the use of concordancing software. This computational analysis, using a specific computer programme called ATA (Aston Text Analyser), is both qualitative and quantitative. It is complementary to the first analysis because as Biber, Conrad and Reppen state in this direction *"corpus-based analysis should be seen as a complementary approach to more traditional approaches, rather than as the single correct approach."* (2000: 9)

In this research, the analysis consists of selecting representative content-words (as nodes) in the form of nominal compounds because these were perceived by students as impediments for their text comprehension. The computational analysis is based on a concordance-based approach which shows the content-words in their near environment (collocation) i.e. what items (lexical or grammatical) precede or follow the node, because *"the goal of corpus-based approach investigations is not simply to report quantitative findings, but to explore the importance of these findings for learning about the patterns of language use."* (Biber, Conrad and Reppen, 2000: 5)

The aim from using a computational analysis is to discover those elements that would remain hidden with other methods, and to complete the manual analysis. The findings of both analyses will be dealt with in 'Chapter Six.'

As a language teacher, I do not aim to teach biology content-discipline. The selection of a biology corpus is used as a tool to reach the objectives of this research related to the need of students having to read authentic texts, and to show how language is used in these extremely specialised texts.

Therefore, identifying and analysing the third and fourth year biology students' needs forms the basis for the teaching of the reading skill: for developing a relevant methodology, selecting appropriate classroom tasks and developing suitable evaluation procedures.

Research Questions

The problem stated in the 'General Introduction / page 3' (the teaching of English based on teachers' intuition, use of 'E.S.P concocted texts', focus of the syllabi on one area of language 'grammar' or 'lexis' at the expense of the other), and the relatively little genre-analytic research into scientific discourse dealing with both grammar and lexis, prompted me to undertake this research.

There is quite a good deal of research into scientific texts dealing with either the grammar or the lexis (but not with both), e.g. Tarone et al. (1998), Gunawardena (1989), Hanania and Akhtar (1985), Salager (1984), Wingard (1981), Bartolic (1978), Herbert (1965) and Barber (1962). As yet, there is very little study of complete texts which shows both the lexis and the grammar of scientific discourse.

The present study aims to fill this gap. It attempts to deal with the study of both the grammar and the lexis of scientific discourse, i.e. the discourse of biology, based on evidence (data collected from the manual and computational analyses).

In this research, the first question involves finding out about how the teaching / learning of the English course is done in this particular community, in order to locate the attendant inefficiencies. It is essential to analyse the teaching / learning situation to be able to suggest solutions to existing inefficiencies.

The second research question is related to the English language itself, not only as being too sophisticated (language of science), but also because of its position in the teaching/ learning Algerian programmes (as third or fourth language, after Arabic/Berber + French).

The third research question will focus on the suggested teaching methodology.

Hypotheses or 'Anecdotal Evidence'

The methodology most adapted to probe this problem has to answer the specific aspect of this investigation. Indeed, the present study has to be exploratory in orientation, and as such hypotheses are not plainly stated, and subsequently tested.

The role of quantitative analysis using corpus-based approaches goes beyond the quantitative patterns observed empirically and makes suggestions as for the functional interpretations explaining why the patterns exist. Therefore, it is generally admitted that corpus-based approaches work as quantitative measures that may confirm the truth of intuitive extrapolation or 'anecdotal evidence' (Biber, Conrad and Reppen, 2000:3). It is a fact that most of the process implemented is an exercise of verification and confirmation or invalidation growing out of other kinds of investigations. One may say that this type of processing is nothing but an attempt to transform an 'anecdotal evidence' into a hypothesis to be statistically illustrated, not demonstrated.

Thesis Outline

The present thesis consists of six chapters:

Chapter One establishes the research area. It gives an overview of the foreign language teaching policy in Algeria. It then focuses mainly on the teaching of English, and the underlying reasons for introducing it in Higher Education together with the problems related to it.

The chapter states the research problem and introduces the research questions with reference to a specific scientific community used for a sample study.

Chapter Two deals with the review of the literature of the two concepts related to this research namely 'discourse community' and 'discourse analysis.'
The first part gives the definition of the notion of 'Discourse Community', and it then specifies the discourse community characteristics with reference to the community of biologists.
The second part is devoted to discourse and genre analyses as our students (as learners) are concerned with reading scientific texts in their own discipline.
'Discourse' is the element through which these learners interact. Therefore, definitions and examples of discourse and genre analyses in E.S.P are given, and the description of one specific genre of concern to our students, i.e. biology journal articles is done.
Finally, the chapter ends by specifying the research methods that will be used in 'Chapters Three, Four and Five.'

Chapter Three is concerned with the first concept of our research i.e. discourse community, and it deals with the assessment of the English course through the analysis of two questionnaires (one for the E.S.P. teachers and the other for the students).
The aim is to find out how the discourse community under study is organised in terms of English language teaching / learning practices.
It also highlights the students' reading problems and strategies, and specifies the type of discourse and the genre used by Algerian biology students.

The second concept underlying our study i.e. discourse analysis '*How does the English language function in extremely specialised texts'* is investigated in two separate but related chapters (Four and Five).

Chapter Four gives a detailed manual analysis of the biology corpus, thus showing the different lexical and grammatical structures that constitute the biology discourse, and unveils aspects of complexity of this discourse represented in 'lexical density' (in the form of nominal compounds) and 'grammatical metaphor' (use of nominalisation).

Chapter Five deals with the computational analysis of the same corpus used in 'Chapter Four'. It is complementary to the analysis done in 'Chapter Four'. A concordance-based study of 100 most frequent lexical items i.e. content-words (nodes) in the form of nominal compounds (NCs) is done, thus showing collocation of recurrent and important items (what language items follow or precede the nodes). The chapter ends with the presentation and analysis of the data obtained from the concordance-based study.

Chapter Six is devoted to the pedagogical implications**.** The data obtained from both analyses in 'Chapters Four and Five' will be used as an input for the suggested task-based methodology which also uses concordancing.

The 'General Introduction' has brought to the fore the suggested points to be discussed, in this research, one after the other in the successive chapters.

CHAPTER ONE

STATEMENT OF THE PROBLEM AND CHOICE OF COMMUNITY

CHAPTER ONE

STATEMENT OF THE PROBLEM AND CHOICE OF COMMUNITY

1.1 Introduction

Scientists, in Algeria or elsewhere, form a discourse community that needs to update its knowledge by reading the most recently published information in the scientific field. The latest information available is almost all written in English. It is claimed that 80% of the world's scientific production is written in English.

The present research focuses on 'discourse community', as access to societal power for the discourse community under study is through education and technology, based mainly on familiarity with the norms of the written language of scientific discourse.

Therefore, it is important for Algeria's future development that scientists be 'properly' educated in English to become accepted members of the international scientific community.

Scientists need to be 'properly' trained in English because English occupies the third position (after Arabic and French) or fourth position (after Arabic, Berber and French) in terms of language instruction in Algeria, and inappropriate or inadequate language training programmes could impair the efficient and effective development of the country.

An overview of the foreign language teaching policy seems to be necessary in order to state the research objectives. I do not aim to give a detailed description of the national language policy, but to deal with those points that would be of relevance to this research.
In Algeria, at independence in 1962, the foreign language issue was dealt with in the following way.

1.2 Overview of the foreign language teaching policy

Like many former European colonies, independent Algeria has had to deal with problems of choosing which language(s) it had to encourage its citizens to learn. Besides Classical Arabic (Al-fusha) taught as the first language and as the language for national identity and integrity, other languages (second and foreign) were to be taught as part of the new language planning policies and practices of the post-colonial era. The foreign language teaching policy seems to have been influenced by two main factors: historical and economic.

1.2.1 Historical factors

In 1962 When Arabic was declared the national and official language, French remained the provisional official language before the completion of the arabisation process.

1962-71 period witnessed the bilingual instruction (Arabic and French) as a transitory phase before the achievement of total arabisation. At this period, French was considered as a second language. After this period, and as Djité clearly describes it:

> *By 1971, a 'système-ponctuel' (a step-by-step system) was adopted. According to this system, all subjects were arabized at a given level one year at a time. By 1974, the entire public school system was arabized and French was taught only as a foreign language in the third year* [of the primary school.] (Djité, 1992:21)

At this period, French became the first foreign language.

French played and is still playing its role as it is still taught from primary to university level. From 1993 to 2004, French was introduced in the fourth year (instead of the third year) of primary school (age 10). In 2004-2005, French was introduced in the third year of primary school. Then, in 2005-2006, it was introduced in the second year of primary school. In May 2006, the Minister of Education, Mr Benbouzid, declared that French will be introduced again in the third year of primary school in the academic year (2006-2007).

This ‘swing of the pendulum’ i.e. the teaching of French at primary school (once in the third year , then in the second year, then in the third year again), sometimes without even justifying such a policy, seems to be decided without any previous planning and shows a real improvisation.

The Ministry of Education seems to take decisions without any analytical work as regards the linguistic, social and economic factors, and that long-term planning (or even short-term planning) is not part of their plans.
This kind of decision-taking and lack of long-term planning would obviously affect everyone in the society: the teachers, the pupils (with all the increasing number of drop-outs), the textbook designers, the pupils’ parents, etc.

Miliani clearly summarises such a policy when he states that:

> *changes are always conceived in a limited and localised stage of the system, never as part of a whole where transformation at a level may have impacts on others. Besides, decisions for change seem to come out of the blue. The idea of long-term planning, of thinking in terms of decades, is something that decision-makers do not seem to be capable of.(Miliani, 1996: 8)*

At university level, French is taught to students majoring in the ‘French language and literature’ at the Department of French. French is still the medium of instruction in the Medical Science Faculty (Medicine, Pharmacy and Dentistry) as well as in the different engineering faculties (computer science, civil and electrical engineering).

The teaching of Berber ‘Tamazight’ was introduced in ‘Primary School’ in 1993-94, but was discontinued. It will be reintroduced in the fifth year of ‘primary school’ and in the second year of ‘secondary school’, in 2006-2007, as part of the new education reform.

The teaching of other foreign languages (English, German, Italian, etc.) was introduced in the second year of the Middle school. Students could choose to study the foreign language

they wished if different foreign languages were taught in the school where they studied. The teaching of English (as a foreign language) is provided in all schools in Algeria (in cities and villages), whereas the teaching of other foreign languages was provided only in the main cities like Algiers, Oran, Constantine, Mostaganem, etc.

1.2.2 Economic factors

As Algeria's wealth of oil and gas was growing, Algeria needed to develop the oil and gas industry as well as other industries. To develop the industry, the Algerian people needed to be trained to cope with the new techniques of the new technologies. Therefore the Algerian authorities had to take further decisions as to which foreign languages they had to introduce in the educational system.

Classical Arabic is used for national unity as a prerequisite for political stability and this in turn can enhance economic prosperity. To reach economic prosperity, Algeria needed a language to play that role because Arabic (the language for identity and integrity) could not do that. Ager observes:

> *The Arabic language itself needed change and development, it had not been used in advanced scientific and technological spheres for it lacked many of the needed terms and expressions in which to describe processes, inventions and procedures applicable to a modern society.* (Ager,1984: 6)

Therefore, French was used instead, in many different fields. As Djité (1992) put it "*the knowledge of French was seen as a guarantee to socioeconomic well-being.*"

However, nowadays, even French cannot play fully that role, as almost all scientific information is published in English (as stated previously). The French scientists themselves have to write their work, or part of it, in English if they want it to be worldwide used. So, in Algeria, students are increasingly encouraged to take the learning of English more seriously than in the past because of its even greater importance vis-à-vis French, as a language for wider communication in the scientific and technological spheres.

1.3 Role and Status of the English language in Algeria

At independence, French was given importance (for reasons explained in 'Section 1.2.1'), and English was only the second foreign language after French. Changing attitudes began to be noticed in the 1980s (with the advent of the 'Ecole Fondamentale' as will be explained later). From this period, English began to be given greater importance than French as a medium of learning, especially at university level. The role of English in Algeria has been that of a second foreign language and has been taught from grade 8 (age 14) to both levels of university education (undergraduate and post-graduate), with the aim of enabling learners to benefit from 'external experiences and contemporary cultures.' (Fetni, 1987: 10).

In the 1990s, English gained more ground. In 1993-94 English was introduced for the first time in Algeria in some primary schools. Pupils (age 10, at the fourth year of primary school), or rather the pupils' parents, had to choose for their children between English or French as a first foreign language. But this 'dual' foreign language teaching at the primary school was discontinued as it did not have much success, and only French remained on the timetable.

In attempting to define the role that English plays in the educational system, I shall have to briefly present the school levels at which English is taught.

1.3.1 English at school-level

In Algeria, English is taught in the Middle and Secondary Schools at different levels, and in all streams (literary, scientific and technical teaching).

1.3.1.1 English in the Middle School

During the period of 'Enseignement Général', pupils in the Middle Schools (Collège d'Enseignement Moyen 'C.E.M') had the possibility to choose English as a foreign language or any other foreign language such as Spanish, German, Italian or Russian, in addition to Arabic as first language and French as a second language or first foreign language. The teaching of foreign languages was introduced in the ninth year of education (age 15).

The status of the English language has changed since the 1980s. The foreign languages mentioned above have been gradually replaced by English. Since the creation of the 'Ecole Fondamentale', English has been the only foreign language taught and has been moved to second rank next to French. It is introduced in the eighth year of education.

For the other foreign languages (Spanish, German, Italian, etc.) they were not taught anymore in the middle and secondary schools. As the Minister of Education[1] stated in an interview in 1992,

> *As regards the other foreign languages such as Spanish, German or Italian, a new policy would be put in use for their re-implementation i.e. 'redynamisation[2]'. For this purpose, contacts are being taken with the foreign embassies*. (El Watan, 1992)

1.3.1.2 English in the Secondary School

English is taught in secondary schools during the three years of education in all streams.
In the literary stream, English is compulsory and is taught for 4 hours a week (2 sessions), together with another foreign language (Spanish or German) that the students choose. The 'coefficient' of English in this stream is three (3). In the scientific stream, English is compulsory and it is the only foreign language taught for three hours a week (two sessions). The 'coefficient' in this stream is two (2).
As part of the new education reform (March 2006), a new schedule for teaching English in the secondary schools has been developed where the numbers of teaching hours as well as the coefficients have been changed because of the new 'Architecture' of the Baccalauréat. There will be six main branches for the 'Baccalauréat 2008': Foreign Languages, Letters and Philosophy, Experimental Sciences, Mathematics, Technical Mathematics, and Management and Economics (Le Quotidien d'Oran, 2006:09). The number of teaching hours and the coefficient of English have increased in some branches thus showing the

[1] "*S'agissant des autres langues étrangères telles l'espagnol l'allemand ou l'italien, une politique nouvelle serait mise en application pour leur redynamisation. A cet effet, des contacts sont déjà en cours avec les ambassades étrangères.*" (in El-Watan, 1992)

[2] 'Redynamisation' : This is a diplomatic formulation to say that the teaching of these languages will be implemented again.

importance given to English by the Algerian authorities. The number of teaching hours and coefficients are as follows:

- Foreign languages: 5 hours a week / Coefficient 4;
- Letters and Philosophy : 4 hours a week / coefficient 3
- Experimental sciences and Technical Mathematics: 3 hours a week / coefficient 2, but in the third year of secondary education the coefficient is 3 for both branches;
- Mathematics and Management / Economics: 3 hours a week / coefficient 2.

(Teaching Department of National Education, March 2006)

1.3.1.3 English at university level

English is taught as a main subject to students majoring in English at the Section of English of the Department of Anglo-Saxon Languages (Faculty of Letters, Languages and Arts) which aim is to train students mainly for a teaching career (in Middle and Secondary Schools), or for post-graduate studies for a university teaching career or research, as well as for working in other economic sectors either state-owned (SONATRACH, Mobilis) or private (Djezzy, etc.). The Section of English has always counted the biggest number of students in the four years of study, as compared to the other departments of foreign languages.

The table below gives the total number of students in the different foreign language departments of the Faculty of Letters, Languages and Arts of the University of Oran - Essénia. The figures given in Table 1.1 have been almost the same for the last five years.

Department	Number of Students
English	1380
French	1280
German	352
Spanish	323
Russian	66

Table 1.1: Number of students in language departments (2005-2006)

The curriculum of the English degree includes language modules (Grammar, Phonetics, Conversation, Written Expression, Reading Comprehension and Listening Comprehension) and content modules such as Linguistics, African Literature and Civilisation, American Literature and Civilisation, and British Literature and Civilisation, together with T.E.F.L and Educational Psychology in the fourth year to prepare students for their teaching career.

English is also a compulsory subject (not a medium of instruction) in all science and social science institutes of Higher Education at both undergraduate and post-graduate levels. English is taught in these institutes in the form of E.S.P. (English for Specific Purposes). By E.S.P., we mean English taught to a group of learners for academic or professional purposes. In Algeria, the purpose is to help students have access to academic scientific and technological information in English publications and contribute to international meetings and conferences in order to avoid isolation from new discoveries and university standards. E.S.P. is also incorporated as an option into the 'Linguistique Appliquée' Magister programmes (post-graduate level) at the three main universities of Algiers, Oran and Constantine.

Three E.S.P. Centres (run by the Algerian Ministry of Higher Education and the Overseas Development Administration 'O.D.A' in London) were set up between 1988 and 1994 at the universities of Algiers, Oran and Constantine, and the objectives of which were agreed upon by the Algerian authorities, and summarised by the former British Cultural Attaché in Algiers, as follows:

1) To promote the use of English in tertiary level establishments so that it can become a tool, for further academic purposes and an effective means of assisting Algeria's economic and commercial development;

2) To help raise the status, orientation and objectives of E.S.P. in Algerian tertiary level establishments so that it may become a viable part of the higher education curriculum in key priority areas;

3) To help Algerian cadres in key priority areas to benefit from postgraduate training and research in Britain.

Therefore, between 1988 and 1994, English language teaching programmes to prepare students for the IELTS (International English Language Testing Services) exam were delivered every year, for one year, to students from science institutes who got scholarships from the Algerian government to go abroad (United States of America, Great Britain and Canada) to further their studies. Unfortunately, this training course was discontinued because of the events that took place in Algeria during the period called 'la décennie noire', literally meaning 'the black decade' (the 1990s).

1.3.2 The teaching of E.S.P. in Algeria

In Algeria, English is considered as a means for achieving scientific and technological development and its present introduction in different sectors attests to this.

E.S.P. courses are provided in different institutions such as I.A.P. (Institut Algérien de Pétrole), I.N.E.L.E.C. (Institut National d'Electronique)., and the C.P.E (Centre de Perfectionnement de l'Entreprise) funded by SONATRACH (a state-owned oil company) whose aim is to train its staff to become able to use English in different contexts.

1.4 Reasons for teaching English in Algeria

As English has never been the medium of instruction in Algeria (except in few institutes as the I.A.P and I.N.E.L.E.C), we can group the reasons for teaching English in Algeria into two:

1.4.1 English as the language of science and technology

English is considered as the language of science and technology because most of the scientific literature is either originally written in English, or is translated into English.

Hayane[3] (1989) stressed this reason when he said,

[3] « *Si à partir de 1923 une partie des Algériens s'était mise à voir dans les langues étrangères un moyen de progrès et de communication avec le reste du monde, après 1962 les autorités ne leur accordèrent qu'un crédit d'utilité limitée à la science et à la technique moderne.* » *(Hayane, 1989:45)*

> *From 1923 part of the Algerians had started seeing in the foreign languages a means to progress and communication with the rest of the world; after 1962 the authorities only assigned them a functional status limited to science and modern technology.*
> *(Hayane, 1989: 45)*

According to Taleb Ibrahimi (a former Algerian Minister of Foreign Affairs), to develop science and technology in Algeria, we need to master English. He stated that "science and technology need English."[4] (1976:32)

However, English is a code which is mastered only by a small fraction of the society, and the achievement of its generalisation is conceivable only if the necessary conditions are satisfied. The provision of these conditions is pertinent to avoid falling in the same 'vicious circle' of students' failure followed by drop-outs.

Therefore, the development of the Algerian society is crucially dependent on an efficient and effective command of the English language.

1.4.2 English as a substitute for French

This is clearly stated in an official document[5] which reads that:

> *The generalisation of the teaching of English should be considered as a substitute for French.* (Teaching Services, Ministry of Higher Education, 1986)

This is so mainly because of its freedom from 'colonial connotations' in Algeria. As Treffgarne (1994) put it,

> *.....renewed emphasis on arabisation in Algeria since 1990 has affected the role of French as the medium of higher education,*

[4] « La science et la technologie ont besoin d'anglais .» (Taleb Ibrahimi, 1976 : 32)
[5] « *La généralisation de l'enseignement de l'anglais doit être examinée en termes de substitution à la langue française. » (Direction des Enseignements, Ministère de l'Enseignement Supérieur, 1986)*

> *rather than English. In addition to its freedom from colonial connotations, the dominance of English as the language of science and technology has accentuated its less emotionally charged function as a language of wider communication.* (Treffgarne, 1994: 7)

The Algerian Minister of culture, as quoted by Taleb Ibrahimi (1976:164) also emphasised the importance of using English because it has no colonial connotations when he stated that '*parce que l'anglais est sans aliénation politique*.' This means that the objectives of teaching English to Algerian university students is a desirable end, especially in the situation whereby pre-university education is largely arabised while science and technology in the world is largely anglicised.

The problem of this research will be discussed after examining the current model of education and its attendant inefficiencies to see how these can affect the learning of English in Algeria, particularly for science students.

1.5 The current model of education and its attendant inefficiencies

The broad objectives of the current model of education have been explained in 'Section 1.2/ Overview of the foreign language teaching policy' and may be described as consisting of statements of desirable ends for educational practices.

Practices consist of means that are assumed realised ends. It is in these means where the inefficiencies attendant to the current model are located.

In this study, two main educational ends (i.e. objectives) are isolated and their attendant inefficiencies of the means used to achieve them are discussed one after the other.

1.5.1 Objective 1 of the current model of education: E.S.T

The objective of the gradual replacement of French by English in Algeria (as stated in 'Section 1.4.1') is to help learners access information fundamental to their subject-speciality in science and technology, as it is claimed that most scientific literature is published in English.

But in practice, reality is quite different. Arab[6] observes:

> *The objectives assigned [to the teaching of English] are rarely reached because of the quantitative and qualitative deficiency, as well as the instability of the staff teaching in this field; to these difficulties are added the non-existence of appropriate pedagogical techniques and methods and the chronic absence of books and other didactic tools.* (Arab, 1990)

This was also emphasised in the official document which reads:

> *Such a substitution is not obviously possible in the short and middle terms. It implies a total reorganisation of the teaching of languages at primary and secondary levels, and would consequently be confronted to a shortage of teaching staff and problems of availability of didactic tools.*[7] (Teaching Services, Ministry of Higher Education, 1986)

Furthermore, in spite of the fact that " *40% of the national budget was devoted to educational and vocational training as by the mid-seventies"* (Djité, 1992), considerable problems have accompanied the arabisation of primary and secondary education, and democratisation of primary education. According to Bennoune:

> *In 1977-78, of 100 pupils enrolled in primary schools, 20 would drop out before the end, and 40 would fail to pass the examination for the certificate of primary education (and therefore would not be allowed to stay at school).*

---

[6] « *Les objectifs assignés [à l'enseignement de l'anglais] sont rarement atteints en raison de l'insuffisance quantitative et qualitative, ainsi que de l'instabilité du corps enseignant dans ce domaine ; à ces difficultés s'ajoutent l'inexistence de techniques et méthodes pédagogiques appropriées sans parler de l'absence chronique de livres et autres supports didactiques. »* (Arab, 1990)

[7] « Une telle substitution est à l'évidence irréalisable à court et moyen terme. Elle implique en effet une totale réorganisation de l'enseignement des langues au niveau primaire et secondaire, et se heurterait de ce fait à des obstacles d'encadrement et de disponibilité de supports didactiques. » Direction des Enseignements, Ministère de l'Enseignement Supérieur, 1986)

Thus,

> *....of the remaining 40 pupils about 18 would be admitted to high schools, 16 of which would be candidates for the baccalauréat examination*. (Bennoune, 1988: 227)

and as a result, "*The pass rate was 25 percent in 1978-79. This meant that only 4 out of 100 would have the chance to go to the university." (Ibid)*

Whereas the arabisation of primary and secondary education was dictated by the demands of a wish to increase access to education and reduce elitism, it is obvious that this policy had also resulted in "*the absence of qualified teachers, adequate textbooks, laboratories, and even classrooms"*, in that " *the quality of education was impaired, a fact that undermined particularly the scientific and technical option."* (Bennoune, 1988: 223).
It is not the fact of using Arabic that affected the quality of teaching / learning, but because the prerequisites, sine qua non conditions for such an educational decision were not available.

In the same line of thought, Miliani finds that,

> *The democratisation of primary education that is a shift from highly selective and competitive, to a stage where it became considered as an inalienable human right for the benefit of all, ... was not also supported by the pedagogical, didactic and human resources necessary to provide a quality education for all.* (Miliani, 1996: 7)

This mass schooling was also accompanied by school-drop-outs which generated a new phenomenon that affected the whole society. According to Miliani (1996: 7):

> *These drop-outs and other young unemployed who developed a kind of anti-establishment attitude, became known as 'hittistes' (those who lean on walls).*

Literally, but essentially jobless, idle and available for delinquency, or the trip abroad.

Funnily enough, this situation induced as a corollary an increasing interest for foreign languages, contrasting the debasement of learners for education in Arabic and a kind of under-evaluation of local cultural values. Indeed, French and English came to mean a chance to go abroad, to have a qualification for a job, and climb in the social ladder of life, thanks to a 'decent earning.'
English, more than French, has supplied thousands of unemployed young Algerians with dreams of a better life, not only because it represented the might of the Anglo-Saxon world namely the U.S.A (the American Dream), but certainly, because it offered much wider geographical horizons of fulfilling a career, as France became a more and more saturated destination of emigration.

Another reason contributed to the growing passion for English, beyond the absence of the emotional burden the recent 'common' or 'shared' history of Algerians, the French language drags on. As France has in the past represented a place of emigration for unqualified labour, physical human capacities were needed for low rated jobs left over by the French. The English world required by necessity a minimum level of education, and therefore the English speaking Algerian community postulating for emigration, tends to be more and more qualified. This criterion gave a larger choice of destinations in the world, including areas where English is the official language. This status, we are thinking of the Arab and Asian worlds, gave Algerians almost the same linguistic competence as the local populations of the host countries, and drastically increased needs away from French, back to a non-native English community in which they found themselves competing with others, with equal chance of getting a fair deal of opportunities.

The choice of Arabic as an instructional medium and democratisation of primary education would thus seem (for reasons mentioned above) to have resulted in the deterioration of the quality of primary and secondary education. Inevitably Higher Education was also likely to be affected by these developments in primary and secondary education.

1.5.2 **Objective 2 of the current model of education**: English as a substitute for French.

As stated in 'Section 1.4.2' English has been taught in Algeria as a substitute language to French because of its freedom from colonial connotations.

The consequences of a real replacement of French by English at different levels of the educational system (i.e. English as a substitute for French) may affect not only the educational system, but the Algerian society as a whole. Such a decision will result in a society :

> *Where French would still be used but in an inappropriate way, and where English would be spoken only by a minority, in a mediocre way, and quickly forgotten*. (Kapil[8], 1989, in Slougui, 1993:3).

In a word "*une anglicisation qui sera inévitablement sauvage*" (Arab, 1990), because the linguistic vehicle chosen to facilitate development is the language in which Algerians do not have efficient command.

My concern, if this is true, is that we need a description of the situation, a diagnosis of problems and a prescription.

1.6 Statement of the research problem

Algeria is well aware that the economic and commercial development cannot be achieved through the Arabic language as this latter itself needs change and development. To gain membership in the world community and be part of this 'global village' where mainly the English language is used for such interchanges (economic and commercial), students (especially those in science faculties / departments) should be well-initiated in English to meet their own needs and the needs of the national education which aims at such

[8] « *où le français sera encore parlé mais mal, et où l'anglais sera connu de peu, de façon médiocre et bien vite oublié. » (Kapil, 1989, in Slougui, 1993:3).*

development. To meet these needs, appropriate courses should be designed for students at university level.

However, and as stated in the 'General Introduction', designing a syllabus for science students (in the different departments where English is not the medium of instruction) has long been based on teachers' intuition and common sense. And as such, some syllabuses, have focussed on grammar only and others on lexis only, using 'E.S.P. concocted texts' (instead of authentic texts). Hunston *et al.* (1997) note:

> *It is common to see grammar and vocabulary as separate areas of language teaching and learning. Many course books have separate sections on grammar and vocabulary; syllabuses list grammatical structures and key vocabulary items separately; students are described as being 'good at grammar' but having a 'limited vocabulary', or vice versa; grammar and vocabulary are often tested separately. (*Hunston *et al.*, 1997:208).

In science institutes in Algeria, the great majority of the teachers of English are language teachers (newly qualified with a B.A. degree), and to a much lesser extent, English is taught by few subject-specialists. As a result, we get two separate teaching methods: one focusing on grammar and the other focusing on lexis.

1) Focus on grammar only: Most syllabuses for teaching / learning English in science institutes were/are mainly based on teaching grammar for two main reasons:

a) Most teachers of English in science institutes were / are not members of the scientific community, and as such they themselves did / do not master properly the frame of reference i.e. the scientific lexis and concepts.
b) It is taken for granted that science students already master scientific lexis and concepts.

This was true, to some extent, as 'former Algerian students' mastered French well and this fact helped them in learning English. Much of the scientific lexis is the same in both languages i.e. French and English because they are Latinate names.

The problem nowadays is quite different as science students are completely ‘arabised’ and their knowledge of French is quite low.

2) Focus on lexis only: In teaching the English component, subject-specialists focus more on lexis and the explanation of experiments and processes, and as such ignore the teaching of grammar.

Therefore, focusing on one area of language teaching at the expense of the other may lead to an inefficient language training.

The methodology I would adopt to investigate this problem would be to answer the following research questions.

1.7 Research questions

The results of the needs analysis (done in ‘Chapter Three’) provided the motivation for formulating the research questions which will be investigated throughout this study. These results show students’ dissatisfaction related to the English course. The students found that they have not acquired much knowledge in their specific field because either the course was not well-taught, or because they found the English language itself too sophisticated.

Therefore, the three questions that the present research addresses are:

1.7.1 How is the discourse community under study organised in terms of teaching / learning practices ?

1.7.2 How does the English language function in extremely specialised discourse ?

a) How are concepts realised in biology discourse ?
b) What are the ways of expressing them ?
c) What grammatical structures do these concepts combine with ?

1.7.3 How can the research findings be dealt with in developing an appropriate teaching / learning methodology for university science students ?

In order to achieve this purpose, the community of biologists (undergraduate students at the university of Oran) has been selected.

1.8 Choice of community (undergraduate biology students)

The research population profile (i.e. undergraduate biology students) is based mainly on information collected through questionnaires and informal interviews of the Head of Department, some teachers and the students. In this part, we try to give information on how a student becomes a biologist in Algeria by answering the following questions:

(i) What are the conditions for becoming a biologist ?
(ii) Where are the students of biology trained ?
(iii) What future perspectives do these students have ?

(i) The conditions for becoming a biologist

In order to study biology in Algeria, students need to have their 'Baccalauréat' (science stream) with a good mark in 'Natural Sciences' at the Baccalauréat exam.
They, then, have to do a four- year course which includes a two-year common core course and a two-year speciality study (Biochemistry, Plant Biology, etc.), with English taught as a separate subject.

The students have to attend biology lectures, tutorials and laboratory work (Travaux Pratiques 'T .P') in their respective departments (Arabophone or Francophone), with English taught to help them access scientific information necessary for completing the lecture knowledge or for using it in the 'mémoire / project' writing- up. The students have to pass all their on-going exams and 'the mémoire viva' to be able to join the community of biologists.

At the end of the four-year course, the students get their D.E.S. (Diplome d'Enseignement Supérieur). This is so at the university of Oran. In some universities (Annaba and Mostaganem, for example), the biology course lasts five years, at the end of which students get a degree of 'Ingéniorat d'Etat.'

(ii) The Department of Biology

The students follow their courses at the Department of biology / Faculty of Sciences. At the university of Oran, the 'Department of Biology' (previously called 'Institut des Sciences de la Nature', then 'Institut de Biologie') is composed of two main departments:

a) The 'arabophone' department where biology is taught in Arabic; and
b) The 'francophone' department where biology is taught through the medium of French.

For the former programme (1984 to 1997), each department runs courses in different specialities (in the third and fourth years of study) as the tree-diagram below shows:

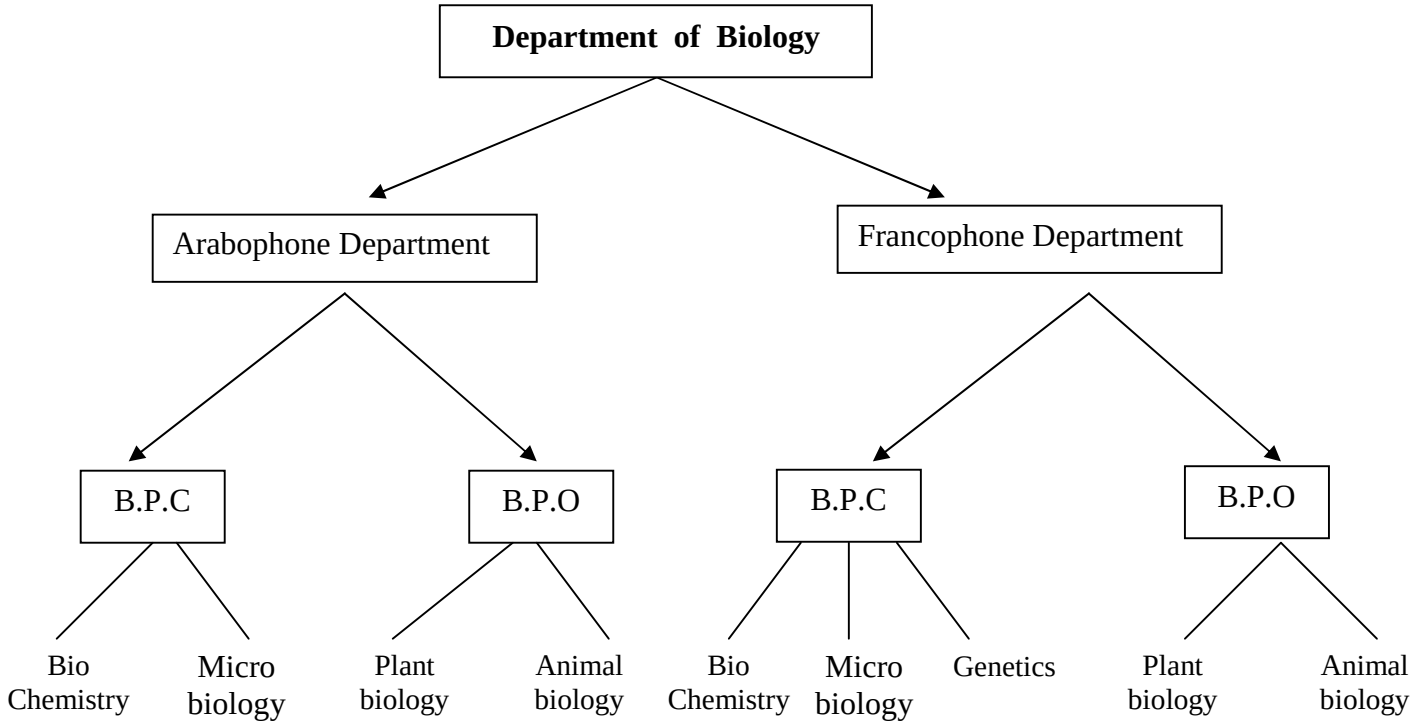


Figure 1.1: Specialities taught at the Department of biology

('*B.P.C*' stands for 'Biologie Physico-Chimique' and '*B.P.O*' for 'Biologie des Populations Organiques')

Figure 1.1 shows that 'Genetics' is taught only in the 'Francophone' Department, inducing a discrepancy in syllabus resulting from a linguistic difference, giving a privileged status to the students of the francophone Department.

The programme introduced in September 1997 (and which is still in practice nowadays) shows few changes. The same specialities are still taught in both departments (Francophone and Arabophone). However the number of students in the Arabophone department has decreased.

The total number of students at the 'Department of Biology' is 1700 for the academic year 2005-2006 (according to the 'Vice-Doyen de la Pédagogie') and it is distributed as follows:

a) First year: 530 students ;

b) Second year: there are '280 students registered at the Francophone Department' and '15 students at the Arabophone Department.' A number in total contradiction with the objective of the national education i.e. Arabisation of the educational system, and the number of students in the arabophone department that was superior to the number of students in the francophone department in 'the fourth year' in the 1990s. In the academic years 1993-94 and 1994-95, there were:

- 250 students in the Arabophone Department in the four specialities; and
- 230 students in the Francophone Department in the five specialities.

c) Third and Fourth years : 640 students.

There is also a three-year course study called 'D.E.U.A' (Diplome d'Enseignement Universitaire Appliqué) where English is taught as a compulsory component during the three years. The number of students registered on this course is 170 for the academic year 2005-2006.

In 2005-2006, The L.M.D (Licence / Master / Doctorat) programme was introduced at the Department of Biology for students in 'Sciences de la Nature et de la Vie' (SNV). The number of students registered on this programme in 2005-2006 is: 60 students;
These figures could be represented in a table as follows:

Department of Biology	
Programmes	**Number of students**
D.E.S	1470
D.E.U.A	170
L.M.D	60
Total	1700

Table 1.2: Number of students on each biology programme 2005-2006

The content of the LMD programme in the first year (semesters one and two) shows the different modules to be taught e.g. Animal biology / plant biology, chemistry, etc., together with the introduction of the teaching of the French language (22.5 hours, each semester). This is so because the students come with a completely arabised background, and the teaching in the LMD programme is done in French.
In the second year (Semesters three and four), English is introduced (22.5 hours, each semester) together with the other biology content modules.

The introduction of English in the LMD programme also shows one year void between the secondary school and the university instruction. The same practice is followed in the D.E.S. curricula (1984-1997 and 1997 to date), where English is introduced in the third year of university study , thus showing a two-year void.

(iii) The biology students' future perspectives

The students who graduate may:

- work in laboratories (state-owned or private), or in the agricultural sector; or
- teach 'Natural Sciences' in secondary schools; or

- teach at the university under the following conditions:

The students who are ranked the first of their department may undergo a national exam. If they pass, they get scholarships to further their studies (to prepare their 'Magister') and later, they can prepare their 'Doctorate'.

1.9 Reasons for the choice of community

Five main reasons motivated the choice of this specific community.

(i) This community more than any other at university level relies on scientific data that is almost entirely carried out through the English language; since even the French community of biologists handles a good quality of English, and tends to communicate outside the European world using this language.

(ii) **The most sophisticated and fairly recent, updated discoveries in this field** are the result of the leadership of the U.S.A. and Great Britain because of the huge amount of money they invest in this industry, giving an unquestioned role of leadership.

As part of this research, in the year 1993-94, I listed and hand-counted all the resources available in the 'fourth year biology library' at the University of Oran, and found out that there were 175 books, magazines and journals. Out of these 175 documents, 136 were published in English (i.e. 77.71 %)

(iii) **The Instrumental value of English in a biology discourse community**

In Algeria, in scientific education, it is agreed that the test of 'membership' is to write a 'project/mémoire', in the final year of study, in Arabic or French (in the Arabophone and Francophone Departments respectively) using English language source material.

To become an accepted member of the community of biologists, students have to be able to demonstrate the ability to read scientific literature written in English in their subject speciality.

Slougui (1993) attempted to study the instrumental value of English in a biology discourse community in Constantine (Algeria) and found that citation patterns in Magister theses written by Algerian students are distributed as follows:

- Arabophone theses: English is the prevailing language of citation as it represents 88.9% of the total input. French, which ranks second, constitutes 9.02% of the total references i.e. 1% of the English citations. Arabic is almost absent (2.79 %).
- Francophone theses: English is predominant (61.62 %), French partially meets the francophone needs (37.94 %), almost half of the English input. Arabic is totally absent.

This reason (**instrumental value of English**) is important as it shows that the undergraduate students in both departments will also need English in the future, if they undertake post-graduate studies.

(iv) A personal involvement in the teaching of E.S.P. in general, and my short experience of teaching English to this community, stimulated my interests and provoked queries that gave to my pedagogical duties, challenging quality determined by convex research activities.

The remedial function / objectives of this reflection will certainly boost experiences in this subject and bring in better approaches to the problems of 'scientific English' preoccupations.

(v) Lack of research in this field

To the best of my knowledge, no study has attempted a detailed examination of the 'biology genres' written in English and read by biology students in Algeria, in order to find out what difficulties these genres may present to non-native students of English in order to suggest remedial work.

- Hadj Moussa, (1998) in her Magister thesis, suggested a method for teaching English to the students of biology. She reviewed the literature on discourse analysis and examines the implications of some concepts of discourse analysis for E.S.P teaching with special reference to teaching the reading skill and concludes by a sample lesson.

- Baka, in her PhD thesis (1989), analysed spoken biology texts in a native-English environment. She used nine recorded biology lectures (of 43,000 words) given by native-British lecturers for first year undergraduate students at the University of Aston. She then identified features of discourse organisation.

These experiences remain isolated attempts, and cannot provide us with a systematic squaring of the whole field of research under investigation. My work may not pretend to do a much higher impact from this point of view. Nonetheless, it will put forward a query, that may have some statistical degree of 'truth', and will represent a landmark for further investigation to be implemented in the frame of a research programme.

1.10 Conclusion

'Chapter One' establishes the research area by defining the research problem in reference to a particular academic discourse community i.e. the undergraduate biology students.
It states the research problem and objective, and introduces the research questions.
It specifies the community for study (i.e. the undergraduate biology students) as far as the English language course is concerned and explains the reasons for such a choice.

We, then, described the community of biologists in general terms. We showed briefly the process through which, under normal conditions, a student goes in order to become a biologist (i.e. from a neophyte to a member of the community of biologists).
'Chapter Two' tries to develop the notions of 'discourse community' and 'discourse analysis' relating to this research.

CHAPTER TWO

REVIEW OF THE LITERATURE AND RESEARCH METHODS

CHAPTER TWO

REVIEW OF THE LITERATURE AND RESEARCH METHODS

2.1 Introduction

Three elements are necessary for the description of the teaching / learning situation under investigation: the participants (teachers and students), the place (the classroom) and the institution (the department where the participants teach or study). These three elements are continually in contact and interact with each other (interaction between teachers and students, and between students and their peers, in the classroom where the teaching/learning takes place and which is part of an institution, the Department of Biology). They, thus, form a 'microcosm' represented in a 'discourse community' (the scientific community) which interacts through 'discourse', and not through language only.

Therefore, and based on the research questions (in 'Chapter One'), the theoretical framework of this research is developed through two concepts:

1) The notion of discourse community
   We shall describe the community of biologists through the notion of 'Discourse Community' because we stated in the 'Introduction of Chapter One' that 'access to societal power for this community is through education and technology, based mainly on familiarity with the norms of the written language of scientific discourse', this being the main objective of the English language course; and

2) Discourse Analysis (the description of how the English language functions in extremely specialised discourse i.e. Analysis of scientific discourse)

In this chapter, the review of the literature of these two concepts (Discourse community and Discourse Analysis) is given, and the research methods are specified.

2.2 The scientific community's language impact on social communication

As a language is not studied in isolation but within a social group and / or cultural setting, in this part of the research, we try to show how members of the scientific community communicate with each other and what type of language they use among themselves.

We stated in the General Introduction that the aim of this study is to achieve a richer understanding of the nature of the scientific discourse in order to improve the quality of tuition for non-native speaker students who are preparing to write their 'projects/mémoires' and become active members of the society. Therefore, it is necessary to explain the contexts in which this discourse is used.

As the scientific community is part of the society as a whole, I found it useful to describe the notion of society before I describe the notion of 'discourse community', because it seems, from some definitions, that the characteristics of a society apply exactly to those of a discourse community, the latter representing a smaller unit.

2.2.1 The notion of Society

Barakat [9] (1986) sees society as being dynamic because:

a) A society lives in a continuum i.e. it changes continuously because of some internal and external factors and forces;

b) A society is continuously developing;

c) Human beings are social and this means that they live and develop in a society through social interaction (and not in isolation from other members of the society.) (Barakat, 1986:29)

To Barakat's definition, we can add Thom's (1972) mainly because his description of society includes other notions useful to the description of the discourse community under study. According to Thom, a society is 'a metabolic form' when he notes that,

[9] The author's translation from Arabic.

> *Une société est une forme métabolique d'abord parce qu'elle survit aux individus qui la constituent; il y a un flux permanent d'individus, constamment renouvelés, qui assurent la permanence de la forme Sociale. (Thom, 1972:318)*

Thom, then carries on stating that,

> *De plus, il existe entre ses membres une interaction continuelle qui assure la stabilité de la forme. Cette interaction peut souvent se matérialiser par une circulation de complexité, d'information à travers tout le corps social. (1972:318).*

Thom's definition is three-fold:

a) According to his view, a society is a metabolic form because there is a constant renewal of members of this society.

   This notion of member renewal is useful for the development process of the society and of the discourse community, and is therefore possible only under normal conditions .

This point is further emphasised by Thom when he describes the global form of a species (i.e. the individual). He said :

> *Il y a lieu de lui [the global form] associer la totalité des mécanismes stabilisateurs qui assurent, en des conditions approximativement normales, la multiplication de l'espèce.* (Thom, 1972:161).

Le Corbusier[10] makes similar observations when he states that 'existence' is 'occupation of space' in:

> *Prendre possession de l'espace est le geste premier des vivants, des hommes et des bêtes, des plantes et des nuages, manifestation fondamentale d'équilibre et de durée. La preuve première d'existence, c'est d'occuper l'espace.*
> (Le Corbusier, in Thom, 1972)

b) Moreover, there needs to be interaction between the members maintaining the stability of this form.

c) That this interaction often materialises as a circulation of complexity, of information, through the social body. Not only is the element of member renewal important but there also needs to be interaction between the members to maintain the existence of the discourse community.

This interaction is often done through the circulation of information either socially, through conferences and meetings, or through other mechanisms of intercommunication (Swales, 1990: 24) i.e. journals, books, correspondence letters, etc.

Thom's notion of society is most useful for the description of the discourse community under study because our discourse community needs a constant renewal of members in order not to die out. This element of renewal is essential in the development of a society in general, and in that of a discourse community in particular.

[10] Le Corbusier (Charles Edouard Jeanneret, 1887-1965) a French architect and town planner, is one of the most famous architects of the 20th Century. His buildings have become today typical monuments that illustrate his theories of space and human movement dynamics. He contributed to many urban projects with new social distribution of space. He also built one of the most curious and intelligent combinations of light, space and volume for metaphysical concentration, the 'Cathedral of Algiers: Le Sacré Coeur'. It is conceived as a space of communion i.e. total communication between members of the same group, in this case the congregation, through mutual comprehension and apprehension.

A social network analysis is also another method that shows how an individual relates to other individuals in a society.

2.2.2 Social networks analysis

According to Brierley and Kemble (1991:163), in any society, whether it be the wider speech community of a political state or a particular working environment, there is always a need to pass and exchange messages. The means which this is achieved represents a form of information system. An information system can be thought of as a network used to convey:

- new information;
- signals;
- existing information; the norms and values of a society;
- expectations that society has of an individual, group, or groups.

Also to Brierley and kemble, because language is used and developed by people who are subject to changing environment(s), any information system is dependent on the people who use it. An information system is therefore very much a social network which uses a sophisticated form of coding, i.e. language (Ibid). This is the case of the biologists in Algeria (who form a social network) who need to use the English language to pass and exchange information.

Two academic works (from the field of 'social networks') quoted by Scott (1991) and which include useful notions for the description of the discourse community are selected and discussed below. These are: Warner's and Mitchell's codification of social network analysis.

I do not intend to use the mathematical techniques of social networks analysis, but I shall use only a few concepts from the field of social networks analysis in relation to this research.

2.2.2.1 Warner's work

Warner states that *"the social organization of a community consists of a web of relations through which people interact with one another."* (Warner, in Scott, 1991)

Interaction is vital between the members of the discourse community under study. The existence of this discourse community entails absorption of members between whom there needs to be interaction. Crane (1972: 98) notes that *"...in technology as well as in basic science, the social interaction facilitates the diffusion of information."*

In the present research, social interaction of the whole discourse community of biologists is done through informal discussions at work (face to face), in national and international conferences and meetings, or by reading journal articles, books, and so on. from the biology field.

2.2.2.2 Mitchell's codification of social network analysis

Mitchell's codification of social network (1969:36-9) is useful because it focuses on some key elements necessary for the survival of the discourse community under study.

For Mitchell, interpersonal networks are built from two different ideal types of action: 'Communication' and 'Instrumental or 'Purposive' type of action.'

(i) Communication

So that our discourse community survives, there needs to be communication between members of the discourse community through transfer of information. This type of action joins Warner's web of relations through which people interact with one another.

This notion of communication through transfer of information is the backbone of the discourse community.

Crane also speaks about the importance of knowledge diffusion from one person to another and states that, *"the growth of scientific knowledge is a kind of diffusion process in which ideas are transmitted from person to person." (Crane, 1972).* She then adds a concept of '*contagion'* which she explains as the need for individuals to involve other members of the community, a social system in adopting innovations, and make them share the benefit of intellectual production as the process of social improvement, and cultural progress.

This ‘contagion’ effect is important to take place among the members of the discourse community under study so that mature members transmit their knowledge to novice members.
This also suggests that if a member of a discourse community is being influenced, he is then absorbed in the discourse community, and this is a sign of a good social integration.

However, if the opposite case happens (if no contagion effect takes place) the situation could be disastrous. Crane warns that, *“...lack of interpersonal communication seriously weakened the intellectual development of the field. ... and knowledge failed to cumulate.”* Crane (1972, 25-6). This could, thus, cause ineffective development of the whole nation e.g. causing drop-outs and leading again to the problem of ‘hittistes’ mentioned in ‘Chapter One’, i.e. to the exclusion of whole social categories of the national group entity.
This is why education remains one of the most efficient and unavoidable institutions of social cohesion. It determines the individual’s ambitions, and provides him with the means to fulfil his objectives and be recognised by his peers as a positive participant of the social network and the community wealth.

(ii) **The instrumental or "purposive" type of action,** which involves the transfer of material goods and services between people.

In the case of the discourse community under study, material goods and services are books, journals, etc. used to transfer information necessary to the members of that community. This instrumental type of action is relevant to the present research as the latter is interested in investigating the genre-type read by members of the discourse community studying biology.
The concept of social networks as seen by Warner and Mitchell is particularly useful because it focuses on social interaction through communication and transfer of material goods and services.

This in turn is useful to the present study as its main focus is the scientific community being “*an association of persons bound together neither by law nor chain of command but by the communication of information, through journals, conferences, informal discussions and other channels.*”(Kneller, 1978: 190)

2.3 The notion of community

The concepts of society and social networks are mentioned above to show how their defining characteristics apply to those of a discourse community. However, because these concepts seem too large for the present study, I therefore believe the analysis of the problem would be more useful at the level of the 'community' as it is the focus of the present study.

2.3.1 Definitions of Community

Hymes reserves the notion of community for "*a local unit, characterized for its members by common locality and primary interaction.*" (Hymes, 1974: 51, in Milroy 1990: 15).

Miller sees a community as:

> *A rhetorical construct constituted by its characteristic joint rhetorical actions, its genres of interaction, of getting things done, including reproducing itself.*
> (Miller, in Swales, 1992)

These two definitions are particularly useful as they highlight the notions that the community under study include. These notions are that the community under study is a local unit and that it is characterised by interaction. It is also a '*rhetorical construct'* (Miller) as we are interested in its genres of interaction, mainly in those including reproducing itself.

'Rhetorical construct' is understood as a pattern of communication developing several channels of interaction, verbal and non-verbal. The non-verbal covers up all the codes that the traditions and rituals perform, whether in happy or sad ceremonies, that the group institutionalises, in order to help the individual go through them, with the feeling that he remains a member of the group, and that the group identifies itself to his experiences.

The above two definitions are also useful because they incorporate the concepts of '*interaction*' and of 'r*eproduction*' as does the definition of society mentioned in earlier parts.

Roe (1993) defines a community as "*a group of individuals defined by their purpose*" and that "*language is used in-vivo through community communication because it mediates a purpose between people.*" He carries on saying that "*learning a language is gaining acceptance by a community.*" This is true for the Algerian scientific community where E.S.P. is taught to help these scientists communicate in foreign communities.

As a researcher, I am primarily interested in my learners performing these roles in those foreign discourse communities, be they invisible colleges, international conferences and meetings, journals, books, and so on.

2.3.2 'Speech' versus 'linguistic' community

Mey argues for a distinction between 'linguistic' and 'speech community' and finds that:

> *The linguistic community is analyzed on a metapragmatic level and concentrates on norms of speech,*

while,

> *The speech community is characterized by its concentration on the pragmatic rules of speaking.*
> (Mey, in Hudson, 1993)

Other writers have further emphasised this definition of speech community according to which it can mean "*shared linguistic rules*" (Bloomfield), "*Shared norms*" (Labov), or "*shared regularities of use.*" (Fishman).

Swales's definition seems to be more holistic as it includes all these notions. According to Swales '*a speech community*' would be described as "*a homogeneous sociolinguistic assemblage of people who share place and background*" (1992:19), and that "*a speech community typically inherits its membership by birth, accident or adoption.*" (Swales, 1990:24). The common feature of the notion of a 'speech community' and 'discourse community' is as Bloor (1998:57) notes "*a contribution to a model of how language is used in society.*"

The present study is neither concerned with 'speech community', nor with 'linguistic community' because in Algeria, the English course for the community of biologists has as its stated objectives the teaching of skills which would enable students access information fundamental to their subject speciality in order to write their 'mémoire' and become efficient members of the community of biologists.

It, therefore, focuses on 'discourse community', as access to 'the target discourse community' is based mainly on familiarity with the norms of the written language of scientific discourse.

2.3.3 Discourse community: Definition and criteria

2.3.3.1 Definition of Discourse Community

A discourse community would be described as:

> *a heterogeneous sociorhetorical assemblage who share occupational or recreational goals and interests"*, and that "*a discourse community recruits its members by persuasion, training or relevant qualification.*
> (Swales, 1990)

In this regard, I am taking the Algerian undergraduate university students as such a community. This is so at least in relation to the English course which is aimed at developing their English language competence through training. This English course has as its stated objectives the teaching of skills that would enable students to have access to information specifically needed to carry out the purpose of their academic career that shapes in the form of a 'mémoire' i.e. an end-of-degree dissertation. This qualification is the first step of acknowledged capacity to adhere to the community of biologists.

Herzberg's definition is also useful to our context because of the idea of 'discourse' that it includes. He states that discourse communities are also used to:

Signify a cluster of ideas: that language use in a group is a form of social behaviour, that discourse is a means of maintaining and extending the group's knowledge and of initiating new members into the group.
(Herzberg, 1986: 1, in Swales , 1990: 21)

2.3.3.2 Criteria for describing a discourse community

In this research, twelve criteria for describing a discourse community have been selected: Swales's model (1990) and six management science elements under the acronym 'C.A.T.W.O.E.'

The selection of six management science elements in the description of the university discourse community is justified by the fact that in this research we aim at managing the teaching / learning situation, as stated in the 'General Introduction' of this research.

In management science, *systems thinking* which is "*sometimes referred to as the systems approach, recognises that the relationships, processes, connections and inter-connectivity are what underlie the dynamics of any system.*" (Brierley and Kemble, 1991:169)

System thinkers believe that "*a system viewed as a whole is always greater than the sum of its parts.*" And that "*all systems are wholes whose structures emerge from the interaction and interdependence of their parts.*" Their view is opposed to 'the philosophy of Reductionism' which view is that everything in the universe could be understood if it was studied in small parts.

According to Brierley and Kemble, it was acknowledged by several systems thinkers (mainly represented by Checkland, Mumford, etc.) that people, and not the mechanical tools in the work place, are the centre of the information systems (1991:168), and the same fact applies to foreign language teaching

In foreign language teaching, people are central in that system i.e. teachers and students interact together. So, for systems thinkers, how and why people are central needs to be analysed and understood in terms of the overall system, as will be done in 'Chapter Three.' Also, as far as teaching is concerned, methodology cannot exist apart from the system it operates in, as it has to include everything that people do, or are likely to do, as it is explained in the following part.

2.3.3.3 Arguments for using Swales' s model and CATWOE

As "*the true discourse community may be rarer and more esoteric"*, (Swales,1992: 20), the present study proposes twelve criteria to describe a discourse community.
Swales (1990: 24-7) proposed six defining characteristics of a discourse community. These are: *'goals', 'mechanisms of intercommunication', 'participatory mechanisms', 'genres', 'lexis' and 'a threshold level of members.'*

For the present study, Swales's characteristics are useful but not complete. I, therefore, propose to add six other management science elements (summarised by the acronym 'CATWOE', Checkland and Scholes,1990:35) that are necessary for the description of the discourse community under study, and which can be applicable to the description of some other academic discourse communities.

The 'CATWOE' elements are used to complete Swales's criteria for the description of a discourse community. I am proposing to use both the CATWOE elements and the Swales' criteria because the former does not include '*genre*' and the latter does not include *'customers'* and '*constraints.'* These three elements (*genres / customers / constraints*) are essential for 'the transition phase' which is socially important in Algeria. By 'transition phase', it is meant the phase where '*the transformation process'* should take place, thus allowing the undergraduate students in the biology institute, who are not yet members of the discourse community, become members of this discourse community.

In his model, Swales ignored the element '*customers*' which is important as it can represent either the students themselves, the institutions where they study or the institutions that they will join for work. The students as '*customers'* (Checkland et.al.,

1990) are important because the essence of the present research is to train these students '*customers'* to become members of the 'Target Discourse Community.' (Swales, 1990)

The institutions are also important *'customers'*. We have to consider them in order to know what 'product' to give to the students i.e. what do these institutions (as customers) want / need their students to be able to perform ?
'*Constraints'* also represent another element that was ignored by Swales.

In Algeria, biology students meet impediments in using scientific literature written in English. The aim of the present study is to find out these 'constraints' that prevent the transformation process (referred to earlier) from taking place.

The management science elements "CATWOE" do not include '*genre*'. This latter is important for the present study because it is the element through which the students interact in order to become members of the "target discourse community." (Swales,1990). This interaction is important for this discourse community because it facilitates the diffusion of information.
Therefore the 'CATWOE' elements and the Swales's model allow one to view the discourse community under study i.e. the community of biologists as including the following criteria:

The **"CATWOE"** elements: *Customers / Actors / Transformation process / Weltanschauung / Owners / Environmental constraints.*

- **Customers:** are the beneficiaries (when the results are positive) or the victims (if the results are negative). These could be either 'the students' or 'the institutions', or both;
- **Actors:** Those who would do the transformation. In our case, 'the actors' are 'the teachers';
- **Transformation process**: the conversion of input to output (the process by which undergraduates join the members of the target discourse community) that is the biology course and the English language course.

Yet in this research, we are concerned by the English language course only. We are aware that ideal research conditions would require a parallel investigation through all biology courses so that the 'Transformation Process' develops in complementarity and homogeneity, since 'compartmentalisation' in separate teaching units, affects the overall efficiency of this transition from a neophyte position into a full community member.

- **Weltanschauung:** the worldview which makes this transformation meaningful in context;
- **Owners:** those who could stop the transformation (the decision makers, the institution);
- **Environment constraints:** constraints that could stop the transformation from taking place.

Swales' s criteria for establishing a discourse community (undergraduate students) permit one to add another six elements . These are:

- **Goals:** the common goals of the undergraduate biology students are 'assumed' to become efficient in using the English language in order to gain knowledge in their subject speciality.
- **Their mechanisms of communication among its members:** are lectures and tutorials, and laboratory work (Travaux Pratiques).
- **The participatory mechanisms** are used for getting information only, because undergraduate students are 'assumed' not to be mature enough to give feedback. This statement may often be an a priori. This issue should be investigated seriously.
- **Genres**: what are the genres which mediate the in-flow of information?

  The genres used by undergraduate students are assumed to be books, textbooks, journal articles, manuals, etc.
- **Lexis:** This community has acquired some specific lexis, but it is not very developed yet;
- **A threshold level of members**: new members need to be initiated into this community so that it does not die out, that is why this phase, 'the transformation phase' is socially important. That is also the reason why the notion of 'society' and 'social networks' have been mentioned previously.

Through the information given in ‘Chapter One / Choice of community of biologists’, and the criteria developed in this chapter, the community of biologists in Algeria, under normal conditions, can be represented diagrammatically as follows:

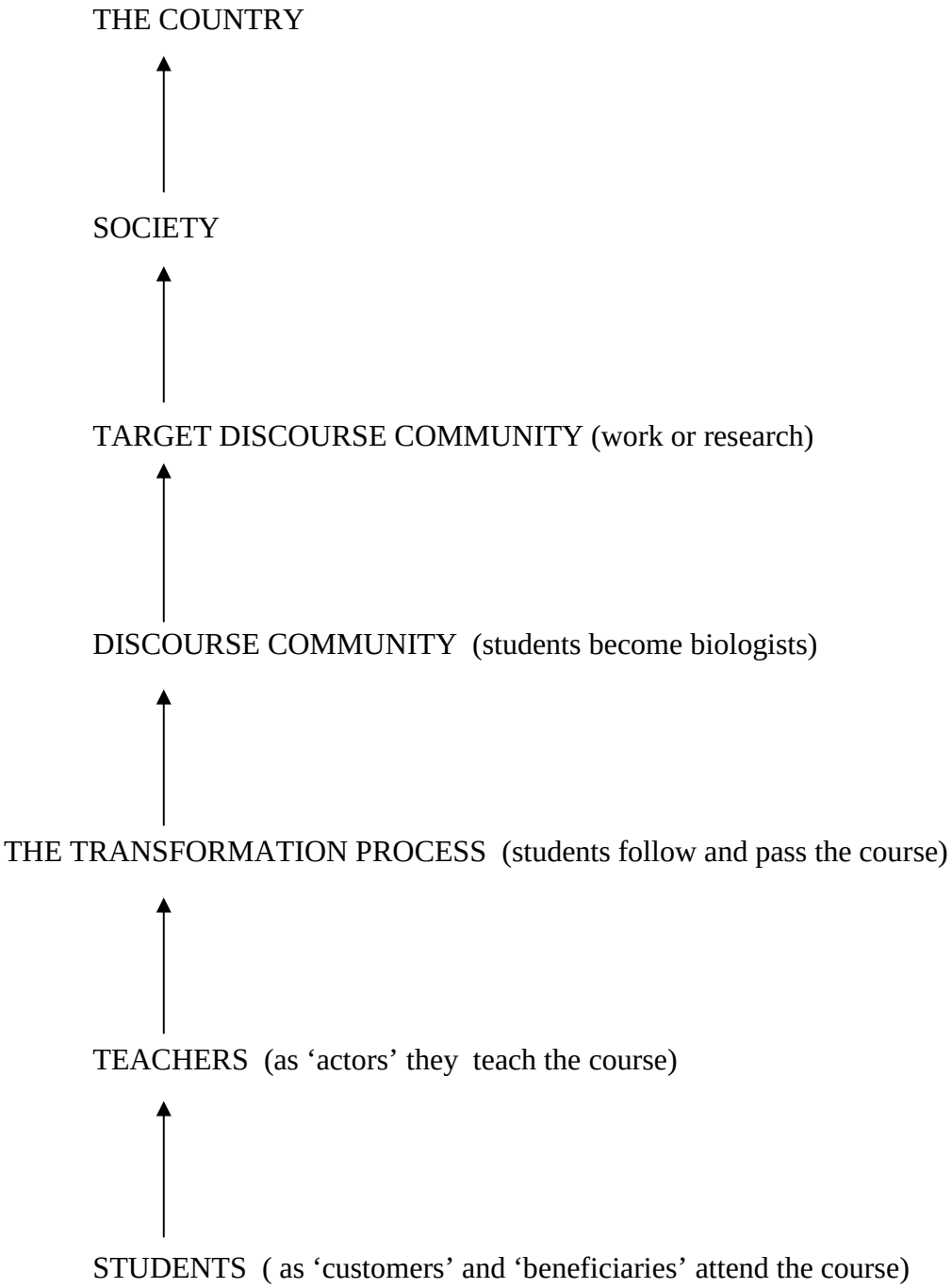


Figure 2.1: Evolution of the university discourse community.

As the community of biologists goes through a process, therefore it could be best represented as follows:

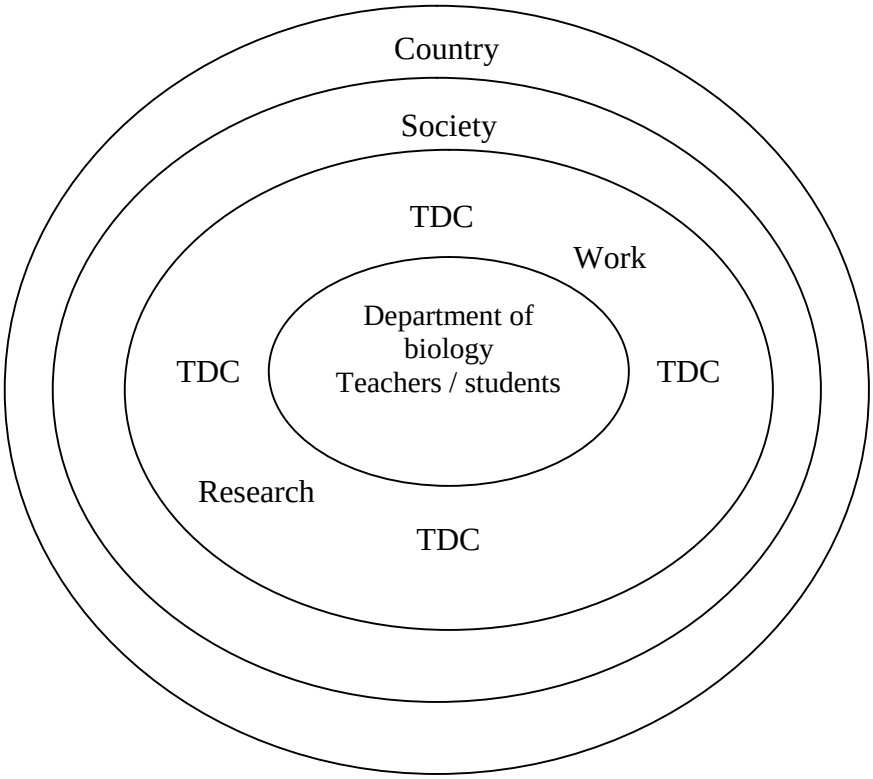


Figure 2.2: Role of the Transformation Process

The circles of 'Figure 2.2' illustrate the various groups and sub-groups that the society includes. Pedagogically, it is hoped that the direction of progression goes from the inner circle towards the outer circle. Each line represents an abstract border that can be realised (seen or felt) only through the shifts from linguistic to speech to discourse community '*under normal conditions'* (i.e. the student moving from a neophyte to a member of the discourse community).

However, this evolution does not necessarily imply a movement through a hierarchy of values, since each community has its own codes of integration. In other words, a good biologist may happen not to be integrated in his society, or his country as a nation, e.g.

those who leave the country, are lost for the society and the nation as a whole, since the investment in their education has proved to be of no impact on the communities (social and national).

Such an issue, although secondary to our purpose, remains of a crucial importance to the Algerian institutions since the national scientific community has lost their competence plus their expertise. Fortunately, those who leave the country represent a small fraction of the whole scientific community.

We are also well aware that the concept of a discourse community which is built around the twelve criteria listed above is not easily applied to the university community as a whole, as there might be what Thompson calls "*different degrees of engagement*" (2001: 21). In this context, Thompson (Ibid) explains that it is not certain that all students, especially those at undergraduate level, can be said to aspire to participation in an academic community; their motivation may be more instrumental than integrative. Many students take degrees that will enhance chances for promotion in order to further their careers in industry, or as mentioned above (they leave the country), and have no intention of remaining in an academic community. Others may study for a doctorate for intellectual stimulation and to pursue their interests in a particular subject. In this case, their desire may be to engage *with*, but not necessarily to engage *in*, a community.

The notion of 'discourse community' has been explained above through definitions and criteria. We also focused on the role of discourse in a discourse community.

In the following part, we move to the second concept of the research i.e. Discourse Analysis, by first restating the research problem.

2.4 Restatement of the research problem

Twentieth-century discourse theories witnessed the most important developments in the teaching of English as a foreign language. Attention was given to the development of reading knowledge of the language for the students and scientists who needed access to information published in English.

Many universities throughout the world (in the Maghreb, in Asia, etc.) where English is not the medium of instruction, include in the curriculum of the science faculties the teaching of English, which in most cases, is basically related to the teaching of the reading skill.

However, for many years, the teaching of the reading component has been inadequate because it has been done in a traditional way, i.e. the focus of some teaching methods has been on the sentence level, putting more emphasis on lexis and syntax and ignoring semantics. Consequently, students become *"unable to use this knowledge to recognise the roles which these grammar structures play in the communicative structure of an extended text"* (Riley, 1985:88).

Reading is considered as a communicative act, and as such it involves more than a mere grammatical (lexical and syntactical competences) knowledge. Teaching a foreign language summons so many other aspects of the complex communicative act. Indeed the language remains nothing but a code, but what transforms this code into a semantic content seems to be more determined by the channels this code implements, whether oral , written, verbal or non-verbal. This issue opens so many doors of reflection and as many perspectives of research, that much controversy and stimulating debates illustrate the obsolescence of traditional approaches and the rather up-to-date and yet unpredictable perspectives discourse analysis may provide ahead.

Many scholars in different disciplines, Labov (1972); Sinclair and Coulthard (1975) and Applied linguists, Candlin *et al*. (1976) have found:

> *...inadequate the type of linguistic analysis whose investigations are strictly limited to the internal functioning of the verbal code-phonology, morphology and syntax-independent of the circumstances in which the code is used.* (Riley, 1985: 1).

The authors mentioned above consequently urge other scholars to widen the field of research so as to include *"the external functioning of the verbal code as well, what people do with words."* (Riley, 1985: 1). This quotation suggests that emphasis should be put on

function and communicative competence, rather than structure and grammar as traditionally implemented in most of the teaching contexts.

This point is also backed up by the works of Riley, when he proposes to consider that learning a language is certainly more learning how to *"do things with utterances rather than a mere exercise of assembling sentences"* (1985: 1). Having said that, one does not rule out the importance of the role of structure and grammar in any language learning (particularly in a foreign language). These two elements are essential in language learning, but should not account for the only elements to be taught to aim for the mastery of a language.

Mastering a language in this case must probably assume establishing a balance between the two approaches. It is a fact that such prospect remains determined by the classroom situation, involving a homogeneous level of the learners in the group, the course objectives i.e. a syllabus adopted to their needs (whether long term or immediate) and the pedagogical means available, when and if, since specifically designed programmes are yet to be conceived, often out of field practice.

This is why experiences, today, are likely to teach and unveil many unknown and unpredicted aspects of the journey across to a foreign language. It is admitted now, that theory in this field progresses too quickly for adequate applications in laboratory conditions, since the 'guinea-pigs' are learners who do not wish to put their career at stake, and jeopardise their future with methods that may be original, new, but not necessarily efficient.

It seems the more obvious for foreign language teachers, as they proceed in their practice, that they are faced with a dichotomous process whereby they acknowledge their inability to establish a distinction and a priority between whether they only teach a language, or they also teach a culture. A series of disturbing questions, often come recurrently in this debate.

- Do we teach a language, and then a culture ?
- Do we teach a culture, and then a language ?
- Can we teach one without the other ?

The same questions arise in teaching languages for specific purposes (L.S.P.):

- Do we teach the language ? or
- Do we teach the content-discipline ? or
- Do we teach both ?

No one would question the fundamental objective that teaching a foreign language, aims to give access to some sort of information, some sort of content, shall we say to 'meaning'. Therefore, 'semantics' is the other trait, which should be dealt with in any language learning as it covers 'meaning'. This latter is directly related to 'discourse analysis.' Discourse analysis focuses on 'the analysis of meaning' because *"the domain of discourse analysis is the description of the process whereby we create and relate, organise and realise, meaning."* (Riley,1985:1).

2.5 Discourse Analysis

Discourse analysis has attracted the attention of many scholars in the last decades.
Much work has been done about discourse in general, and discourse analysis in particular.

Among these scholars, Widdowson who states that

> *Discourse ... means a mode of social practice, in particular how institutions establish ideologies for the control of practice. (Widdowson, 1990:38)*

Fairclough (1992) in his social theory of discourse regards discourse as "*a mode of action, one form in which people may act upon the world and especially upon each other, as well as a mode of representation."* Both Widdowson and Fairclough stress the social role played by discourse in a society, because what is learnt depends in large part on the experience of learning within a particular social environment.

Gee (2005), would call a 'Discourse' with a capital 'D' because of the important role it plays in society. Gee (2005:20) states that people use different styles or varieties of language for different purposes, and that people also use different varieties of language to

enact and recognise different identities in different settings, and as such he would call each such variety a 'social language.' To illustrate this point, Gee gave the example of a student studying 'hornworms' who might say in everyday language, a variety of language often referred to as 'vernacular language', something like '*Hornworms sure vary a lot in how big they get'*, while the same student might use a more technical variety of language to say or write something like '*Hornworm growth exhibits a significant amount of variation.*' The vernacular version is one of social language and the technical version is another.

According to Gee "*people build identities and activities not just through language but by using language together with other 'stuff' that isn't language."*
He concludes by saying that he would reserve the word 'discourse' with a little 'd', to mean language-in-use or stretches of language (conversations or stories). 'Big D' Discourses are always language plus 'other stuff' (Gee, 2005:26), as is the case of the biology discourse which is formed of everyday words (language) and 'other stuff' to give new meaning in science.

Gee states that the term 'Discourse' (with a big 'D') is meant to cover important aspects of what others have called 'discourses' (Foucault, 1966); communities of practice (Lave and Wenger, 1991); cultural communities (Clark, 1996) and discourse communities (Miller, 1984).

2.5.1 Role of Discourse in Society

According to Gee (1990a)

> *Discourse is a socially accepted association among ways of using language, of thinking, feeling, believing, valuing, and of acting that can be used to identify oneself as a member of a socially meaningful group or 'social network' or to signal that one is playing a socially meaningful role.*
> (Gee, 1990a)

This quotation is meaningful because it includes two key terms related to this research 'discourse' and 'social networks' (already referred to in this chapter).
The use of 'network' may be an alternative concept for the time being, since 'group' (used in the quotation as well) may sound too much general, not adequate and its paradigmatic lexical equivalent 'community' is still to be investigated, and therefore it would be too early in this process, to make it perform a conceptual function.

As in this research, we are primarily concerned with teaching / learning a foreign language, and more precisely E.S.P, discourse is the element through which people (in this case biology students) interact and gather to form a 'social network.' Therefore, teaching / learning takes place within a specific social environment and as such, it presupposes socially constituted collaboration between the learner and others, i.e. the learner and the teacher, the learner and his peers, or the learner and his more capable peers (the well-established scientists).

In the same line of thought, Purvis and Hunt (1993) describe discourse as follows,

> *Discourse is constitutive of social relations in that all knowledge, all talk, all argument takes place within a discursive context through which experience comes to have, not only meaning for its participants, but shared and communicable meaning within social relations.* (Purvis and Hunt, 1993: 492)

It seems from this quotation, that social relations in any society are dependent on discourse, and that discourse is in turn dependent on social relations. They are inclusive of each other, and this also stands for the teaching / learning situation.

Related to this point is Vygotsky's social theory, which although originally applying to child learning and mental development, has also been explicitly addressed in discussion of 'Second Language Learning.'

Vygotsky theorised (1962, 1978) that there exists a 'Zone of Proximal Development' (ZPD) which he described as being *"... the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers."* (Vygostky, 1978:86, in Simpson, 2004: 63)

In our case, the implication for language learning is that collaboration, either among learners or between the learners and their teacher, is vital for the learner's development. Since discourse is dependent on social relations, and that these relations may be multiple and varied because of the different contexts of use, therefore the situations would presuppose the use of different types of discourse.

2.5.2 Types of Discourse

The definitions given above show the role of discourse in society in general , or within a community in particular. 'Discourse' is the focal element through which members of the society (as explained in 2.2.1) or members of the community interact.

Althusser (in Purvis et al, 1993) goes even further by stating that "*it is through discourse that individuals are interpellated as subjects."* This part will discuss the most important examples of discourse that exist in different communities. It will, then, focus on the written discourse as it is the main concern of this research.

In social theory, there seems to be three types of discourse classified as: *spoken, non-verbal and written.*.

Purvis and Hunt's (1993) account on discourse lies in its importance as a means of communication either through language and in this case the authors refer to either spoken language (speech systems) or written language discourse (texts), or through the medium of non-verbal sign-systems. This last type of discourse is represented by Purvis et al. in the example of "*practices in which males open doors for females.*"

Purvis et al. find that these practices are elements of a discourse whose organising framework is a strict sexual division of labour.

This statement may be true in the western world, but may need a different interpretation in a different cultural background. In the traditional and mostly rural society, the same attitude may be considered as 'rude' to the male category.
Rituals and customary manners would deserve a side attention in this reflection, since the classroom context often requires non-verbal systems of signs that could complete or improve a didactic purpose. It has given birth to various theories on role-playing in pedagogical situations, and tends to be more and more considered as a pedagogical means.

Ignorance of 'non-verbal systems' in pedagogical situations, may also lead to counter didactic effects. As an example, when a boy hands up first to answer a question, and is then followed by a girl, and the teacher with a smile says, "Ladies first", this order of priority may be counter producing. The teacher is therefore in the situation whereby, he does not know which system should be implemented in the pedagogical setting, which remains neutral, neither cultural i.e. Algerian, nor foreign, i.e. English. This frame of relations is new for both the learner and the teacher, and often a discoursal choice has to be made before the choice of a language. This is where 'interlanguage spaces' take place appealing 'intercultural' initiatives to be negotiated by the teacher and the form, in the classroom context.

2.5.2.1 Spoken and non-verbal discourse

Gee (1990) also refers to spoken and non-verbal types of discourse when he says

> *Discourses are ways of being in the world, or forms of life which integrate words, acts, values, beliefs, attitudes, social identities, as well as gestures, glances, body positions, and clothes.*
> (Gee, in Clarke, 1990:142).

He goes on to argue that "*such discourses can only be acquired through apprenticeship and observation and cannot be learnt in formal settings.*"

In his quotation, Gee is particularly referring to the spoken and non-verbal discourse which individuals acquire by observing, that is, paying attention to people of the same community saying it (spoken discourse) or doing it (non-verbal discourse).

This in a way calls people to conform to the rules of the community they are living in. The hidden idea, in Gee's quotation, of *"people having to conform to the rules of the community"* is clearly reflected by Hall (in Purvis et al, 1993:485) when he defines *" discourse as sets of already -made and preconstituted 'experiencings' displayed and arranged through language*", in the sense that discourses impose frameworks which limit and influence what can be said (spoken discourse) and what can be done (non-verbal discourse). That is the reason why some specialists state that discourse cannot be learnt in a formal setting as this latter could distort the real meaning; because it is not said or done naturally, it will lead to a sort of artificiality of both the spoken and non-verbal discourse. This is true especially in the classroom situation, where the teaching of some components is done in an artificial setting, e.g. some role plays in conversation module may be inadequate to students when not geared towards a kind of near to reality situations. Such practice may not only lead to artificiality, but would also restrict thinking and creativity.

2.5.2.2 The written discourse: Text and Discourse

'Written discourse' is another type of discourse, and it is the focal element of the present research. The term 'discourse' is reserved for the "*coherent, dynamic, communicative function of a text: for text as a communicative entity."* (Crombie, 1985:56). This difference called for by Crombie allows one to emphasise the fact that the focus of the current reflexion is related to the study of the communicative dynamics of discourse; that is to make students aware of the way in which various elements of discourse function in relation to one another in order to communicate meaning. Crombie adds,

> *When we look at text as discourse, we look at its communicative dynamics: at the way in which its various elements function in relation to one another to communicate patterns of integrated meaning.* (Crombie, 1985: 56)

This quotation is significant because of its relevance to the purpose of the present study, as our learners are interested in looking at the communicative dynamics of the text as discourse, this being the main function of documentation and the various literature reviews that support their training.

2.5.3 Function of Discourse

In 'discourse' (whether spoken or written) it is meaning that should play an important role and that should be focused on. Meaning does not exist as such in a text but it is negotiated. For Widdowson (1980),

> *Communication is called for when the language user recognises a situation which requires the conveyance of information to establish a convergence of knowledge, so that this situation can be changed in some way. This transaction requires the negotiation of meaning through interaction. I refer to this negotiation as discourse.* (Widdowson, in Riley, 1985: 7)

In the present research, we are primarily concerned by teaching students the way in which various elements of discourse function in relation to one another in order to communicate meaning, and our capacity to transmit onto them procedural devices to help them recognise useful meaning, and negotiate access to it.

Therefore, what is more essential to us is to know how students interact with discourse in order to negotiate meaning. In this case , how do biology students interact with scientific literature to negotiate meaning ? as *"the function of discourse is seen as the creation of a common meaning through the participants' sharing and comparing information"* of scientific discourse essentially important for the discourse community under study, and for the target discourse community

2.6 Scientific and Technical Discourse

The term ‘discourse’ has been so far defined. This part deals with scientific discourse by first providing definitions of ‘scientific discourse’ and then present essential features that form scientific discourse.

If one takes any scientific literature, the first striking feature is the division of discourse into verbal and non-verbal organisation. However, each discipline is written in a more or less particular way. In ‘Economics texts’, more use is made of tables and graphs; in ‘Chemistry’, formulae and equations are very much used; and in ‘Engineering (gas treatment)’ mainly formulae, equations, tables and different kinds of graphs are used (Belkhenchir, 1986: 46-7).

Therefore, Riley (1985:107) defines scientific discourse *“as the verbal and non-verbal realisation of the communicative system of science”;* thus scientific discourse is seen as being composed of verbal realisation i.e. ‘text’ (as defined by Crombie, 1985:56), and non-verbal realisation as graphs, diagrams, charts, formulae, equations, etc. to communicate and diffuse scientific knowledge.

There is another aspect of non-verbal system in scientific discourse which falls within the same field. It is the entire experimental process that is transcribed into written discourse and borrows description and expository modes of composition. As such, the non-verbal covers up visual, hearing, smelling, touching and tasting senses which are bound to render a physical / tangible phenomenon into its textual transcription. There is a lot more to investigate in this direction, since it concerns the shift from empirical knowledge to theoretical knowledge i.e. from the world of physical and natural manifestations to the world of abstraction, and to a certain extent intellectual speculation.

There are reasons to worry about the issues at stake, since research is a way to approach ‘Truth’, and our ability to unveil this truth varies from a methodological point of view whether we use language to do it, or the elements of the concrete world to reach such a target.

The realisation of "*communicative system of science*" (Riley), follows certain norms agreed upon by scientists. As Widdowson expresses it,

> *Scientific exposition is structured according to certain patterns of rhetorical organization which, with some tolerance for individual stylistic variation, imposes a conformity on members of the scientific community no matter what language they happen to use.* (Widdowson, 1976: 61)

Myers (1994) in an earlier study (1990a:142) compared scientific articles published in New Scientist (the popularising articles) and articles published in Scientific American (professional articles) and found that:

- *The professional articles* create what he calls "*a narrative of science*" i.e. they follow the argument of the scientist, arrange time into a parallel series of simultaneous events all supporting their claim, and emphasise in their syntax and vocabulary the conceptual structure of the discipline.

- *The popularising articles*, on the other hand, "*present a sequential narrative of nature, in which the plant or animal, not the scientific activity, is the subject, the narrative is chronological, and the syntax and vocabulary emphasize the externality of nature of scientific practices."*

The importance of Myers' findings lies in the fact that it highlights the main point which the present research will focus on. That is, to find out how scientific discourse of biology is structured and in which of these two constructing views of science it is represented. This can only be done through a study of the 'genres' (Swales's fourth criterion in Section 2.3.3.3) used by the community under study.

2.7 Genre(s): Definition and role in language teaching / learning

The term 'genre' has been given different definitions as 'genre' may concern a 'speech event' or 'a particular written piece of work.'

Richards et al. (1985) defines 'genre' as a " *particular class of speech event which has certain features common to that particular event."*

Similarly, Swales defines 'genre' as "*a class of communicative events, the members of which share some set of communicative purposes.*" (Swales, 1990: 45).

Therefore, "*communicative purpose*" is in fact "*the defining feature by which a genre such as the academic article is distinguished from other genres... and by which the consideration of genre is distinguished from the consideration of register."* (Dudley-Evans, 1994: 219).

In the present research, a genre can be defined as a communicative event through which:

- Either the learners, of the discourse community under study, have to negotiate its meaning to reach a specific objective (i.e. these novice students of biology have to negotiate the meaning of scientific texts in order to be able to write their 'mémoire' and then join the target discourse community); or

- experts (scientists in the Target Discourse Community) writing these articles to communicate new ideas, new findings, to their colleagues. Hence, Swales saying "*members of individual professions, or discourse communities will share common purposes of communication, or genres."* (Swales, 1990).

2.8 Genre Analysis in E.S.P

The application of genre analysis to language teaching, and especially to teaching English (i.e. E.S.P.), has grown in importance recently. Because English is worldwide used, it is gaining more ground in different disciplines (as shown in Swales' simplified version (1990: 97) of Baldauf and Jernudd's findings of the advance made by English 1965-81):

	1965	*1981*	*GAIN*
Chemistry	50	67	17%
Biology	75	86	11%
Physics	73	85	12%
Medicine	51	73	22%
Maths	55	69	14%

Table 2.1: Advance made by English in different disciplines from 1965 to 1981

Table 2.1 shows that the English language gained more ground in different disciplines from 1965 to 1981 (and is still gaining more ground as most of the scientific literature nowadays is mostly written in English). This gain is represented in percentages. The percentage of gain in biology is the lowest (11%) compared to medicine (22%). This is quite representative as it stresses the fact that biology is the discipline which has almost always been written in English.

English is also the language of most "*invisible colleges*" (Crane, 1972), and the language used in international conferences, organisations, etc.
This 'Discourse Community Pressure' (D.C.P. Swales, 1990), that is non-native speakers of English having to use English, is the assumption that more and more scientists whose mother tongue or whose language of wider communication is not English, need to be competent in this latter as "*Effective communication and dissemination of scientific information is crucial."* (Ebel *et al.*, 1987: 57).

We have to be careful here and make it clear that it is true that we observe an expanding spread of the use of English at international meetings, yet make a proviso for varieties of English to be more faithful to the field of practice, the 'scientific English' itself shifting from terminology, jargon to discourse of the scientific community.

At present, most people who need to learn English, they usually need it for academic or professional purposes (E.S.P), hence the need to use some specific genres of the English language, for one aim or another. However, learners of E.S.P. do meet some impediments during the learning process, either in writing or reading the target language. These impediments could be due to several factors: the teaching approach, or the methodology, etc. Therefore, English language programmes have to become more explanatory by trying to make the learners aware of the rationale of the text-genre that s/he is required to read or write through 'genre analysis'.

'Genre analysis' is the study of how language is used within a particular context. Its particular focus is on the distinguishing features of different texts (be they spoken or written) in all areas of E.S.P.

2.8.1 Genre-types and their Structures

In different literature of different disciplines, 'genres' can be classified as spoken or written and they *"vary in terms of the mode or the medium through which they are expressed: speech or writing."* (Swales, 1990: 62).

In this section, I shall firstly, give some examples of spoken and written genres, and secondly list some structures of written genres only, as the focus of the research is based on one written key genre (journal articles).

a) **Spoken genres** range over a spectrum including**:** board meetings, the business negotiation, slide presentations, lectures, and tutorials, to name but a few.
b) **Written genres** include: seminar papers, research articles, reports describing experiments, and corpus material.

The structures of each type of genre (be it spoken or written) show clearly that *"genres are communicative vehicles for the achievement of goals."* (Swales, 1990:46)

Mitchell (1957:75) specified elements in the structure of shop transactions in Libya (in Flowerdew, 1993:307). The '*shop transaction'* is a 'spoken genre' used to achieve one's goal of buying or selling goods, and is therefore considered as a '*transactional'* form of communication to differentiate it from '*interpersonal*' communication. (Chomsky).

The structure of different genres has been identified by several scholars, and it is well represented in some academic works. Swales shows how genre analysis can be used effectively to describe the types of discourse found in academic settings.
Swales's pioneering work (1981) on 'Aspects of Article Introductions' (the four move structure) and 'Genre Analysis' (1990) (the three move structure) showed how research article introductions are structured. He considered that the way the writers structured and commented upon references to previous research (Discussion of Previous Research 'D.P.R' / Move 2) was an aspect of academic writing.
Swales' work led to parallel research into other sections of the research article such as: Results, Discussion of Results and Abstracts.

Dudley-Evans (1994 and 1998) presented an approach to the analysis of the 'Discussion Section' of scientific research that draws its inspiration from Swales' original work.
Bhatia (1991) proposes seven moves for the genre of sales letter.
Davies and Green (1984) have identified a four part structure for "physical structure" texts.
In the present study, the genre used by biology students as a communicative vehicle is 'scientific journal articles' (information collected from the student questionnaire answers/ 'Chapter Three') and the *"achievement of goals"* is to use the information they extract from these articles to write up their 'mémoire' (end-of-degree project).
The following part will refer briefly to one written key text-genre 'journal articles' read by biology students aiming to use English in academic settings, and the role of the genre within the discourse community that regularly uses it.

2.8.2 The Written genre-type in this research (i.e. scientific journal articles)

'Scientific journals' play an important role in the development of a society. According to Ebel *et al*., scientific journals "*exist primarily to serve the scientific community"* but "*they*

also have important function: providing individuals engaged in science with a vehicle for announcing and disseminating their research results." (Ebel *et al.*, 1987: 59).

In this research, journal articles, as a written genre, are chosen for study for two main reasons.

1- Biology students (whether in the francophone or arabophone departments) read only articles from different biology journals to get scientific information related to their subject-speciality in order to write their 'mémoire' (as the questionnaire answers tend to confirm statistically); and also because

2- The written communication is indispensable in science as "*the written word is the dominant medium for lasting communication and careful study."*(Ebel *et al.* 1987:5)

It has another fundamental function for the members of this community, that of citation. The written information (text) remains the only tangible point to back up the construction of their reasoning. References to other researchers' work become a guarantee of truth, established by methods acknowledged by the community, since they are published and institutionalised, i.e. 'announcing' and 'disseminating' results. This information is confirmed in the 'computational analysis' (Chapter Five) where the first and most frequent non-grammar word in the biology corpus is *'et al.'* with 317 occurrences .

This is how the scientific community accumulates production of new ideas, and use them to proceed further to improve and / or develop knowledge.

2.9 Survey of research methods in E.S.P.

The development of E.S.P. (from the 1960s to date) has been marked by the influence of different types of analyses.

2.9.1 Register analysis

Register analysis which "*operated on the basic principle that the English of, say, Electrical Engineering constituted a specific register different from that of, say, biology or of General English"*, (Hutchinson and Waters, 1987:9) took place in the 1960s and 1970s. It had

focused on sentence grammar and was associated with the work of Peter Strevens (Halliday, McIntosh and Strevens, 1964), Jack Ewer (Ewer and Latorre, 1969) and John Swales (1971).
In the first stage of its development, E.S.P had also focused on language at sentence level as the work of 'Barber (1962), Herbert (1965), etc.' testifies.
The aim of the analysis was to identify the grammatical and lexical features of these registers.
Useful research in this field also tends to confine itself to very limited, clearly defined, areas for example: prepositions in chemical abstracts, Latinate names in biology.

2.9.2 Discourse analysis

Whereas in the first stage of its development, E.S.P had focused at the sentence level, the second phase of development shifted attention to the level above the sentence (from register to discourse analysis) to *"understand how sentences were combined in discourse to produce meaning."* (Hutchinson *et al.*, 1987:11).

The leading lights in this movement were Henry Widdowson in Britain and the so-called Washington School of Larry Selinker, Louis Trimble, John Lackstrom and Mary Todd – Trimble in the United States of America.

'Register Analysis' was seen more structuralist and 'Discourse Analysis' more communicative because it was first concerned by the study of spoken discourse, e.g. Austin, Searle and the 'Speech Act Philosophy.' Nowadays, 'discourse analysis' refers to "*the analysis of spoken and written stretch of language using two different methodologies: conversation analysis and discourse analysis."* (Seedhouse, 1997). The analysis covers aspects of sentence connection, or cohesion. Therefore, and according to Brown and Yule (1985:26) doing '*Discourse Analysis'* certainly involves '*Doing syntax and semantics'*, but it primarily consists of '*doing pragmatics'*. In 'Discourse analysis' as in pragmatics, we are concerned with "*what people using language are doing"*, and accounting for the linguistic features in the discourse as the means employed in what they are doing.

'Discourse analysis' includes 'Genre Analysis' and most recently 'Corpus Analysis'.
Genre analysis adopts a discourse analysis approach.

2.9.3 Genre Analysis

According to Hyland (1992: 15) "*genre analysis is the study of how language is used within a particular context.*" This means that each genre is structured according to the goal it wants to achieve, for example, the way to write the 'Introduction' of a scientific research article may be different from that of writing the 'Introduction' of a ' Magister / Master Thesis'.

In teaching 'Genre Analysis', we try to show how texts are structured i.e. '*how writers conventionally sequence material to achieve particular purposes'*, Hyland (1992:15), in order to offer positive help to our students, in terms of organising their written assignments.

"*A genre approach offers students essential information on language in use (*as is shown in the following most known E.S.P genre samples), *not a means of packaging ideas."* (Hyland, 1992:17). In E.S.P, the pioneering work dealing with 'genre analysis' is Swales' (1981) 'Aspects of Article Introductions.'

A discourse analysis approach is used in Swales' study of the introduction section of forty-eight examples of the genre 'academic journal article', drawn from a wide range of academic disciplines (Physics / Electronics / Chemical Engineering / Medical field / Biology and Social Sciences) where he suggested a four-move structure.

His most recent model 'Genre of Research Article Introductions, 1990) suggests a three-move sequence structure: Establishing a territory / Establishing a niche / Occupying the niche.

Extending Swales' approach, Dudley-Evans (1994) has established an 11 move structure in the 'Discussion Sections' of research articles (as explained in Section 2.8.1.2 of this research).

2.9.4 Corpus Analysis

Whereas 'linguistics' is the description of aspects of language use (e.g. syntax / semantics, etc.), 'computational linguistics', which saw the light in the 1950s, is associated with the use of computers for machine translation between languages, and it includes scholars from computer science, artificial intelligence and linguistic theory. The methodology used in this field is 'corpus linguistics'.

The development of modern technology has also reached the field of E.S.P, where we can find a considerable amount of research using 'corpus linguistics'.

'Corpus linguistics' deals with corpus analysis, i.e. the selection of a corpus (following certain criteria) for language study. The criteria for corpus design are:

- Representativeness,
- Sampling,
- Diversity,
- Size,
- Copyright permission (for written texts), and transcription of spoken data (Biber *et al.*, 1998: 246).

The main reasons for using corpus analysis could be summarised in this research as follows:

a) Most of these studies have found that there exist differences between what textbooks are teaching and how native speakers actually use language as evidenced in the corpora; (McEnery and Wilson, 1996: 104);
b) ... a number of corpora have been collected, and researchers are using them to work toward empirically based descriptions of varieties of languages (Chapelle, 2001: 37)
c) The linguistic facts about language use can be helpful for developing dictionaries, syllabi and teaching materials, particularly in English for specific curricula. (Chapelle, 2001: 37)

As the aim of this research is to develop an appropriate teaching methodology and create some teaching materials for the community under study, the conceptual framework of this research will deal with discourse and genre analyses followed by a computational analysis. The results from these analyses will form the foundation stone for the suggested methodology.

2.10 Research methods

The review of the literature of the two concepts namely 'discourse community' and 'discourse analysis' given in the previous sections of this chapter leads us to the determination of the kinds of research methods. The latter will be of two kinds:

2.10.1 Questionnaires to assess the teaching / learning situation

Two questionnaires were distributed to teachers and students in order to find out how the teaching of the English course was done.

The teacher questionnaire deals with questions related to the course content, the methodology used, etc.), and the student questionnaire tries to assess the students' real needs for learning English, and what problems they might have met in learning English.

2.10.2 Discourse analysis

A detailed description of discourse / discourse analysis and genre / genre analysis has been given in the previous sections of this chapter.

The investigation of how the frame of reference is represented in biology discourse ('How does language function in extremely specialised discourse?') will be done using two methods: one manual and the other computational.

2.10.2.1 Manual analysis

In 'Chapter Four' the specification of the corpus of texts used for analysis will be given. This includes: Rationale for the corpus, limitations of the corpus, types and sources of data and the corpus text characteristics.

The manual analysis shows the grammatical and lexical features used in this corpus.

a) The grammar used in this corpus: A frequency count of tenses, voice (active / passive) and modals, is done and the results are presented in table forms.

b) Lexis: Because the lexis of biology discourse consists of everyday words in the form of nominal compounds 'to give new meaning in science', we therefore focused on analysing lexis represented in the form of nominal compounds. The analysis deals with two aspects of nominal compounds: forms and functions.

   (i) Forms of nominal compounds (based on Williams'categories, 1984) include NC specification and NC frequency .

   (ii) Functions of NCs in scientific discourse (based on Bartolic's categorisation, 1978).

2.10.2.2 Computational analysis

Because lexis in biology discourse is formed of everyday words in the form of NCs that tend to 'chunk' and 'collocate' to give new meaning, a computational analysis is seen necessary for this research as it will unveil all the language elements in the near environment of the nominal compounds under study.

A specific computer programme called ATA (Aston Text Analyser) will be used to do this analysis in 'Chapter Five'. The analysis will be based on 'concordancing' to find out how concepts are realised in scientific discourse and which grammatical structures these concepts combine with (that is 'Research Question 2'). As Stubbs put it

> *Meanings are conveyed directly, by the choice of particular words, but they are also conveyed indirectly by patterns of co-occurrence: which words collocate, and which words occur in which grammatical constructions.* (1996: 97)

Therefore, the aim from the use of a corpus, together with the analysis of collocational frameworks, is to raise consciousness of both teachers and students of the different ways we might develop to present and explain language patterning, focusing on one specific discipline

2.10 Conclusion

This chapter dealt with the literature review of 'discourse community' and 'discourse analysis', two key concepts used throughout this research.

Firstly, it gives the definition of the notion of 'Discourse Community' and its different characteristics, represented in both Swales's six criteria and the management science elements 'CATWOE'. Each element is encapsulated in the other, forming a chain, if respected, would hopefully lead to a betterment of the situation.

Then, we refer to the notion of 'Target discourse Community' that the students will join after successfully completing the 'Transformation process' ('T' one main element of 'CATWOE') i.e. the English language course, which would help undergraduate biology students join the 'Target Discourse Community' and become active members of the society, because the first objective of any university is to serve the society.

Secondly, and in relation to the notion of 'discourse community', this chapter dealt with 'discourse analysis' as 'discourse' is the element through which members of the discourse community under study interact. Research on discourse and genre analyses is reviewed, and the implications of this for the present research are discussed; then, reference to the kind and genre of discourse used by the undergraduate biology students is made.

The chapter ends by suggesting the research methods (questionnaires, manual and computational analyses) that will be used in 'Chapters Three, Four and Five.'

CHAPTER THREE

THE TEACHING AND LEARNING CONTEXT

CHAPTER THREE

THE TEACHING AND LEARNING CONTEXT

3.1 Introduction

In 'Chapter Two' the discourse community of biologists was described through the twelve criteria (Swales's criteria and the six management science elements C.A.T.W.O.E) focusing mainly on the 'Transformation Process' (that is the English course) which helps the learners to move from a neophyte to a member of the Target Discourse Community. This chapter deals with the first research question, i.e.: 'How is the discourse community under study organised in terms of teaching / learning practices ?'

In scientific education, it is agreed by the institutions (the Ministry of Higher Education, the university and the university departments) that the teaching/learning of English is a compulsory component in the general curriculum. As stated in 'Chapter Two / Section 2.2', language is not studied in isolation but within a social group, and as such would therefore need to be analysed within that social group. Such analysis would normally have recourse to ethnographic data collection procedures represented in the use of the triangulation: questionnaires, interviews and classroom observations. However, in this research we managed to get answers to questionnaires only, and to some 'informal' interviews. The research population objected to formal (recorded) interviews and classroom observations.

In 'Chapter One / Statement of the problem', we have observed that in science institutes, English is most of the time taught by language teachers, who as non-specialists would focus more on grammar (at the expense of lexis); and if English is taught by a subject-specialist, this latter would focus more on lexis and scientific experiments (at the expense of grammar).

To check how far this is true, two questionnaires were designed: one for teachers and the other for students. Therefore, the following part will deal with the analysis of the questionnaires that were answered by both teachers, and the students.

3.2 The Teacher Questionnaire (see Appendix A)

The questionnaire addressed teachers (as 'Actors', Chapter Two/Section 2.3.3.3) of the English language course at the department of biology. It was designed to find out the English language course objectives and organisation, and the methodology used by both language teachers and subject-specialists teaching English (i.e. Transformation Process). The questionnaire originally included questions about the teaching of the four skills. However, as the number of teaching hours in the biology department varies between one and a half hours in the third year and one hour in the fourth year, so the skills actually taught are 'reading' and, to some extent, 'writing.' Also, more emphasis is put on teaching the reading skill because of the ministerial decision that states that the objective of teaching English in science institutes is "to learn how to read technical literature written in English"[11] (Ministry of Higher Education, March 1992)

For these reasons, we shall focus on the 'reading' and 'writing' skills in the analysis of the questionnaire data.

3.2.1 The format

The questionnaire was divided into two parts:

- **PART A** deals with questions related to the teacher's background and the English course objectives.

- **PART B** deals with questions related to the methodology used by language teachers and subject-specialists to teach E.S.P.

[11] « *Apprendre à lire et comprendre une documentation technique écrite en anglais. » (Ministère de l' Enseignement Supérieur, Mars 1992)*

3.2.2 Analysis of the questionnaire results

3.2.2.1 The teachers' background and the English course objectives

(i) The teachers' background:

The questionnaire was answered by two teachers teaching English at the department of biology:

- One language teacher (B.A. in the English language);
- One subject-specialist (a microbiologist).

(ii) The English course objectives:

The language teacher states that the aim and objective of the English course is to show how the English language works, whereas, the subject-specialist states that the objective is to improve scientific knowledge (he gives lectures and workshops about the subject-speciality in English)

The language teacher states that the students' main problems in E.S.P. are:

- grammar (tenses, concord, connectors, etc.), and
- general vocabulary.

The subject-specialist teaching E.S.P. finds that the students have difficulty mainly with

- scientific terminology, and
- sometimes with grammar and general vocabulary.

3.2.2.2 The teachers' methodology

This part reports on the methodology used by the language-teacher and the subject-specialist for teaching English. The questionnaire answers are given in a table form to compare the methodology of both teachers of English. For reasons given in 'Section 3.2', the analysis will focus on the teaching of 'the reading component' and 'the writing component.'

Teaching E.S.P at the Department of Biology

The reading component	The language-teacher	The subject-specialist
1) Types of texts	- Suitably mixed	- too specialised - suitably mixed
2) The approach used to teach reading	**The communicative approach** - Looking for linking words to help students plot their way; - Devising questions to encourage students to infer meaning; - Introducing strategies to identify main points of a paragraph/complete text. - There is no explicit grammar teaching.	**The traditional approach** - the teacher reads the text aloud; - The students read the text silently and then answer questions related to terminology (orally or in writing)
3)Reading exercises	- Reordering (of sentences/ paragraphs); - Answering inference questions; - Completion exercises; - Transfer of written information to non-verbal form (tables, graphs, etc.)	- Answering 'literal' questions related to the text; - Making sentences; - Explanation of difficult words.
4) The exercises are done	- Individually; or - In pair / group work	- They are done individually only.
5) Evaluation	- Transfer of written information to non-verbal forms (charts, tables, graphs, etc.); - Multiple choice questions; - True / false questions ; - Cloze test; - Jumbled sentences / paragraphs;	- Questions related to terminology used in the text; - Definition of words (synonyms, antonyms, etc.); - Making sentences with specific words;
6) Assessment	- 'Correct' answering of questions.	- Scientific knowledge.

Table 3.1: The methodology used in teaching the 'Reading component'

Teaching E.S.P at the Department of Biology

The writing component	The language-teacher	The subject-specialist
1) Sub-skills taught	- Information transfer (from tables, graphs, etc.) to text ; - Writing summaries of published scientific articles and paragraphs; - Writing essays; - Writing reports.	- Writing summaries and paragraphs; - writing essays; - Writing reports; - Information transfer.
2)The approach used	**The communicative approach** - Writing is taught through texts and units of thought (Isolated sentences are not good examples for teaching writing) - Study a model text; practise various types of sentence writing based on the model (introductory sentences; sentences containing new ideas, development of the ideas, results of those ideas, inferences, conclusions).	**The traditional approach** - Explain the main principles of writing (how to start, how to develop ideas, how to conclude) then ask the students to write an essay; (According to this teacher 'students must experience writing paragraphs through doing').
3) Evaluation	- Continuous assessment through exercises and writing tasks that accompany every writing unit (to check students' progress); - The final writing report (an individual work related to the student's speciality)	- Students are asked to write essays and reports (no continuous assessment; just the examination at the end of semester / year)
4) Assessment	- To decide on the relative value of content and grammar.	- To assess scientific knowledge.

Table 3.2: The methodology used for teaching the ' Writing component'

3.2.3 Conclusions from the questionnaire results

a) ESP course aims:

- The language teacher teaches how the language works and cares a lot for the students' progress through the course;
- The subject-specialist teaches the subject-matter and cares a lot for scientific knowledge.

Not at any time does the subject-specialist mention anything about the course gradation or the students' progress.

b) The approach:

- The language teacher teaches English at discourse level (through the communicative approach);
- The subject-specialist uses the traditional method and teaches English at sentence level (and sometimes even at word level).

c) The assessment:

- The language teacher assesses his students through the whole teaching period. It is a continuous assessment to check students' progress. What is assessed is language (the content as well as grammatical / lexical and cohesive correctness.)

- The subject-specialist assesses the students once a semester. What is assessed is scientific knowledge.

d) The materials:

- The language teacher uses published and in-house materials.
- The subject-specialist uses published materials, and sometimes produces his own materials.

3.3 The student questionnaire (see Appendix B)

3.3.1 The format:

The questionnaire was divided into three parts: A, B and C.

- **Part A** serves to gather general information on the students' background namely, age, gender, the department they belong to (Arabophone / Francophone), the number of years of study of English, and the English level proficiency.

- **Part B** deals with questions related to the general objectives of the English course at the Department of biology.

- **Part C** deals with questions related directly to the English course (organisation, teaching methods, etc.)

3.3.2 The questionnaire methodology

The number of students surveyed was 100, all in their fourth (and final) year, studying either 'Biochemistry' (in Arabic) or 'Genetics' (in French) at the department of biology (University of Oran) in the academic year 1994-95.

3.3.2.1 Aims of the questionnaire

The questionnaire was set up with the following aims :

(i) To identify the students' real needs (not those that have been 'assumed' so far) and the problems faced by students in learning English in order to develop a teaching methodology accordingly. As Potts put it:

> *It is surely the whole point of the ambitions of an E.S.P. programme that its syllabus and methodology objectives, should be as closely geared as possible to the needs, both potential and ongoing, of the students we teach.*
> (Potts, 1988)

(ii) To assist the institute of biology with its English course review;

(iii) The questionnaire was also a way to tell students that as teachers we are concerned in them learning the language and in the way they are learning it. Because as Chalon [12] questioned and answered,

To the question " Teachers, what for ?" We would be tempted to answer, " Teachers, for listening."

(Chalon, as quoted by Riley, 1985)

3.3.2.2 Organisation of the Questionnaire

The questionnaire was composed of nine rubrics with different kinds of questions:

- Yes/ no questions;
- Multiple choice;
- Factual questions;
- Ranking scale questions (where students are asked to attribute a number - one to five- according to the relative importance given to each item) ;

Open-ended questions were avoided because they might be ambiguous, and time-consuming for the students.

The questionnaire was written in English but the students were free to use a language of their choice (Arabic , French or English) to answer the questionnaire.

3.3.3 Results and Discussion of the Questionnaire

3.3.3.1 Rubric 1: The Students' Background

The following information was obtained when the questionnaires were analysed.

1- Distribution of the students by sex

More female students participated in answering the questionnaire.

[12] A la question « Des professeurs, pourquoi faire ? » Nous serions tentés de répondre, « Des professeurs pour écouter . » (Chalon, as quoted by Riley, 1985)

Level	Respondents	Respondents	Total
	Male	Female	
Fourth year	40	60	100

Table (i): Distribution of respondents according to sex

2- Distribution of the students by age

The age of all the students who answered the questionnaire varied between 20 and 25.

3- Medium of instruction of the students

The questionnaire was given to 100 students in 'the Arabophone Department' and 50 students in 'the Francophone Department.' Seventy eight (78) students who answered the questionnaire study biology in Arabic, and twenty two (22) students study biology in French (a total of 100 students).

4- The students' foreign language studied at the middle and secondary schools

Eighty six (86) students studied English, two (02) students studied Spanish and another two (02) students studied German. Ten (10) students did not answer.

Foreign language studied	Number of students
English	86
Spanish	02
German	02
No answer	10
	TOTAL ***:*** ***100***

Table (ii): Foreign languages studied

5- Number of years of English instruction

Eighty six (86) students have been studying English for six years, and four (04) students for one year only (i.e. students who studied another foreign language at school).

6- Self-evaluation of English proficiency

Eighty six (86) students (those who have studied English for six years) stated that their level in English is intermediate. Four (04) students stated that their level is beginner (those who studied Spanish or German previously).

Questions 5 and 6 were asked to find out the level of students in order to develop a teaching methodology accordingly.

Number of students	**Number of years of English instruction**	**Level**
86	6 years	Intermediate
04	1 year	Beginner
10	No answer	No answer

Table (iii): Students' self-evaluation

3.3.3.2 Rubric 2: General Instructional Objectives of the English Course

1 - The year of study of English

English is introduced in the third year study at the Department of biology. It is taught for three semesters: two semesters in the third year, and one semester in the fourth year.

Year of study	English / Semester
First year	No English
Second year	No English
Third year	Semester 1 and semester 2
Fourth year	Semester 1 only

Table (iv): Year of study of English

From the table above, it can be noticed that a void of two years in English study exists between secondary and university levels.

2- Types of English classes

All the students agreed that the type of English they had in the third year university level was scientific (as opposed to literary English).

3- Necessity to study English

All students agreed that it is necessary for them to study English for the following reasons:

- To understand scientific literature written in English in their field of study;
- For further research as all the scientific literature in biology is published in English;
- It is the language of wider communication;
- It is the language of science.

4-The main objective of the course

All the students agreed that reading is the most important skill. The importance of skills was ranked as follows by the students:

1- Reading

2- Speaking

3-Writing

4- Listening

5- Number of teaching hours of English per week

The students are taught English for one and a half hours a week (one session) in the third year, and one hour a week in the first semester of the fourth year.

6- Was there a suitable number of English lectures on the course?

To this question, the students answered that there were not enough English lectures because they find:

- Their level too low to improve in such a small number of English sessions;
- They could not study all the scientific terminology necessary for their speciality;
- They could not get all the information necessary for their study, as according to them, their teacher's main aim was to finish the programme.

3.3.3.3 Rubric 3: The English course organisation

1- Was the course well organised (content , duration, activities, etc.)

The students gave the following answers to show how well the course was organised:

- Content	54 positive answers
- Activities	48 positive answers
- Tasks	06 positive answers
- Duration	06 positive answers

(some students ticked more than one answer, and some did not answer this question).

2- What did the English course consist of ? here, again the students ticked more than one answer to show what the course focused on:

- Teaching subject speciality	78 answers
- Reading comprehension	65 answers
- Writing exercises	24 answers
- vocabulary exercises	21 answers
- Grammar exercises	21 answers

3- Students' ability to understand their English teacher's lectures

- Partly 65 students
- Completely 24 students
- Not at all 03 students

Rubric 4: English Assignments

1- Types of English assignments that the students had to do.

The students stated that their English assignments consisted of:

Types of English assignments	**Number of answers**
Explaining scientific terminology	75
Reading texts and answering comprehension questions	39
Doing grammar exercises	27
Reporting an experiment	12
Writing an essay	04

Table (v): Types of English assignments

2- Degree of difficulty of the assignments

Sixty five (65) students found the assignments easy, and fifteen (15) students found them difficult. Twenty (20) students did not answer.

3- How often were students assigned work out of class

Fifty five (55) students stated that they were sometimes assigned work out of class, fifteen (15) said that they were rarely assigned work out of class, and thirty (30) stated that they were never assigned work out of class

Number of students	Frequency of work out of class
55	sometimes
15	rarely
30	never

Table (vi): Frequency of assignments to do out of class

Rubric 5: The Different Skills Taught at the Department of Biology

In this part of the questionnaire, I tried to find out what skills were taught and what problems the students faced with each skill.

A-The Reading Skill

1- Frequency of reading English

Number of students	Reading frequency
55	Sometimes
30	Never
15	No answer

Table (vii): Frequency of reading English

2- Is reading in English recommended in your field of study ?

Eighty (80) students find that reading in English is recommended in their field of study. Twenty (20) students did not answer.

3- Type of literature they have to read in English

Type of literature	Number of answers
Articles in scientific journals	84
Books and textbooks	48
Song texts	01

Table (viii): Types of reading literature

4-Difficulties the students meet when reading English literature (the students ticked more than one answer)

Areas of difficulty	Number of answers
general English	45
scientific English	36
grammar structures	24

Table (ix): Areas of difficulty in reading English

5- Which strategies do you use when reading ?

The students gave the following answers:

- Translate word for word using a bilingual dictionary;
- Scan the texts (look for specific information only);
- Send the whole article to a French laboratory abroad to get it translated and sent it back.

B- The Writing Skill

1- Frequency of writing English at the Department of biology

Seventy (70) students stated that they sometimes write in English, twenty four (24) students stated that they never write in English, and six (06) gave no answer. The use of 'never' remains ambiguous because it is not clear whether these students were never asked to write in English, or whether they were not able to write in English.

2- Kind of writing the students had to do

Forty two (42) students reported that they had to write an essay. By 'essay' they meant a short paragraph, not a project.

3- Difficulties the students meet when writing in English

The students find difficulties in writing at:

- word level (choice of lexis) 66 answers
- sentence level 51 answers
- paragraph level 48 answers

C- The Speaking skill

1- Necessity of speaking English at the Department of biology.

The results of this question indicate that twenty (20) students find it necessary to speak English and seventy two (72) students find it unnecessary.

2- Do you need to speak English ? and Why ?

Twenty (20) students find that they need to speak English in order to:

- Discuss scientific matters with foreigners as it is the language of international communication;
- To go abroad to study;
- To greet people ? (*my question mark*)

3- Frequency of speaking English

Frequency of speaking English	**Number of students**
sometimes	70
never	20
no answer	10

Table (x): Frequency of speaking English

The students speak English mainly in the teaching /learning classroom. They sometimes speak it with foreigners living in Algeria, and with friends on the phone.

4-The difficulties they generally have when speaking English are: Sentence construction and pronunciation.

D- The Listening Skill

1- Is listening necessary in your field of study ?

Sixty four (64) students find it unnecessary to listen to English in their field of study; and twenty six (26) students find it necessary, and ten (10) students gave no answer.

2- Frequency of listening to English

Frequency of listening to English	Number of students
Frequently	03
Sometimes	70
Never	12
No answer	15

Table (xi): Frequency of listening to English

3- Where / when do you listen to English ?

The students listen to English:

- In the classroom;
- On the phone, while talking to friends;
- On the radio;
- On the television;
- In songs;
- On the bus ? (my question mark because it is not clear with whom the person speaks English on the bus, and for which reasons)

Rubric 6 /Course Content: Assessment of the standard of teaching in English lectures.

1-The standard of teaching in English lectures was evaluated as follows :

Quality standard of English lectures	Number of students
Very good	16
Good	40
Poor	20
Very poor	18
No answer	06

Table (xii): Quality standard of English lectures

2- Relevance of English class content to the subject-speciality

Degree of relevance	Number of students
Very relevant	42
Relevant	37
Not relevant	03
No answer	18

Table (xiii): Relevance of the English class to the subject-speciality

3- The teacher's role. S/he helped the students to:

- Learn scientific terminology 66 answers
- Understand scientific facts 54 answers
- Learn the grammatical rules of the English language 33 answers

4- Types of teaching aids used by the lecturers of English

According to all the students, the only teaching aid used by the teacher is: the blackboard.

5- Was / were the teaching aid(s) used constructively? and Why ?

Some students found that the teaching aid (the blackboard) was used constructively and some others find that it was not used constructively for the following reasons.

- Reasons for the positive answer:
  - As some students' level was too low, the teacher helped them by writing the difficult words on the board;
  - To explain difficult words;
  - To explain scientific formulae (chemical / mathematical) etc.;
  - To write the exercises.

- Reasons for the negative answer:
  - The board was not enough to explain and communicate all the scientific information;
  - The teacher did not write any formulae or equations on the board but just 'nice' literary sentences.

6- Learning activities

Seventy five (75) students stated they sometimes do individual work in class; fifteen (15) of them stated that they often do individual work in class; and ten (10) did not answer.

It was noticed from the students' answers that they never do pair/group work (compared to the teachers' answers (Table 3.1) who stated that they gave the students pair/group work activities). The reason for that is the group size (big groups) and the limited time allocated to the English class (one and a half hours/a week)

7- The availability of resources that the students may have needed for the English course :

There are two libraries at the Faculty of Sciences: a main library, and one specific library in each science department. The students mentioned only one type of resource available to them i.e. library books (at the main library)

Types of resources	Number of students	Availability
Library books	40	good
Library books	22	poor
Library books	30	very poor

Table (xiv): Availability of teaching/learning resources

(This table should read that forty students (out of 100) find the availability of resources (i.e. library books) good, etc.).

In the 'fourth year biology library', there are books as well as journals.
In the 1990s, Algeria did not subscribe to many scientific journals, therefore students had to look for, and find the necessary and up to date scientific journals using a 'friend to friend' strategy.

The laboratory facilities for teaching English are inexistent at the Department of Biology.

8- Did the English course help you acquire the necessary skills?
Fifty (50) students find that the English course offered to them, helped them acquire the necessary skills and, forty two (42) students find that it did not.

9- If not, the problems in learning English are due to the
- Lack of adequate material 54 answers
- Lack of teacher's experience 48 answers
- Lack of appropriate syllabus 42 answers. (The students ticked more than one answer)

Rubric 7: Students' Assessment

1- Assessment methods that the teacher used
All the students stated that they were assessed through written examinations only. Another contradiction with the teacher who asserted that he used continuous assessment 'Table 3.2 / Evaluation.'

2- Assessment methods ranked in order of preference of the students:

- Written assignments (essay) 51 answers
- Project 27 answers
- Exams 20 answers
- Continuous assessment 15 answers

(The students ticked more than one answer).

3- Relevance of assessment methods to the course as it was conducted

Seventy (70) students think that the assessment was relevant to the course as it was conducted. However, nine (09) students find that the assessment was not relevant, but they did not state the reason(s).

Rubric 8: Overall Course Evaluation

1-The overall view of the teaching quality of the English course in each year of study.

- **Teaching quality of the English course in the third Year**

Quality of teaching	**Number of students**
Very good	15
Good	45
Poor	20
No answer	20

Table (xv): Teaching quality in the third year

- Teaching quality of the English course in the fourth Year.

All the students who answered the present questionnaire agreed that the teaching quality of their English course was very poor, without really explaining why it was so.

2- Aspect(s) of the English course having been the best taught.

Most students agreed on the following:

- Scientific terminology;

- Study of scientific texts;
- Detailed scientific explanation of concepts, etc.;
- Translation of words / processes in Arabic helped them a lot.

3- Aspect(s) of the English course having been the least well taught

- Not enough grammar and grammar exercises of sentence structure;
- Discourse structure;
- Not enough writing and speaking;
- How to translate scientific texts;
- To read general English.

4- Aspects of the English course that the students regard as having been irrelevant to the course

Some students find that some items taught on the English course did not correspond to the speciality they were studying. However, the students did not state what these items were.

5- Main sources of difficulty for students in pursuing their English course

- Inexistence of syllabus ;
- Inadequate teaching methodology;
- Understanding and speaking (pronunciation) English correctly;
- No balance between the load of information given by the teacher and the time allocated to the English session (too much information given in one session).

6- Suggestions for improving the English course

The students made the following suggestions to:

- Provide them with a syllabus;
- Introduce English in the first year of their biology degree;
- Increase the number of teaching hours;
- Recruit a full-time 'competent' teacher.

The students defined **a competent teacher** as a teacher who: 'masters the English language and the subject-speciality'; 'has a good teaching methodology'; 'takes into consideration the

different levels of students' competency in the same group; but not a teacher whose main aim is only to finish the programme.

- Provide more recent books and textbooks for learning English;
- Use authentic texts (not 'ESP concocted' texts);
- Have more exercises / tasks in order to understand the terminology and the different equipment used in the description of scientific experiments to become able to read texts without much difficulty.

Rubric 9: Some more General Questions

1- The two strengths of the English course

The students focused mainly on the:

- Study of scientific terminology;
- Study of scientific texts.

2- The two weaknesses of the English course include:

- Lack of grammar teaching;
- Insufficient study of scientific terminology.

3- In the last question of the questionnaire, the students were asked what they have gained from the English course and their answers were as follows:

- Study of scientific texts and knowledge of some scientific terminology;
- How to read scientific texts;
- To understand and write English, and to some extent we learnt how to speak English;
- How to use a dictionary;

However, four students (the four beginners) found that they have gained 'nothing' from the English course without justifying such a statement.

3.3.4 Main results of the student questionnaire

From the students' answers, one can draw the following general conclusions:

- The students who answered the questionnaire are mostly female (60 female and 40 male).

- Most of them (86 out of 100) had studied English at the middle and secondary schools and their level in English is intermediate. Four students were beginners because they had studied other foreign languages (Spanish and German) in the middle and secondary schools.

- Most students (78 students) come from the arabophone department and twenty two (22) from the francophone department.

- All the students agreed that English is necessary in their field of study (Rubric 2 / Question 3) and they all stressed the importance of the reading skill (Rubric 2 / Question 4). Much emphasis was made on the need to develop their reading skill through the study of authentic texts and scientific terminology.

- The students stated that the English course (in the third year) consisted of teaching them scientific English through scientific text study and scientific terminology and that it was not the case in the fourth year (Rubric 8, Question 1, very poor teaching quality in the fourth year).

- Most students (70) find writing in English necessary but they mentioned that the main problem they meet in writing in English is mainly at word level (problem of lexis).

- As far as speaking and listening are concerned, the students think that they are useful but not necessary in their field of study.

However, the students expressed some of the problems related to the English course. These are:

- The unbalance between the time allocated to the English course and the level of students (which is generally low);
- Inexistence of syllabus;
- Inadequate teaching methodology;
- Lack of adequate teaching material;
- Lack of teacher's experience.

The suggestions they made for improving the English course were as follows:

c) administrative; and

d) pedagogic .

a) Administrative suggestions

- Increase the number of teaching hours of the English course;
- Introduce English in the first year of the biology degree in order to avoid the void of two years where English is not taught;
- Recruit a full-time teacher;
- Provide more recent books and textbooks for learning English.

b) Pedagogic suggestions

- Provide a syllabus;
- To use authentic texts (instead of 'E.S.P. concocted' ones);
- To have more practice when it comes to the use of the target language (more exercises/tasks, more assignments, etc.) in order to become efficient independent readers.

The problems and suggestions mentioned by the students are significant as they build the foundation stones of the research. This reflection will seek to find solutions to both problems (administrative and pedagogic).

We aim at solving the administrative problems (time allocated to the English course, group size, recruitment of full-time teacher, etc.) through discussion and consultation with the administrative staff (Dean of the Faculty, Head of Departments, etc.)
However, this done alone would not suffice. We should also solve the pedagogical problems.

By solving the administrative problems, we can help in better organising the English course in order to improve its tuition.
One way, to reach such an improvement is by providing an appropriate methodology.

These two conditions are but objectives, when they come to be implemented, they change into major discoursal discrepancies. The person in charge of the administrative ruling tends to magnify the importance of the management of the transmission of knowledge. The teacher, the one in charge of the transmission of knowledge, is often required to bend pedagogical objectives to administrative objectives (e.g. transform a tutorial (T.D) into a formal lecture because of the number of students, as mentioned in Rubric 6 / point 6 'learning activities: group size, limited time, poor pedagogical means, etc.). Therefore, we are aware that what we are going to propose may have a sample 'plausibility' and may be considered as experimental data.

3.4 Specification of the reading literature used by biology students

The methodology aimed at, will be developed to help students improve their reading skill in English when it comes to using scientific journal articles, the discourse of which is highly sophisticated and hence difficult to understand by non-native students. Therefore, the following part will specify the subjects using the reading literature i.e. journal articles, and their reading strategies. It will then examine aspects of the biology discourse, focusing on aspects of complexity.

3.4.1 The Subjects using the reading literature (i.e. journal articles)

In this research, the subjects using the texts under study (biology journal articles) were fourth year biology students at the university of Oran (Algeria) who were taking an English reading course. The aim of the course (as stated previously) was to help them in their reading performance of biology journal articles.

The students were native speakers of Arabic/Berber, who had studied French as a first foreign language, and who had studied English as a second foreign language for six or seven years (with the exception of some students who had studied another second foreign language e.g. German or Spanish)

These students studied English at school (for five years) with a two-year void of English instruction at university. English is not taught at the Department of biology in the first two years of university education. The five years of English at the secondary school are completed by three semesters of English (two semesters in the third year, and one semester in the fourth year) of university education.

The students have some dormant knowledge of English from school, and have some familiarity with some lexis and syntax of biology discourse.
These students were taught biology through Arabic or French medium (and not through English). They are undergraduate and therefore they are not very much exposed to sophisticated English language. However, they, at the same time, have to read journal articles dealing with biology, the language of which is quite complex.

3.4.2 Reasons for using journal articles

Lodge observes "*no choice of a text for illustrative reasons is innocent.*" (Lodge, 1981:23, in Swales, 1990:50). As such, there are three reasons why these students need to read journal articles:

a) Specific scientific information

To get scientific information necessary for the writing-up of their 'mémoire' (written assignment);

b) Availability of resources

Different journals written in English dealing with biology are available at 'the fourth year biology library' as compared to books which cost more for Algeria to import. The students also used 'a friend to friend technique' to borrow journals.

c) Frequency

Journals are published more frequently than books, and they therefore contain more up to date information, and more targeted topics, which correspond to specific learners' queries.

These reasons explain why undergraduate biology students read journal articles. However, as mentioned above, the language of scientific journal articles is, on one hand, quite

sophisticated and difficult to understand by undergraduate students. Halliday explains this point by stating that,

> *It would not be possible to represent scientific knowledge entirely in commonsense wordings; technical terms are not simply fancy equivalents for ordinary words, and the conceptual structures and reasoning processes of physics and biology are highly complex and often far removed, by many levels of abstraction, from everyday experience. (Halliday, 1993:70).*

This statement (as well as Gee's i.e. 'use of language plus other 'stuff'' in 'Chapter Two') establishes the distance between everyday language used for domestic purposes, and technical language for rational and methodological processing of information, as available in scientific journals.

As such, the language in which these journals are constructed is specialised and sophisticated and, is therefore bound to be difficult to follow by non-native students of English.

On the other hand, these students find biology texts hard to read although they know many of the words and grammatical structures used in the texts because, in the case of scientific writing, "*it seems that there are certain features of the way meanings are organized, and the way they are worded."* (Halliday, 1993:124) that present special problems for these learners. It confirms their inability to use procedural devices, to make their terminology a piece of discourse. This fact may be due to the teaching/learning experience they have had. It is, as Carrell observes "*the result of learning the elements of the language without understanding the processes which one utilizes to communicate with those elements."* (Carrell *et al.*, 1988:120). The consequence of such a practice leads the students to the following reading strategies.

3.4.3 Students' reading strategies

As stated in the questionnaire in this chapter, the students read biology articles written in English only when they do not find articles written in Arabic or French. However biology

articles are rarely published in Arabic or French, and as a consequence students have to read these articles in English. When students read articles in English, they either:

- Process them word for word 'local reading' (Cohen *et al.*, 1988:16), and in this case the quality of reading could be questioned: 'Do students achieve satisfactory comprehension ?' (cohesion processing), or
- They scan them ('lire en diagonal' as the students said). When they use this strategy, they may miss pertinent information (coherence processing), or
- They give the articles to an expert to translate. This operation is cost-effective in terms of time and money (discourse processing).

According to the students, these ways of processing their texts are due to the complexity of the discourse of biology.

3.5 Discourse of Biology

The discourse of biology, though it uses general English i.e. common grammatical structures and non-technical lexis, it also uses a very specialised English (biology terms, much lexical density and grammatical metaphor).

This specialised language, 'Langue de Spécialité' (Pavel) or 'Langue Spécialisée' (Lerat), is considered by some linguists as part of the general language, and therefore serves a certain purpose. Pavel states that,

> *The specialised language is part of the general language used for the transmission of knowledge relating to a specific field of experience. It shares the grammar of the general language and part of its lexical and semantic inventory (morphemes, words, syntagms and combinatory rules) but makes a selective and creative use of them...* (Pavel[13], 1993:1, in Gledhill, 1995:28).

[13] *La langue de spécialité est un sous-ensemble de la langue générale [LG] qui sert à la transmission du savoir relevant d'un champ d'expérience particulier. Elle n'existe qu'en partageant la grammaire LG et une partie de son inventaire lexico-sémantique (morphèmes, mots, syntagmes et règles combinatoires) mais* ***en fait un usage sélectif et créatif...*** (Pavel, 1993:1, in Gledhill, 1995:28).

In the same line of thought, Lerat (1995:21) emphasises that, the specialised language (langue spécialisée) stems out of a professional context. It is language as such that develops *"an autonomous system"* that contributes to the transmission of knowledge, and the dissemination of creative ideas.

Our aim in this research is not to dispute the claim whether a specialised language is part of the general language, or whether it is not. We agree that a specialised language is part of the general language "*un sous-ensemble de la langue générale*". Our aim is to focus more on the objective that we, as people, want a specialised language to serve a specific purpose.
It is important to mention that it is not the language itself that serves a certain purpose and also that a language cannot be *'selective'* and '*creative'* (Pavel, 1993:1). Rather, It is people (users of language: scientists, linguistists, etc.) who use a language to achieve purposes, and that only people are 'creative' and 'selective' when they use a language. Crucial to this point is Roe's notion that *"language cannot have a purpose." "**Only people*** (here referred to as 'Discourse Communities' (Swales, 1990:23) ***have purposes."*** *Also for Roe**, "language is what they use purely in order to achieve those purposes.***" (Roe, 1993:1).

One of these purposes crucial to our argument, as stated in Pavel's and Lerat's quotations above, establishes that the specialist language conveys knowledge gathered through a field of experience (Pavel, 1993:1). Furthermore, a language develops a major function, i.e. the transmission of knowledge (Lerat, 1995 :21).

This knowledge transmission is pertinent to our subjects (biology students) as this is the key to help them join their target discourse community. It is also the key process by which our learners (neophytes) become members of the Target Discourse Community (TDC). Therefore, moving from non-members to members of the TDC requires knowledge transmission to take place at the transition phase, one of the CATWOE elements (T: Transformation, i.e. the English language course) referred to in the description of the discourse community in 'Chapter Two.'

The aim of this research is to help this knowledge transmission i.e. understanding biology discourse without "*external or internal mediation*" (knowles, 1995:2) to take place at the

transition phase, since these procedures will certainly make the learner waste time in strategies for deciphering semantic content, and scatter his/her energy over secondary activities of acquisition of knowledge, rather than topical understandings that possession of discourse implies. Therefore, our learners need to be trained in "*the way meanings are organized*" and "*the way they are worded",* (Halliday, 1993:124), to become professionals in their field of study and to be able to join the target discourse community.

This training is crucial for our learners because it is socially important. This social importance is two-fold:

1. So that a community survives and does not die out, neophyte members need to be trained and initiated into this community. There needs to be a threshold level of members (Swales's sixth criterion for the description of a discourse community/ 'Chapter Two / Section 2.3.3.3')

2. These learners need to be aware of the established concepts negotiated and agreed upon by professionals in their discourse community to be able to communicate with each other. Knowles explains this point stating that,

> *... Most of all, names, normally named 'words', are at the very basis of social intercourse. Of course, much social intercourse takes place on a distinctly professional level, invoking/evoking - and instantiating by words-established concepts, phenomena or artefacts current within particular discourse communities for the simple reason that they have been professionally introduced, negotiated and subsequently validated by them.* (Knowles, 1995:2)

This type of communication between members of a scientific discourse community is necessary for its survival. In this regard the importance of the use of 'social networks analysis' (mainly Warner's and Mitchell's works) referred to in the description of a discourse community in 'Chapter Two / Sections 2.2.2.1 and 2.2.2.2', provides us with a useful tool for understanding such phenomena.

The learners need also to know these words and concepts because, as members of the discourse community, they are bound to share the same lexicon, to become a terminology that allows access to discourse as nuances in meaning.

As such, and by helping our learners to bridge this gap i.e. to overcome the reading difficulties caused by scientific texts, and become able to join the target discourse community and communicate with each other "*without internal or external mediation*", we help increase "*the discourse community's sophistication, first of all encyclopaedically (at least for the pioneers) and then terminologically. The route from term to concept is immediate for DC members.*" (Knowles, 1995:2). However, before our learners reach the stage of "*immediate route from term to concept*", they stated (in the questionnaire answers) that they go through the stage of problem facing i.e. the complexity of the discourse of biology.

3.6 Aspects of complexity of biology discourse

According to the students, this complexity of the form of English lies mainly in '*Lexical Density'* and '*Grammatical Metaphor'* (Halliday, 1993:71) widely used throughout the texts by the authors of biology discourse.

These features (Lexical Density and Grammatical Metaphor), which make comprehension difficult to follow, are discussed one after the other.

Halliday's terms '*lexical density'* (LD) and '*grammatical metaphor'* (GM) are referred to in order to come up with a definition specific to the point of study. The former (lexical density) is defined by Halliday as *"a measure of the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure"* and it can be measured as "*the number of lexical words per clause."*

The latter (grammatical metaphor) is defined as "*a substitution of one grammatical class, or one grammatical structure, by another."* (Halliday, 1993:79).

For the present purpose, 'lexical density' will be used to refer only to those lexical items (content words) which form a pile-up of nouns (a noun string) without use of any preposition, or any relative clause, etc. between these nouns (e.g. Ltk mouse fibroblast cell line, a consensus polyadenylation signal sequence, a rat liver expression cDNA library, etc.)

The second criterion of lexical density used by Halliday and which reads "*the number of lexical words per clause"* will not be considered as it is not relevant to the present research.

'Grammatical metaphor' will refer to the substitution of verbs, etc. by nouns, in order to form 'lexical density' i.e. the principle of the metaphor implies a lexical combination (NC) which is within the limits of this discourse, decoded with the same referent.

3.6.1 Lexical Density

Because the language of scientific journal articles is planned and formal, lexical density tends to be quite high. This is why students find the discourse of biology journal articles quite dense.

Halliday (1993:76) refers to some general tendencies to use lexical density by explaining that in formal spoken language, lexical density does not play an as much important role as in the written form which is highly planned as shown in the next grid:

- Spoken language ⟶ two lexical words per clause
- Written language ⟶ six lexical words per clause.

He then adds, that in scientific writing 'lexical density' may be much higher.

This can be easily understood since scientific writing is entirely geared by the need to inform, to manipulate data in a great quantity , and as such it excludes all forms of 'interpersonal formulas'

a) Because of a permanent worry: not to be subjective;

b) And because of the need of accuracy that excludes all references to cultural environment (nothing as the narrator/author, weather conditions, historical moment, etc.) outside the focal interest of the biological item under study.

Thus, in biology journal articles, scientific writers tend to condense their texts by using a pile-up of nominal compounds (NCs). These NCs form the lexical density which will be examined in this research.

3.6.2 Grammatical Metaphor

This use of nominal compounds presupposes nominalisation, i.e., turning almost any linguistic item (verbs, etc.) into nouns (grammatical metaphor), another feature being examined together in this research.

This discourse complexity does in fact present major problems for the present learners because they have not been educated through English. As stated before, English is taught in science departments, just as one component of the curriculum to become a tool for reading scientific literature. For this reason, the students' problem is twofold:

a) The nature of the language used as a medium of instruction.

Halliday observes,

> *For those who are taking up English as a language for science and technology, the problem may be greater or less depending on the degree and kind of grammatical metaphor found in the language(s) they have used as medium of education before.* (Halliday, 1993:82)

And the present learners have been educated through Arabic or French, two languages which *"would resort to a number of different propositions and relative clauses."* (Salager, 1985:140), instead of using grammatical metaphor.

b) The lack of students' exposition to English in general, and to scientific English in particular. In Algeria, the only place where our learners hear, read or speak English is in the classroom (for one and a half to two hours/a week during six/seven years of their educational life). As such, they have grasped only some basic English to survive in their educational field.

So, when they read journal articles full of lexical density (LD) and grammatical metaphor (GM), their problem is obviously more important than that of native speakers considering Halliday's statement that 'nominalisation' can be a problem even for native speakers. This is explained as follows,

> *It seems likely that part of the difficulty arises, however, because these metaphorical expressions are not just another way of saying the same thing. In a certain sense, they present a different view of the world. As we grew up, using our language to learn with and to think with, we have come to expect (unconsciously, until our teachers started to give us lessons in grammar) that nouns were for people and things, verbs for actions and events. Now we find that almost everything has been turned into a noun.* (Halliday, 1993)

This quotation states that the nominalisation process used in written scientific English seems to be a problem even to native speakers.

To summarise, nominal compound (NC) frequency, their rarity in the readers'L1 and L2 and their complexity, represent the major problems for our learners. Therefore, our immediate goal as EFL reading teacher is to minimise reading difficulties, to maximise comprehension, and to save time and money for the students who pay for their articles to be translated.

It is agreed that learners' awareness of frame of reference (their subject-speciality) is obviously richer than that of the language teacher because as Knowles notes,

> *Those who observe and analyse discourse communities because of an interest, even a utilitarian interest, in their technolects are almost exclusively never members of them.* (Knowles, 1995:2)

And language teachers teaching science students' communities have almost never been members of them. The choice of their being appointed as teachers is always determined by the type of code used by a scientific community i.e. the language namely here: English, and never discourse, since in this case the teacher would be chosen for the content of his teaching i.e. biology.

It happens that language and discourse (English + biology) meet in single individual and rare instances (as seen in 'Chapter Two', the teacher Questionnaire answered by a subject-

specialist, a biologist, teaching English), and in Algeria, when the teacher is sent abroad (Great Britain or the U.S.A) to study and complete his/her postgraduate studies in biology.

Often when they come back, they neither have the time nor the preoccupation of developing a reflection on the problems that the biology discourse community encounters. However, the primary concern of this work is to investigate (in 'Chapter Four') how this frame of reference is represented in language, i.e., the English language.

3.7 Conclusion

In this chapter, we tried to assess the teaching / learning of the English course (as an answer to the first research question) through the analysis of two questionnaires (one for the E.S.P teachers and the other for students).

The questionnaires' results showed some strengths as well as weaknesses of the English course. Both teachers and students stressed the fact that reading is the basic skill taught in class. The students also stated that they need to read scientific texts, focusing mainly on journal articles.

The type of discourse (written) and the genre (scientific journal articles) used by biology students have been specified and reasons for such a choice are given.

We then discussed the aspects of complexity of biology discourse represented in 'Lexical Density' and 'Grammatical Metaphor', two aspects that will be dealt with in 'Chapters Four and Five.'

CHAPTER FOUR

FRAMING THE BIOLOGY CORPUS FOR MANUAL ANALYSIS

CHAPTER FOUR

FRAMING THE BIOLOGY CORPUS FOR MANUAL ANALYSIS

4.1 Introduction

This chapter starts by specifying the kind of corpus to be used for analysis. Firstly, the principles underlying the selection and design of the specialist corpus are discussed. Secondly, a manual analysis of grammatical aspects (tense, voice and modals) and lexical items (scientific terms in the form of Nominal Compounds using Bartolic's categorisation) of the whole corpus is done.

4.2 The current pedagogical situation at the Department of biology

The students' questionnaire exploitation, and a first sight interpretation found the most important observations that establish a general unease about the English course they have had so far. Without being able to point out the shortcomings of the lectures with accuracy, they expressed some awkwardness about the methodology implemented by their teachers. The students' pedagogical dissatisfaction is due to two main reasons:

4.2.1 The traditional way of teaching

The reading course generally consisted of giving the students a text, followed by a vocabulary study (usually 'study of lexis in vitro', Knowles, 1995, as opposed to the 'study of lexis in vivo' i.e. in real contexts). The students are, then, asked to answer some comprehension questions. Such a teaching situation seems to be under complete teacher's control (selecting texts, asking questions and expecting correct answers from the students), and relegates the student's role to a secondary role i.e. to answer the teacher's questions only. The students are not involved in any decision-taking as far as the content of the English course is concerned.

4.2.2 The reading texts

The texts (concocted for teaching purposes, not 'authentic texts') used as a basis for the reading course are imposed by the teachers of English without any prior consultation of the students. These texts are not representative of the learners' needs, as no needs analysis is generally done before. Working on this basis, therefore, a students' needs analysis (through a questionnaire, (in 'Chapter Three') and informal interviews of Director and head of Department) was undertaken.

To the question 'Rubric 8 /Question 6: What suggestions would you make for improving the English course ?', all the students agreed that the course should include 'authentic reading materials.' This point is related to Swales's fourth criterion 'use of genres that mediate the in-flow of information', which is important for the 'Transformation process'.

In this context 'authentic reading material' is defined by the students, the Director of the Department of Biology and by the Head of the Francophone Department (personal communication) as 'the actual articles that the students need to read in order to get the relevant information to write their 'mémoire', and be able to join the target discourse community.' That means, not those texts imposed by the teachers of English, and which mostly represent 'simplified versions' of scientific texts, produced specifically for teaching purposes.

Therefore, the first step undertaken in this research is the creation of the corpus. This latter had to follow certain criteria because as Needham notes that *"if we were to select the material at random, it would probably turn out that the data was so sparse that nothing useful could be found."* (Needham, 1967: 47).

4.3 Rationale for the corpus of texts

The creation of the specialist corpus will enable us to answer the students' and the institution's real needs (i.e. to use authentic texts) as revealed in the Questionnaire and informal interviews, and at the same time to use this corpus as the foundation stone on which the suggested methodology will be based.

Two points can be made about the decisions underlying the selection of this corpus:

- The objective of the research, and
- the resources available . (Engwall, 1994:50).

4.3.1 Objective of the research (Reintroducing the second Research Question)

The objective of the research is to find answers to the second research question which is *'How does language function in extremely specialised discourse'*. It will consist mainly of the study of the frequency and complexity of the lexical and grammatical features characteristic of biology discourse used in journal articles.

4.3.2 Resources available

In order to control text-type aimed to be used in this research, two parameters are held constant. These are:

(i) Authenticity of texts

The texts that form the corpus are articles that were read by biology students in their final year of study (semesters 7 and 8) and who were at the stage of writing up their project.

(ii) Text genre and provenance

The combination of 'the objective' and 'available resources' made me restrict the study to one genre: journal articles.

It is obvious that most of their needs in terms of corpus are singularly targeted i.e. all texts that are directly compiled for the drafting of their papers, examinations and research queries. This very functional perspective determined the choice of the corpus, even if teachers (as much as myself) may question the quite reducing view of specialisation, knowing that it may end in the form of a terminology, in the best of cases a jargon, which is far from a discourse capacity.

This choice was also found out in the students' answers in the questionnaire ('Chapter Three'). To the question 'Rubric 5 / Question3: In your speciality, do you have to read:

a) Textbooks, b) articles in journals, c) handouts, other (specify)',

most students (84%) answered that they read journal articles to get the scientific information relevant to their project writing. The students rarely read books or any other genre dealing with their subject-speciality.

The articles selected to form the corpus deal with the fields of 'Biochemistry' and 'Genetics' and were extracted from different scientific journals (see Section 4.5).

4.4 Limitations of the corpus

A fundamental decision that had to be taken in corpus creation concerned the optimum size for the corpus. Engwall stressed that "*no scientific criteria exist for determining the size of any corpus"* (Engwall, 1994 : 51), while Meyer (2002:138) gave more instruction when he stated that "*what we plan to do with a corpus greatly determines how it is constructed: vocabulary studies necessitate larger corpora, grammatical studies (at least of relatively occurring grammatical constructions) shorter corpora."* However, the notion of 'large and short corpora' has not been clearly determined.

In the theory of linguistic text analysis, a corpus of 100 thousand words is said to be representative and sufficient for the purpose of qualitative linguistic analysis. In the present research, this figure had to be balanced against other considerations.

The size of the corpus under study (i.e. fifteen journal articles dealing with two biology sub-specialties 'Biochemistry' and 'Genetics') is fifty six thousand words (a definition of the term "word" will be given in the next part of this chapter).

This corpus seems relatively small for the following reasons:

(i) Lack of students' cooperation

In this research, we aim at using "authentic reading material" i.e. biology articles read by students [and not any other reading material that we can select], in order to find out what complexity these articles present for students and where this complexity lies.

(ii) Number of specialties in the corpus

Another limitation (although not very significant) is that the corpus includes articles from only "Biochemistry" and "Genetics" because only students studying these two sub-specialities were willing to cooperate (as compared to the other students doing other sub-specialities like Microbiology, Animal biology, Plant Biology).

Points 1 and 2 are interrelated because they both show lack of students' cooperation.

Incidentally, the two specialities covered in this research are the only specialities that have always been part of the whole biology programmes (in the 1984-1997 programme, in the 1997 programme which is still in practice, and in the L.M.D, as explained in 'Chapter One/ Section 1.8').

(iii) Research aim

The aim of this research is not to create a glossary or compile a dictionary where millions words are needed, but to have a representative authentic corpus dealing with the field of biology, and from which to extract some lexical and grammatical features indicative of the biology discourse, and propose an appropriate teaching methodology as the methodology used before has been so far unsatisfactory.

(iv) Copyright problem

Some copyright editors denied permission for research. Therefore, only journal articles for which copyright permission was granted are used in the corpus. Application of the criteria listed above and consideration of the limitations of the corpus creation, resulted in the following selection of texts.

4.5 Types and sources of data

A collection of authentic texts representative of the language of biology was compiled for language analysis in order to investigate the lexical and grammatical features characteristic of biology discourse and to discuss their complexity. Fifteen journal articles dealing with 'biochemistry' and 'genetics' were collected from students and photocopied. These, then, formed the present corpus which counts fifty six thousand running words / tokens.

All the articles were read by final year biology students at the writing-up stage of the 'mémoire' and before the stage of joining the target discourse community.

These articles were published in scientific journals like: Biochemistry, Nature, Peptides, Biogenic Amines, Neuropharmacology, and Naunyn-Schmedeberg's Archives of Pharmacology, between 1985-1994 (see List of Texts in Appendix C).

Considering the difficulty of getting copyright permission to initiate a new collection of data, we have decided to keep the articles and papers compiled for this research although they date back to the nineties for most of them. They remain references, since their observations on the scientific language have not much changed ever since. Indeed scientific jargons do not change quickly, and their validity often spans over one generation. This is the time needed to see a culture gradually shift from one hierarchy of values to another, usually better, since societies can only improve. When changes or even mutations take place quickly, they often follow a technological innovation, or a methodological evolution that marks the shift from one mental pattern of research to another one. These mutations are more in the field of epistemological reflection, and it does not seem to have happened since we have resumed our investigation. These articles show that they all share specific characteristics which are listed and discussed below.

4.6 Corpus text characteristics

Engwall (1994:55) suggests that during the selection process in corpus-building, six important features must be considered.

4.6.1 Main features in corpus-building

- ***Form*** (written or spoken),
- ***content*** **(**fictional or factual),
- ***continuity*** *(*whether a text appears regularly or irregularly**),**
- ***preparedness*** (whether a text has been edited or spontaneous),
- ***availability*** (relates to the assembling of corpus material and its electronic processing),
- ***delimitation*** (whether a text has clear boundaries or not).

All the texts selected for inclusion in the corpus conform to Engwall's features.

1) ***Form*** : all the texts included in the corpus have a written form because the study is concerned with the analysis of written discourse.
2) ***Content*** : the texts are scientific as they deal with two sub-specialities of biololgy, and are thus characterised by their factual content.
3) ***Continuity*****:** the texts are continuous because they were extracted from scientific journals that appear on a regular basis.

   Scientific journals and periodicals are said to be continuous. As science is evolving, more research needs to be done and written regularly.

4) ***Preparedness*****:** all the texts have been edited before publication.

   This feature is clearly shown in the articles as at the top or the bottom of each first page, it is stated that 'the article was received on ... (followed by a date)' and 'reviewed on ... (followed by a date)' before being published.
5) ***Availability*** **:** The texts were available in their photocopied form;
6) ***Delimitation*** **:** all the texts have clear boundaries i.e. they all have beginnings and ends.

To these features that characterise and at the same time explain the bases on which the texts were selected, a description of the structure of the articles is seen to be necessary because results of the corpus analysis in 'Section 4.7' will be given by referring to the specific parts of the journal articles (that form the corpus).

4.6.2 Structure of scientific articles

Marchand (1991:24) sees a scientific document as consisting of :

- **non-linguistic indications** (Indices non-linguistiques / Informations périphériques) as the layout, and
- **linguistic indications** including : name(s) of author(s) and origin if any – Title - Abstract (if any) – Introduction - Materials and Methods – Results - Discussion or / and Conclusion – acknowledgements - References.

As the specialist corpus under study deals with biology, it is obviously concerned with experiments. According to Martin (1993:192), experiments have a very clear staging structure: Aim – Method – Results – Conclusion.

- The Aim shows the relation of the experiment to the scientific knowledge being constructed;
- The Method section provides explicit instructions so that the experiment can be replicated;
- The Results stage provides for comparability across replications; and
- The Conclusion relates the results to the purpose of doing the experiment.

4.6.3 Structure of biology journal articles

The length of the articles that form the corpus varies between 5 and 11 pages. The number of authors per article ranges from 2 to 5.

All the articles follow the conventional format. Each of the selected articles had been divided by the authors into: Title – Name(s) - Summary or Abstact followed by key-words used in the article – Introduction – Methods – Results - Discussion.

In the Abstract the authors state the principal objective of the investigation and refer to the methods employed, give major results and principal conclusions.

For the Methods section, the authors used different labels. These are: 'Methods', or 'Materials and Methods', 'Patients and Methods' (when the article deals with a disease i.e Parkinson's disease, Alzheimer's, and so on), or 'Experimental procedures'.

In this study, the label 'Methods' will be used to ease the task of reference.

In 'Materials and Methods', the authors refer to the techniques used to carry out the experiment. In the present corpus, these techniques include:

a) The diet used in the experiment;
b) The treatment: either specifying the instruments used (syringe, dissection equipment) or the medicine administered (e.g. children were rehydrated either orally with ORS or i.v. with Ringer's lactate solution, according to WHO guidelines).
c) External factors like temperature, light, ventilation, time, and so on.

‘Results’ and ‘Discussion’ are the longest parts of the biology articles. Most of the non-verbal data (tables, graphs, photographs, mathematical and structural (= reaction schemes) formulae) appear in the parts ‘Results’ and ‘Discussion.’

In the Results section, the authors give the results they have reached without deep interpretation.

In the Discussion section, the authors discuss and interpret the results which either prove or disprove the hypothesis they have built in the Introduction.

In the present research, the articles will be divided into: ‘Abstract’ – ‘Introduction’ – Methods’ – ‘Results’ - ‘Discussion’, and reference will be made to these parts in ‘Section 4.7’ that deals with the analysis of the corpus.

4.6.4 Tense and voice in scientific writing

Thompson, in his PhD thesis (2001:101), gave a clear and relevant summary of research studies on the use of tense and voice in scientific writing. He found that “Lackstrom *et al* (1972) were the first to claim that tense choice in scientific writing has its own special characteristics, being determined by rhetorical concerns, rather than by time lines (as in narrative).

Lackstrom *et al* posited that the past tense is used to indicate lack of generality, while the present perfect gives a *‘good generalization about past events’*, and the present simple makes a general claim.”

Swales’s study (1981:41) of the introduction section of 48 academic journal articles drawn from academic disciplines (16 articles from the hard sciences, 16 articles from biology and medical field, and 16 articles form social sciences) suggests that the 4 moves that form the ‘Introduction section’ are sometimes lexically signalled and sometimes grammatically signalled (tense choice). Swales’s study focused on ‘Move 2’ that is ‘Discussion of Previous Research’ (DPR), the most extensive section of introductions, to relate the history by describing what previous researchers have found, reported, demonstrated, suggested.

Swales suggests that tense selection may indicate the author's stance towards the cited work.

The use of present perfect might imply closeness to the study while past tense would distance the writer from the cited work. This is clearly exemplified in a table in his work (1981:42) where he showed that:

- **Strong Author-Orientation** is expressed by using the simple past, e.g.
  Merceau and Feyman (1986) *investigated* the case of a ferromagnetic sphere ...

- **Weak Author-Orientation** is expressed by using the present perfect, e.g.
  Local failure rates as high at 18% *have been reported* (Cantin et al 1968) ...

- **Subject-Orientation** is expressed by using the present simple, e.g.
  Immunological mechanisms are also *involved* in patients with scabies (Mellenby, 1944; Falk, 1980; etc.)

All the information given above will be explored in the following anlysis.

4.7 Manual analysis of the biology corpus

In order to confirm the aspects of complexity of the written discourse of biology stated by our learners, an overall study of the fifteen biology texts forming the corpus is done to find out what particular characteristics (grammar: tenses, voice, etc), or lexis may biology texts contain and which of these may cause difficulty.

The present study starts first by looking at the use and frequency of grammatical structures, followed by the use of lexis.

We do not aim in this part to do an exhaustive study of all the grammatical structures (tenses, etc.) that occur in biology texts because as Lerat remarks:

> *At first sight, syntax is of little concern to the study of specialised languages, for its field is as general as that of phonetics: word order, determiners, tenses and moods, everything that characterises the 'spirit' of a language, are as common to*

specialised and non-specialised texts, as are intonation and stress to the actor diction and everyday conversation. (Lerat[14] , 1995: 74)

Instead, we try to show a grammar somewhat indicative of biology discourse i.e. which features (tenses, voice, modals, etc.) characterise biology discourse and to what extent these are used.

4.7.1 The grammar (voice, tenses and modals) in the corpus

A frequency count of voice, tenses and modals was done manually and it gave the following information.

a) Active / passive voice

It has long been accepted that the passive voice is used extensively in E.S.P. / E.S.T corpora, as compared to 'general English' corpora. However, in our study, active verbs exceed passive verbs in most texts (in 9 texts out of 15) with the use of 1594 active forms and 1433 passive forms. This finding joins Tarone et al (1981) and Barber's (1962) who found that the active voice was much more frequently used in scientific texts than expected. A fact that contradicts previous research. Thompson (2001) reported that Robinett (1980) performed an analysis on a combination of texts from scientific, as opposed to *belles lettres* fields, and found that 46% of the verbs occurring in scientific texts were passive constructions, while only 29% of those in the *belles lettres* corpus were passive. Therefore, it should not be simply assumed that the passive voice is used more frequently than the active in science texts.

[14] *"A première vue, la syntaxe a peu d'intérêt pour l'étude des langues spécialisées, car son domaine est aussi général que celui de la phonétique: l'ordre des mots, la détermination, le système des temps et des modes, tout ce qui caractérise le 'génie' d'une langue, sont aussi communs aux textes spécialisés et non spécialisés que les intonations et les accents à la diction d'un acteur et à la conversation quotidienne. » (Lerat, 1995:74)*

b) Tenses

- **The past tense** in its two forms (active and passive) exceeds the present tense in 'Methods', 'Results' and is also used to some extent in 'Discussion.'
- **The present tense,** mainly in its active form is used in 'Abstract', ' Introduction', 'Results' and 'Discussion', and its frequency is the highest in the 'Introduction.'
- There is very little use of **perfective and progressive aspects** in all the texts.

c) Modals

The use of modals, mainly in their passive form, is highest in 'Discussion' with some occurrences in 'Introduction', 'Methods' and 'Results.'

All the above information is summarised in the following tables. These tables show the ranking of tenses and voice by frequency of use in the main text in each journal article forming the corpus.

The first column of each table shows the texts (that form the corpus) and which are numbered from 1 to 15 (according to their length i.e. from the shortest to the longest). Full titles of these fifteen texts (the journal articles forming the corpus) will be found in Appendix C.

- Table 1 contains information about texts 1 to 5;
- Table 2 contains information about texts 6 to 10; and
- Table 3 contains information about texts 11 to 15.

The capital letters mentioned in the tables refer to the different rhetorical sections of these journal articles where the grammatical elements are found. Therefore,

- **'A'** refers to '**Abstract'**; **'I'** to '**Introduction**';
- '**M'** to '**Methods**'; **'R'** to '**Results**'; **'D'** to '**Discussion**'.

For example, line 1 of '**Table 4.1'** should read 'In 'Text 1', the present tense / Active voice is used in 'Abstract', 'Introduction' and 'Methods'; the present tense/passive voice is used in 'Abstract', 'Introduction' and 'Methods' ; etc.

The last column of each table shows the total number of the use of the active and passive voice in each article.

	Simple Present		Simple Past		Perfect Tenses		Modals		voice
Texts	Active	Passive	Active	Passive	Active	Passive	Active	Passive	Active / Passive
Text 1	A – I M	A – I M	I - M	I - R	A	A - I	I	D	68 / 83
Text 2	A – I M – R D	A – I M – R D	A – M R	A – M R	I	I - R	I - D	I – R D	69 / 38
Text 3	I – R D		A – M R - D	M - D		I	I	D	63 / 51
Text 4	I - D	A – I D	I – R D	A – M R - D		I - D	I - D	I - R	51 / 89
Text 5	A – I M – R D	I – R	R - D	I - M	R		M – R D		89 / 67
Total number of Active / Passive voice in texts 1 to 5									340 / 328

Table 4.1: Tenses and Voice used in Texts 1 to 5

Texts 1 to 5 showed the frequent use of the simple past and simple present tenses in their two forms 'Active' and 'Passive') in all the rhetorical sections, as is the case for tables 4.2 and 4.3. Modals are mostly used in 'Introduction', 'Discussion and 'Results.' No use of the perfect tense is made in the 'Methods' section.

	Simple Present		Simple Past		Perfect Tenses		Modals		Voice
Texts	Active	Passive	Active	Passive	Active	Passive	Active	Passice	Active / Passive
Text 6	I - D		R - D	M – R D			D		61 / 61
Text 7	D	A – I M – R D	A – I M – R D	I – M R - D	R	I	I - D		77 / 70
Text 8	A – I R - D	A – R D	A – I M – R D	A – I M – R D	I	I	I – A	D	136 / 107
Text 9	A – I M – R D	I – M R	M - R	M - R		I	I – R D		109 / 114
Text 10	I – D		A – I M – R D	I – M R - D	I - D	I – D	I - D		153 / 60
Total number of Active / Passive voice in texts 6 to 10									536 / 412

Table 4.2: Tenses and Voice used in Texts 6 to 10

Compared to Table 4.1, Table 4.2 shows that the perfect tenses are mainly used in 'Introduction' and Discussion'. More use of modals is made in 'Introduction' and 'Discussion' sections.

	Simple Present		Simple Past		Perfect Tenses		Modals		Voice
Texts	Active	Passive	Active	Passive	Active	Passive	Active	Passive	Active / Passive
Text 11	A – I R - D	A – I R – D	A – M R - D	M – R D	I - D	I	D	D	100 / 112
Text 12	A – I R - D	A – I D	A – I M – R D	A – I M – R D	I - D	D	I – M D	A – I D	136 / 114
Text 13	A - D	R – D	R - D	M – R D			D	D	96 / 90
Text 14	A – I R - D	I – R D	A – M R - D	I – M R – D	D	I - D	D	D	127 / 151
Text 15	A – I M – R D	I – R D	A – I M – R D	A – I M - R	A - I	I - M	I - M- R - D		259 / 226
Total number of Active / Passive voice in texts 11 to 15									718 / 693

Table 4.3: Tenses and Voice used in Texts 11 to 15

Compared to Tables 4.1 and 4.2, Table 4.3 shows the use of modals in all the rhetorical sections of texts 11 to 15. The active voice exceeds the passive in the 15 texts as the bottom line of each table indicates.

These findings show that passives, modals, perfectives and progressive forms which are said to cause problems to non-native learners of English are not very much used, and thus leave us with the main difficulty that was perceived by the learners (the biology students) and which is: lexical density used in the biology texts.

4.7.2 The lexis of the Biology discourse

The lexis used throughout the biology texts is:

a) General language words, some of which, are common to all the authors especially when it comes to presenting results, or describing an experiment. Examples of these,

- 'As table 5 **demonstrates';**
- 'our results s**uggested** (experiment 1) and subsequently **confirmed** (experiment 2) that...',
- 'This study **tested...',**
- 'similar results **were reported'**, etc.

b) Specialised language: Biology specialised language is of three types:

i) Scientific terms consisting of a single nominal and these are mainly derived from Latin and Greek origins and are almost universally used e.g. in English (bacteria, hormone, etc.); in French (bactérie(s), hormone); and in Arabic (بكتيريا) / (هرمون). That means the same word (sometimes with differences in the phonological system) is used in different languages. These words seem to present no major difficulty for our learners as they are common to almost all languages.

ii) Acronyms like PCR (Polymerase Chain Reaction), APP (Amyloid Protein Precursor), and abbreviations ('DA' for 'dopamine')

iii) Scientific terms consisting of a nominal group compound (plasma ketone bodies, soybean amino acid mixture).

The latter type (i.e. scientific terms consisting of nominal compounds) is seen by our learners as an obstacle for the overall comprehension of the texts. This point will be discussed in detail in the following part.

4.8 Lexical density (represented in NCs) in biology discourse

> *Whereas in literature and especially in poetry, new words are complex, in science they are compound.* (Salager, 1984:141)

And this is particularly true for the biology discourse used in journal articles. It is significantly characterised by the widespread use of nominal compounds. As Halliday (1993:172) put it "*science could not be science without deploying technical discourse as a fundamental tool*" and also that "*the choice of vocabulary in a text is largely a function of the subject matter*." (Phillips, 1985:31).

In biology context, 'the choice of vocabulary' is based on highly scientific terms presented in noun strings (Nominal compounds). Therefore, because scientific journal articles published in English address experts (and biology journal articles are no exception), they use a language for the expert which poses severe processing problems for lower-proficiency learners i.e. for our learners who are not experts in biology yet, and who are non-native speakers of English. This problem resides in lexical density which is represented in the use of noun strings that can often prevent students' comprehension of the text from taking place, the reason why NCs are dealt with in this research.

4.8.1 Nominal Compound Specification

Nominal compounds, referred to in this research, are those elements (content words) made up of two individual constituents and more, separated one from the other by a space (Ltk mouse fibroblast, cell line, ...) or a hyphen (CSF-brain barrier).

Nominal compounds are those which start and finish with a nominal **(blood** cholesterol **concentrations**) and may sometimes incorporate an adjective (strand-**specific** primers) or a participle (mock-***transfected*** cells).

Throughout this research, 'Nominal Compounds' will be referred to as '**NCs'.**

4.8.2 Dealing with lexical density (represented in NCs)

The complexity of lexical density is mainly due to the frequency of occurrence of nominal compounds (as will be presented in figures in the following part), and to their length (the number of nouns used to form a nominal compound). Because the discourse of biology journal articles is specialised, it uses very long nominal compounds because "*the more specialised the text, the longer the CNPs* [nominal compound phrases]" (Salager, 1984:142).

The complexity of NCs is also due to their rarity in our learners'L1 (Arabic / Berber) and L2 (French). Common to this point is Trimble's notes on non-native users of technical manuals. He states that,

> *Technical compounds create a different kind of learning problem for the users of manuals. First, they are not a common lexical structure in many languages. Secondly, they do not all yield easily to a literal translation or to turning into simple phrases- even to the point of often confusing the native speaker of English.* (Trimble (1977:58), in Salager 1984:142)

Thus, if we consider the following two examples from one of the texts (Protais et al., 1985:862) in our corpus,

- *"Rats were anaesthetised by inhalation of a mixture of halothane N2O,O2...";*
- *"After incision of the skin of the head, a hole was made in the skull at the following coordinates..."*

We notice that the structure of both examples is made easy to follow and the meaning easy to grasp by the use of the preposition 'of' which indicates 'what modifies what' and 'what is referred to what': "inhalation **of** a mixture **of** halothane" and "incision **of** the skin **of** the head" which, if expressed in most usual scientific style would yield NCs like 'halothane mixture inhalation' and 'head skin incision.'

These structures (multiple nouns - three nouns and more - in sequence) seem to be more difficult to understand by our learners (than the 'of' construction) because of the ambiguity they entail. The learners often cannot find out what modifies what.
Nominal compound structure is very frequently used in technical and scientific writing because "*it is shorter and more direct and therefore the information is conveyed in a more condensed form which has a greater impact upon the reader*." (Bartolic, 1977:260).

However, this impact upon the reader can be double-sided (positive and negative):

- **Positive:** for economic reasons (a NC compacts information);
- **Negative:** for our learners who have to struggle to find out what these NCs refer to.

But, if these structures which cause problems to non-native speakers, and sometimes even to native speakers of English, are still frequently used in scientific and technical English, there must surely be reasons for that.

4.8.3 Rationale behind the use of NCs

Through text study of the biology corpus, it was found out that NCs are used for three main reasons.

(i) Economy

Both scientists and linguists agree that nominal compounds are used for reason of economy. This point is highlighted in some quotations from some authors. Trimble (1979:2) finds that "*English scientific and technical compounds usually represent 'shorthand' versions.*"
Halliday (1993:168) finds that "*the point of both the technical terms and the grammar is to compress as much information as possible into a short space*" and that "*without this condensation scientific texts would become very long, and probably unreadable, even for professionals*."
(Halliday, 1993:172).

In the same line of thought, Salager (1984:140) notes that "*numerous compounds prove more economical than the corresponding phrases which would use many function words.*"
In these quotations all the authors agree that nominal compounds are used for reason of economy in scientific and technical English in general, and this also applies to biology discourse in particular.
The two other reasons, for which nominal compounds are used are:

(ii) Scientific meaning

Nominal compounds (although generally composed of every day words) form a lexis that is specific to the speciality in which it is used. Halliday remarks that "*technical language both compacts and changes the nature of every day words*" (Halliday, 1993:172). That is general language words that become "terminologised" (Sager et al., 1980:242) and thus generate a new meaning. A metaphor given by Knowles is quite relevant to this point, "*words are like chameleons. Change their environment and they change their colour*." (Knowles, 1995:10). As such, some every day words used in a specific subject-speciality become scientific terms.
Accordingly Salager states that "*CNPs correspond to new terms, to some kind of technical jargon*" (Salager, 1984:140) and finds that "*the CNP (compound noun phrase) is crystallized into a fixed expression owing a scientific meaning which the individual constituents do not have."* (1984:142). These individual constituents 'chunk' or 'collocate' together, hence generating a new meaning. This process forms the third reason for the use of NCs in scientific texts, i.e., collocation.

(iii) Collocation

'Collocation' is used to some extent in most scientific discourse, but it is most striking in biology texts where NCs tend to 'chunk' and show strong bonding between the different constituents.

However, these NCs form a jargon which our undergraduate biology students have not grasped yet, and when these are quite long (5 or 6 nouns used in sequence) it is difficult for our learners to tell what is referred to what because:

> *...when three or more nouns are in immediate sequence, we encounter* ***the problem of the direction of modification****, in other words, what modifies what, or what is referred to what..* (Bartolic, 1978:270)

In addition to the difficulty mentioned above i.e. length of NCs, the frequency of occurrence of NCs in biology texts is also seen to present an obstacle for non-native speakers of English. The more frequent NCs are encountered in texts, the more our learners have to struggle to understand the text. This point (NC frequency) is described below.

4.8.4 Frequency and Form of NCs in the biology corpus

The frequency and form of the different NCs are given below.

4.8.4.1 NC frequency and length: The Procedure

Each time a nominal compound was encountered in each text (of the fifteen journal articles), it was highlighted and then the number of its lexical constituents was counted. After this, 5 different columns have been drawn. These ranged from: 2NCs, 3NCs, 4NCs, 5NCs to 6NCs. 7NCs were very rare (4 examples in the whole corpus). Here are some examples of each kind of NC:

- 2NCs: Rotavirus antigens, cholera toxin, adenylate cyclase;
- 3NCs: Blood cholesterol concentrations, density gradient ultracentrifugation;
- 4NCs: Antigen-overlay immunoscreening method, rat liver cDNA library, tissue kallilrein-binding protein;
- 5NCs: Rat liver expression cDNA library, rat tissue kallikrein gene family;
- 6NCs: on-line narrow-bore PTH-amino acid analyzer, APP 770 expression constructs (where APP stands for 3 nouns: Amyloid Protein Precursor)

- 7NCs: rat liver \gt-11 cDNA expression library, full-length APP 695 expression constuct, full-length APP 751 cDNA sequence.

In biology discourse (and this also holds true for scientific and technical discourse), it was noticed that NCs are formed of a headnoun (it is the final noun) preceded by a number of other noun modifiers. These, according to Bartolic *"can be freely placed before a headnoun and their number is mainly restricted by the information that is to be conveyed."* (1978:259)

The tables (**4.4 / 4.5 / 4.6)** below show the length of the NCs (ranging from 2 NCs up to 6 NCs) and their frequency of occurrence in each text forming the corpus, i.e., the fifteen journal articles.
2NCs and 3NCs are the most frequent in biology texts, followed by 4NCs and 5NCs, which are also significantly frequent. There are also some examples of 6NCs, and fewer cases of 7Ncs (4 examples in all the 15 texts).

Texts	Text Length (in words)	2NC		3NC		4NC		5NC		6NC		Total
		1st	rep	1st	rep	1st	rep	1st	rep	1st	rep	
1	1104	60	102	29	70	12	14	03	03			104/189
2	1961	34	64	10	16	03	03	02	02			49/85
3	1964	32	52	08	08	01	01	02	02			43/63
4	2540	48	98	11	15	04	04					63/117
5	2805	56	114	20	31	14	24	06	06	04	04	100/179

Table 4.4: Frequency and length of NCs in Biology Texts 1 to 5

(‘1st ’ stands for ‘first occurrence’ of the NCs, and ‘re’ stands for ‘repeated’ NCs. The last column shows two numbers: the top number is the first mention of the NC, and the bottom number is the repeated use of the NC in each text)

Texts	Text Length (in words)	2NC		3NC		4NC		5NC		6NC		Total
		1st	rep	1st	rep	1st	rep	1st	rep	1st	rep	
6	2867	23	28	09	30	01	06					33 / 64
7	2909	46	124	17	33	04	12					67 / 169
8	3712	48	116	20	46	04	06	02	02	01	06	75 / 176
9	4017	76	106	27	40	11	19	03	03	02	02	119 / 170
10	4050	53	101	07	50	08	54	01	01			69 / 206

Table 4.5 : Frequency and length of NCs in Biology Texts 6 to 10

Texts	Text Length (in words)	2NC		3NC		4NC		5NC		6NC		Total
		1st	rep	1st	rep	1st	rep	1st	rep	1st	rep	
11	4296	92	207	40	73	26	40	15	18	02	02	175/340
12	4400	68	126	22	60	07	11					97/197
13	4667	61	151	08	24	04	08	02	02			75/185
14	4893	118	199	56	101	28	47	08	10	01	01	211/358
15	9815	124	373	42	82	07	11	01	01			174/467

Table 4.6: Frequency and length of Ncs in Biology Texts 11 to 15

While doing the text analysis, it was found out that NCs are rarely used in one text only (Protais et al, 1985). The authors used descriptive phrases rather than long compounds. The author of this text probably found that "*the clarity of the meaning is often better achieved with descriptive phrases than with long compounds.*" (Bartolic, 1978:277). This was also confirmed by our students in informal interviews who found that 'descriptive phrases' are easier to understand than long nominal compounds.

The other reason why the authors (Protais, et al) used 'descriptive phrases' instead of NCs is that they are themselves non-native speakers of English.

After the frequency list of nominal compounds, which importance is not only to show how frequently NCs are used in biology texts, but also to show that the more frequent and the longer these NCs are, the more the overall comprehension of the text by non-native speakers of English can be impeded.

In the following part, we shall study the form of nominal compounds.

4.8.4.2 Form of nominal compounds (based on Williams' categories)

Besides their frequency and length, NCs show a considerable variety of form. This latter can also add another feature of complexity to our learners as Williams observes "*length (of NCs) correlates with formal and functional complexity*." (Williams, 1984:150).

A description of NCs, based on Williams' categories (1984:148) of formal analysis of nominal compounds, is done in this part. Although the field of study investigated by Williams is 'engineering' and the field of study in this research is 'biology', the same categories seem to apply to both specialities (Biochemistry and Genetics), with two more categories found in biology. These tables show authentic examples of the forms of NCs used in the 15 biology journal articles.

Composition of NCs	Category	Examples
2 Nominals	Noun + Noun	Alkali ions , serum cholesterol, lipoprotein fractions, ear vein, …
	Gerund + Noun	Yawning behaviour, seeding density, immunoscreening method, immunoblotting procedures, ...
	Noun + Gerund	Protein staining, codon coding, spiroperidol binding, Acid sequencing,...
	Noun + Participle + Noun	Filter-grown cells, breast-fed infants, apomorphine-induced behaviour,...
	Noun + adjective + noun	Cholesterol-free diets, rat genomic DNA,...
	Noun+ participle + adjective + noun	Placebo-controlled clinical trial,...
	Proper name + noun	Wilcoxon Rank, Ussing chambers, ...
	Acronym/abbreviation+ Noun	DA agonists, D2 receptor, cDNA clones

Table 4.7: Form and Examples of 2 Nominals used in the corpus

Composition of NCs	Category	Examples
3 Nominals	Noun + Noun + Noun	Anion exchange resine, Plasma ketone bodies, Fiber rabbit pellets, Blood-brain barrier,…
	Noun +Gerund + Noun	Urine-collecting bags, rate-limiting step, ..
	Noun+Noun+participle+noun	Cholera toxin-induced diarrhea,…
	Noun+Participle+Noun+Noun	Intracellulose-coated sample foil, …
	Acronym + Noun + Noun	CSF-brain barrier, cDNA clone precursor,...
	Noun + Acronym + Noun	Blood-CSF barrier,...
	Abbreviation+ Gerund+ Noun	DA binding sites,...
	Noun+Adjective+Noun+Noun	Casein essential amino acid,...

Table 4.8: Form and Examples of 3 Nominals used in the corpus

Composition of NCs	Category	Examples
4 Nominals	Noun+Noun+Noun+Noun	Soybean amino acid mixture, casein amino acid mixture, symptom blood pressure control, colony plaque screen filters, dideoxy chain termination procedure...
	Noun +Noun +Gerund + Noun	Tissue kallikrein-binding protein, guanine nucleotide-binding proteins, antigen-overlay immunoscreening method, ..
	Noun + Noun + Participle +Noun +Noun	Cholera toxin-stimulated chloride secretion,…
	Noun + Noun + Adjective + Noun + Noun	Adenylate cyclase dependent receptor sites,…
	Adjective + Noun + adjective + Noun + Noun + Noun	Reverse-phase high-pressure liquid chromatography
	Acronym / N+N+N / Abbreviation	SDS-polyacrylamide gel electrophoresis, NADPH-cytochrome P-450 reductase, ...

Table 4.9: Form and Examples of 4 Nominals used in the corpus

Composition of NCs	Category	Examples
5 nominals	Noun + Noun + Noun + Noun + Noun	Rat Kallikrein gene family members, Sanger's dideoxy chain termination method…
	Noun + Noun + Adjective + Noun + Noun	Adenylate cyclase dependent D1-dopamine receptors, ...
	Acronym + Noun + Noun	APP secretase activity (where APP stands for: Amyloid protein precursor); ...
	Noun + Adjective + Noun + Noun + gerund + Noun	Mouse epidermal growth factor-binding proteins, ...
6 Nominals		Line narrow-bore PTH-amino acid analyser, P-450 catalyzing prostaglandin A1 W-hydroxyclase-activity, ...
7 Nominals		Rat liver gt-11cDNA expression library, Taq DNA polymerase lack proof-reading activity, …

Table 4.10: Form and Examples of 5, 6 and 7 Nominals used in the corpus

The preceding tables show the form of NCs used in the biology texts. This was seen important to find out in order to suggest pedagogical solutions to overcome the difficulty caused by those elements involved in the formation of NCs .

Furthermore, NCs not only have a variety of forms, but their use is not arbitrary. They also perform different functions i.e. the authors of scientific discourse use them to express different functions.

4.8.5 Functional analysis of Nominal Compounds (based on Bartolic's categories)

To speak about the different functions of NCs in scientific discourse, it is important to refer to the way scientists write their articles to understand why NCs have functions.

In 'Chapter Two/Section 2.6', we referred to Myers' work (1990a:142) in which he compared popularising and professional articles. He stated that the professional articles (the focus of this research) create what he calls *"a narrative of science"* i.e. they follow *"the argument of the scientist, arrange time into a parallel series of simultaneous events all supporting their claim, and emphasize in their syntax and vocabulary the conceptual structure of the discipline"*.

As stated differently by an Algerian biologist, before writing scientific articles, biologists start by doing the experiment and find results. Then, when writing the article, they refer to:

i) The aim / objective of the experiment;

ii) The subjects used:

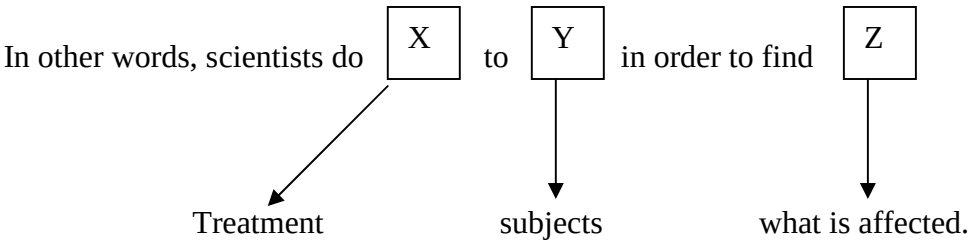
- animals (rabbits / mice / rats,...),
- plants (fungi / geranium,...), or
- human beings;

iii) The method used in the experiment:

- Diet,
- Treatment (specifying the instruments used):
- Kinds of instruments (syringes / scissors / dissection equipment, etc.); and
- Sterilisation of instruments to reduce the error rate of the experiment.

iv) External factors: temperature, light, and ventilation.

v) And they, then prove or disprove the hypothesis.



(Belkhenchir, N., a biologist, 2004, personal communication)

This explanation of how biologists write their articles will lead us to the functional analysis of NCs.

The functional analysis of NCs has been dealt with by a number of authors such as Sager et al (1980), Olsen (1981) and Williams (1984) who actually used Bartolic's categorisation (1978).

For the purpose of our study, it was found that Bartolic's categorisation, although related to a technical field, is more convenient to our analysis of scientific discourse as the same functional categories apply to biology texts and to which some other categories specific to biology discourse have been added.

Bartolic's categorisation (1978) includes:

- *Operation* including the following parts:
    - operating principle,
    - working substance in operation,
    - means of operation,
    - characteristic working part;
- *Materials;*
- *Purpose of Application;*
- *Location ;*
- *Professional Engagement;*
- *Shape / Form.*

These categories (Bartolic's), specific to a technical field, apply exactly to the field of biology.

The following table describes the categories (based on Bartolic's) describing the functions of NCs completed by authentic examples selected from the biology discourse corpus.

To Bartolic's categories, two more categories specific to the biology corpus are added. These are :

- 'Abbreviations / acronyms + headnoun' e.g.

  Abbreviation + Noun: DA receptors (where 'DA' stands for 'Dopamine');

  Acronym + Noun: PCR-based cloning method (where 'PCR' stands for 'Polymerase Chain Reaction') / APP secretase activity (APP: Amyloid Protein Precursor);

- the category 'B of A' e.g. rat brain (brain of a rat), drug intake, stool loss, nucleotide exchange, etc.

TABLE 4.11 a

TABLE 4.11 b

The aim from the study of NCs in biology texts, besides their frequency, was to find out the forms and functions they present so that as teachers and researchers we help 'pedagogically' our learners to overcome the difficulty of recovering NCs so that comprehension of subject-speciality text follows.

It was thought that the only way to help our learners recover the meaning of NCs is to provide them with authentic material for as Williams advises us,

> *NC recovery should be taught in context, for often it is only the context that confirms the NCs underlying meaning.* (Williams, 1980)

In the next chapter, we shall explain how this is significantly true in biology texts forming the corpus under study, where NCS show strong bonding between the different constituents. Lexical density (in the form of NCs) is explained by studying the patterns of significant collocation through 'concordancing.'

4.9 Conclusion

In this chapter, we have highlighted the necessity of selecting a corpus together with specifying the types and sources of data that form the corpus:

- the corpus text characteristics;
- the structure of scientific articles, and more specifically
- the structure of biology journal articles.

Then, a manual analysis of these texts (15 journal articles) is done in details to show what lexical items / grammatical structures are frequently used, and at the same time to confirm (or deny) the aspects of complexity of biology discourse referred to in 'Chapter Three.'

The chapter ends by describing the forms (following Williams' categories) and the functions (following Bartolic's categories) of NCs through authentic examples selected from the fifteen biology journal articles.

CHAPTER FIVE

COMPUTATIONAL ANALYSIS OF THE BIOLOGY CORPUS

CHAPTER FIVE

COMPUTATIONAL ANALYSIS OF THE BIOLOGY CORPUS

5.1 Introduction

'Chapter five' deals with the computational analysis of the same specialist corpus (i.e. the 15 journal articles dealing with biology) that was analysed manually in 'Chapter Four.'
A first description of the corpus of texts (the Rationale, types and sources of data and the corpus text characteristics) was given in 'Chapter Four' to illustrate the manual analysis procedure.

In this chapter, the first part will be devoted to the methodology of designing the specialist corpus for computational analysis. This shows the different stages we should go through to get the specialist corpus ready for computational analysis. The second part will be devoted to the computational analysis of the corpus together with discussion and results of the analysis.

This chapter, the fifth part of this reflection work develops in a more considerable length than the remaining ones. We usually consider that problems of balance (in quantity) between the different parts of a piece of reasoning are methodologically important, since they represent the elements of an intellectual construction. Though length remains no more than an item of form, it may infer sometimes irregular weight in the argumentation and affect the fairness of the ideas handled.
This chapter has a specific consistency, since it represents the backbone of the whole demonstration of our purpose. Indeed, there was a possibility of fragmenting the whole processing of data interpretation as such: e.g.

1) Definition of criteria
2) Typology of nodes

a) –1 node +1 (3 NC)
b) –1 node +2 (4 NC)
c) –2 node +1 (4 NC)
d) –2 node +2 (5 NC)
e) –1 node +3 (5 NC)
f) –3 node +1 (5 NC)
g) miscellaneous (6 NC): this type being rather rare, a few instances would have been gathered in this last item.

The three and four NC being more numerous would have been given more emphasis. One item would have been devoted to each type.

3) Such a distribution of this chapter content in autonomous entities would have sliced the processing of the data and its interpretation. These two steps, in the construction of the reasoning as a whole, are closely related to content, since the corpus, the NC data types, allows us *"to assess the extent to which a pattern is found"* and the interpretation determines our analysis of the contextual factors that influence variability. (Douglas, Conrad and Reppen, 2000: 3)

It is obvious now, that in corpus-based approaches, the core of the analytical inferences matches the bulky amount of data whose quantitative importance can only be dealt with computer calculation. This feature in the methodological approach is thus determined by the use of the computer i.e. the addition of two types of intelligence: the human one (the researcher's) and the artificial one (the computer), as it is admitted as a principle of research that "The essential characteristics of corpus-based analyses are:

- *it is empirical, analysing the actual patterns of use in natural texts;*
- *it utilizes a large and principled collection of natural texts, known as a "corpus," as the basis for analysis;*
- *it makes extensive use of computers for analysis, using both automatic and interactive techniques;*
- *it depends on both quantitative and qualitative analytical techniques*."

(Douglas, Conrad and Reppen, 2000: 4)

5.2 Designing the specialist corpus for computational analysis: The methodology

The corpus of fifteen journal articles from the fields of 'biochemistry' and 'genetics', a representative sample of the biology discourse, is used for qualitative Linguistic analysis. This corpus will be analysed using computer text processing facilities.

The stages followed in this process are described below.

5.2.1 Copyright permission

i) For the present study, the written permission from the authors and editors of different journals from which the articles were extracted, was obtained. The aim was to avoid impinging on the copyright laws as far as the photocopying and scanning are concerned.

We made it clear that reference to these texts, will not provide materials for publications, such as articles or books. These data will be strictly used in the frame an academic reflection, bound to remain a methodological exercise, limited to research perspectives on discourse.

ii) Permission from Dr Peter Roe and Professor Frank Knowles to use ATA (The Aston Text Analyser) was obtained to index the corpus on the computer.

5.2.2 Steps followed in the computational analysis

After the selection of texts based on the criteria mentioned in 'Section 4.3', and the permission for use of these texts was obtained, the computer processing started. Therefore, the texts were optically scanned and converted into the computer readable form with line breaks, ASCII format.

A decision was taken to keep the edited texts identical to the original versions. The sample excluded tables, graphs, photographs, diagrams because:

- They do not form an integral part of the running text;
- The scanner did not recognise completely these non-verbal data.

However, captions of these non-verbal data were included.

The captions were put at the end of the running text in order not to impede cohesion and coherence of the main text.

The reasons for the inclusion of captions can be summarised in two points:

(i) Length of captions; and

(ii) Relevance to the main text.

i) Length of captions

The captions used in the discourse of biology are quite long. They are not just titles, but they vary between 4 and 40 lines. In this context, captions seem to form themselves a whole text.

ii) Relevance to the main text

In these captions, the authors explain the procedure, give results, compare these results with other research, etc. All the information given in captions is related and complementary to the main text.

A biology post-doctoral research student at the University of Aston (personal communication) stated that some biologists look first at the non-verbal data (mainly photographs in plant and animal biology) and their captions before reading the whole scientific article.

There are of course other forms of non-verbal data, such as diagrams and figures that often induce:

a) Either the disregarding of the article;

b) The skimming of the article; or

c) The scanning of parts of it.

It is rather difficult to find statistics of these reading attitudes, since the learners do not keep records of the articles or parts of those they have put aside.

Mathematical and chemistry formulae were retained. Soft-hyphens, those used to break words at line-ends were removed. Hard-hyphens, segment-breaks were retained as well.

The texts were finally indexed to yield the following figures.

- The corpus sample totalled fifty six thousand words / tokens, which yielded
- 5976 types.

A *'token'* is used in computational linguistics to mean *"any word in running text*" and a *'type'* is defined as *"a type of identical token"* (Knowles, 1995). To illustrate this definition, we can say that, in the biology corpus, the word '**gene'** is **a token**, and the word **'genes'** is **a type** (of the token 'gene').

When all the corrections are made and the texts are indexed in the computer, the analysis of the corpus starts.

As the aim of the present research is to find out lexical and grammatical features characteristic of the biology discourse used in journal articles, the analysis will consist of three main steps. These will be explained one after the other.

Step One: The production of word frequency lists using ATA (Aston Text Analyser)
After the specialist corpus texts have been indexed in the computer, the computer converts the running text into frequency lists. These can be selected from the **'file'** menu and are all presented in five columns as follows:

1) All types in alphabetical order;
2) All types in numerical order;
3) Types with a frequency greater than 1/10,000; this list appears under the heading 'non-grammar';
4) Types in 2, excluding 100 'function' words; this list appears under the heading 'frequent';
5) Types in 3, excluding 100 'function' words; this list appears under the heading 'frequent non-grammar'. (Roe, 1995:2)

What these lists mean is that:

- List 1 gives all types in alphabetical order;
- List 2 gives all types from the most frequent to the least frequent;
- List 3 gives the list of 'content-words' (from the most frequent to the least frequent), and excludes 'grammar words'.
- List 4 is the same as List 2 above;
- List 5 is the same as List 3 above.

For the present analysis, the list **'frequent non-grammar'** will be used, excluding **hapax legomina** (i.e. one occurrence in each 10,000 words).

STEP Two: Once we get the list '**frequent non-grammar'** on the computer, we can go to '**the synoptic profile'** and click on any word specific for research.

STEP Three: After clicking on the word in the '**synoptic profile'**, we get the different **concordance lines** necessary for this stage of the research.

This computational process for analysis can be summarised in the following figure as follows:

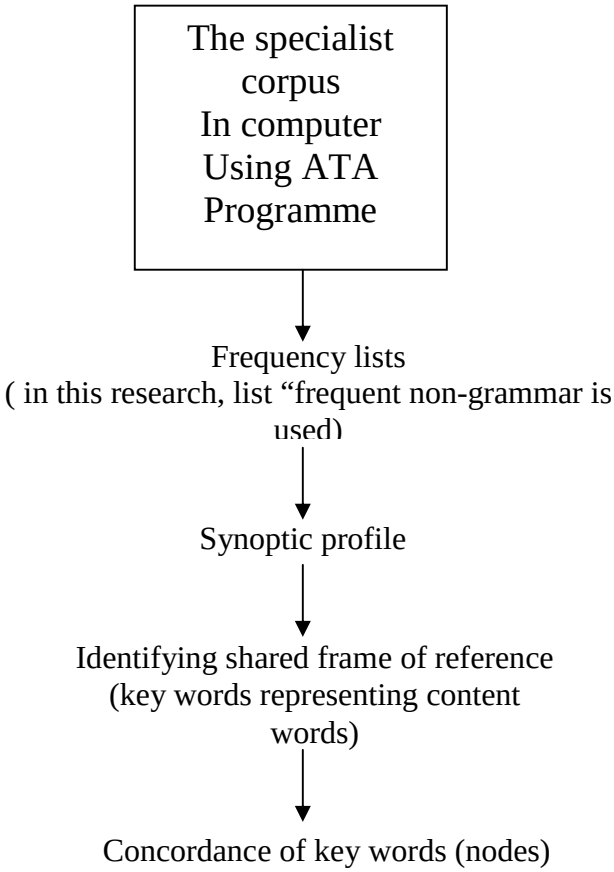


Figure 5.1 : Steps of the computational analysis

5.3 Corpus Analysis – Method and Procedure

In the preceding parts (Chapter Four), we tried to describe how language functions in extremely specialised contexts. We identified 'lexical density' (Halliday, 1993:76) represented in nominal compounds as being the characteristic feature of biology discourse. 'Lexical density' was perceived by Algerian undergraduate students as an impediment for their reading text comprehension.

The concepts of frame of reference are expressed through multi-word segments in biology discourse. These are norms agreed upon by experts in the international discourse community, which our learners are not part of yet, but which they aim to join after the 'Transformational

Process' (Checkland *et al.*, 1990). Therefore, in the preliminary analysis (Chapter Four), all the nominal compounds were counted and classified in categories, and their functions were also explained (following Bartolic's categorisation). These NCs show a strong 'bonding' between their different elements thus forming 'collocational frameworks' in Renouf and Sinclair's terms, 1991.

In the present analysis, we shall try to focus on both the lexis and grammar of the biology discourse. It was noticed that most research using a corpus-based methodology and investigating scientific / technical discourse deals with either the lexis or the grammar of that discourse, but never with the two at the same time. The following works attest to this.

- In his PhD thesis, Gledhill (1996) describes how the English language functions in highly specialised contexts. He tested two hypotheses:

1) The scientific claims of a sample of 10 texts are analysed in terms of reformulation of grammatical metaphor, discourse signalling and posture (Halliday 1985, Sinclair 1981).

2) A corpus analysis of 150 cancer research articles is undertaken to characterise the phraseology of grammatical items in research articles and in the various rhetorical sections of research articles namely Titles, Abstracts, Introductions, Methods, Results and Discussion Sections.

- In his MSc dissertation, Abed Rabbo (1996) used a corpus of 97 thousand running words of electrical engineering discourse and attempted to create a reliable learner's glossary of 400 entries in Arabic/English for electrical engineers trainees at Kalandia Training Centre, Palestine, *"to arm the learner with the confidence needed to survive in the target community where he/she finds the proper learning environment."* (Abed Rabbo, 1992).

- Thouvenin (1996) reports in her MSc dissertation, on computerised investigation of 246,435 words of technical engineering corpus. This latter consisted of the training manuals of naval course in electrical operation and maintenance procedures. She identified frequently-

occurring multi-word units in the corpus and investigated in details nominal groups (ranging from 2 to five individual words) most of which were technical terms:
"The linguistic evidence provided was to assist the creation of multimedia teaching and learning aids for the non-native speakers of English who follow the training courses." (Thouvenin, 1996:3).

Therefore, it is hoped in the present research to investigate both the biology terms and the grammar with which they are used. Hunston and Francis urged us:

> *Because of the emphasis on lexis as the crucial element in language learning, there is a need for a grammar that is lexically based and that can be integrated fully into a lexical syllabus.* (Hunston and Francis, 1998:45)

In this part of the research, three questions related to the research question 2 / Chapter One '*How does language function in extremely specialised texts*' need to be addressed:

1) How are the concepts realised in biology discourse ?
2) What are the ways of expressing them ? (referring mainly to 'collocation', i.e. the co-occurrence of lexical items);
3) What grammatical structures do these concepts combine with ? (referring mainly to 'colligation', i.e. the co-occurrence of grammatical structures).

The answers to these questions will help us to develop an appropriate methodology for non-native undergraduate university students.

5.4 Rationale for the computational analysis

The identification and analysis of biology terms in the form of NCs and their collocates that appear in the corpus was therefore considered an essential step in the development and the improvement of the teaching and learning aids. The role of teaching English in Algeria is, as explained by Thouvenin (1996), "*crucial in ensuring that the quality of training learners matches the quality of the product in both the seller's*" (here acting as actors: teacher, course designer, etc) and "the customers' interests" (in our case the customers are: the students, the administration paying for the students, the government, etc), hence the importance of the two elements (actors and customers) of the CATWOE model referred to in 'Chapter Two'.

It is within the overall notion of improving teaching quality and helping biology learners in the 'Transformational Process' (third element of the CATWOE model and the fourth element in Swales's discourse community model described in 'Chapter Two') that the computerised investigation of the corpus was carried out, for as Hunston and Francis stressed,

> *Because of the more widespread use of electronically-stored corpora among teachers and researchers alike, there is a growing expectation that descriptions of language will be based on quantities of authentic data rather than on a course writer's intuitions and /or language prejudices.*
>
> (Hunston and Francis, 1998:45)

Using an authentic corpus of texts is the essence of this research to avoid having a teaching methodology based on teacher's intuition only.

Two other strengths of using a corpus based-approach have also been pointed out by Biber *et al.*,1994: 169 :

a) Text corpora provide large databases of naturally-occurring discourse, enabling empirical analyses of the actual patterns of use in a language; and

b) When coupled with (semi-)automatic computational tools, the corpus-based approach enables analyses of a scope not otherwise feasible.

Therefore, the main aim from the use of a corpus together with the analysis of collocational frameworks is to raise consciousness of both teachers and students of the different ways we might develop to present and explain language patterning, focusing on one specific discipline.

5.4.1 Process

Frequency lists were produced for the fifteen article corpus. For the analysis of collocational frameworks in this research, the list of 'frequent non-grammar' words of the corpus, produced by ATA, has been used as explained in 'Chapter Five / Section 5.2.2.' For the present study, 100 content words were analysed. These content words were not selected at random, but were selected according to four criteria:

i) **Content-words importance in the field**

These content words were selected by subject-specialists (biologists). The list of all the most **'frequent non-grammar'** words (see Appendix D) was given to biologists. These were asked to underline the most important words in their discipline for analysis.

Being myself a non-specialist, it was thought more reliable to ask specialists to select the important content-words, since they are the ones who are confronted to the complexity of the issue, in practical terms.

ii) **Content-words frequency of occurrence in the corpus**

As a matter of fact, highly significant content words selected by the biologists appear at the top of 'the frequent non-grammar list'. These selected words are tightly related to 'Materials', 'Operation' etc. as explained in the part 'Functions of NCs' in 'Chapter Four.'

iii) **Content-words structure**

The content words that were selected had to be NCs, as these were perceived by the students as impediments for their text comprehension.

iv) **The items under study (content words) as significant figures of the corpus**

when their frequency in the corpus equalled or exceeded 14 occurrences per each 10,000 words.

ATA synoptic profile programme calculates 'collocations' for all items to the left of the node and the right of the node for a span of 4 x 4.

In this thesis, only the two collocates to the left and the right of the node were considered because it was based on the assumption of statistical collocation that *"the closer collocates are to their nodes, the greater the collocational force between them."* (Gledhill, 1996)
In his Phd thesis (1996), Gledhill also states that Phraseological methodologies either treat collocation as directional (either left or right of the node), or informational (collocates are calculated for both sides).

In the present research, the method adopted is informational (as both left and right collocates of the node are considered) to find out what the collocates of the most frequently-occurring items are, in order to categorise these items and develop a teaching methodology accordingly.
Below is an example of an 'informational collocation' of the node 'cells' selected from the research corpus.

***cells**		
central role as receptor	**cells,**	being in close contact
model of chloride secretory	**cells**	devoid of the underlying
cell monolayers when the	**cells**	are being cultured. Thus,
microsomes isolated from the	**cells**	transfected with each cDNA
days following transfection the	**cells**	were assayed for activity

Figure 5.2: Selected concordance lines of the node *cells*

The node 'cells' under investigation is shown at the centre of the concordance lines, between 4 words on either side of the node.

ATA also gives a view of the main concordance lines with a **'view context'** window allowing an overview of the different patterns in which a node may be involved. This is an important tool for pedagogy improvement as we need to study items in a much wider context, and not a study based on concordance lines only.

In the second part of this chapter, we shall begin by looking at a profile of the frequencies of occurrence of the different frameworks and their collocates.
Borrowing specific terms from biology, we can describe NCs in biology as forming templates (the general framework) with a nucleus showing strong bonding. An example of that, would be the NC 'base pair' where 'base pair' shows a strong mutual bonding ('base + pair' always appear together in a text). 'Base pair' is also often preceded by adjectives like 'normal', 'stable' etc., or by chemical formulae (c:06-alkylg, c:o6-ethylguanine). 'Base pair' is generally followed by grammatical words like prepositions, mainly 'with'.
So, to summarise, we can say that 'the template' is '*normal base pair with ...*' and '*base pair'* is 'the nucleus of the template.'

To relate what has just preceded to the field of lexical collocation, and to the purpose of this study, three terms seem to be of utmost importance. These are node, collocate, and span.

5.4.2 Definitions of NODE, COLLOCATE and SPAN

Definitions of the above terms seem necessary at this stage, as these will be referred to in the next part to explain the examples selected from the corpus to illustrate collocation in biology texts.
Phillips (1984:43) summarises the definitions of the three words as follows:

- *A **node** refers to "the lexical item in the focus of attention*";
- **a collocate of a word** is "*a word which co-occurs with the node in the text";* and
- **a collocation** is "*a node-collocate pair*".

To put it simply by using the example ‘base pair’ mentioned previously, we can say that the node is ‘base’, the collocate is ‘pair’ and the collocation is ‘base pair’.

Phillips carries on stating that

> *Collocation is recognised within an environment of a number of words preceding and/or succeeding the node, for example, the 5 preceding and the 5 following words. This environment is termed the ‘span.’* (Phillips, 1984:63)

Depending on the concordance software, the ‘span’ (i.e. the words preceding and following the node) could be either four or five words on either side of the node.

As our main concern in this part is ‘collocation’, so a definition and an explanation of ‘collocation’ are necessary.

5.4.3 Definition of collocation

‘Collocation’ is a word with a number of definitions. Different authors have tried to specify its meaning. Among these authors (Altenberg and Eeg-Olofsson,) who define it as “*more or less equivalent to 'recurrent’ word combination.*” (Altenberg et al. cited in Aarts and Meijs, 1990:3).
Another definition is that given by Cruse and which states that,

> *In the Firthian tradition of British linguistics, however, it [collocation] is generally used in the stricter sense of '**habitually co-occurring lexical items**' or ' **mutually selective lexical items**'* (Cruse, 1986:40, in Aarts and Meijs, 1990:3)

These definitions, although expressed in different terms, state the same idea which refers to those words co-occurring frequently with certain other words in a specific context.

Hoey (2005:3) refers to Partington's definitions of 'collocation' (1998) which he groups into *textual, statistical* and *psychological.*

Partington's '*textual definition of collocation*' is based on Sinclair's and it reads as "*the occurrence of two or more words within a short space of each other in a text.* (Sinclair, 1991:170). For Hoey, this definition is not useful and can result in a woolly confusion of single instances of co-occurrence with repeated patterns of co-occurrence. He prefers using "*lexical co-occurrence"* (instead of 'collocation').

The statistical definition of collocation is "*the relationship a lexical item has with items that appear with greater than random probability in its (textual) context."* (Hoey, 1991a: 6-7, in Hoey, 2005:3). For Hoey, this definition says nothing interesting about the phenomenon; it gives no clues as to why collocation should exist in the first place.

The psychological definition of collocation where Partington refers to two works: Halliday and Hasan, and Leech. Halliday and Hasan (1976: 287) who used 'collocation' as a cohesive device, but they soon found it hard to operationalise and in (1984) Hasan abandons the concept, replacing it with more specific semantic relations (hyponymy, meronymy, etc.)

Leech (1974:20) talks about '*collocative meaning*' which, he says "*consists of the associations a word requires on account of the meanings of words which tend to occur in its environment."* As couched, Hoey (2005:4) finds this definition too general to cover the word in its most common current usage. Therefore, Hoey's definition of collocation is that

> *it is a psychological association between words, up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution.* (Hoey, 2005:5)

Hoey suggests "*priming*" as an explanation of collocation. He explained that in the psycholinguistic literature (e.g. Neely 1977, 1991; Anderson, 1983), the notion of semantic

priming is used to discuss the way a "priming" word may provoke a particular "target" word; and that we can only account for collocation if we assume that every word is mentally primed for collocational use. Also according to Hoey, as a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts in which it is encountered and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. Hoey refers to this property as "nesting" (2005:8). In short, *"collocational priming" is sensitive to the contexts (textual, generic, social) in which the lexical item is encountered and it is our knowledge of a lexical item that it is used in certain combinations in certain kinds of text.* As such *"Primings nest and combine."* (Hoey, 2005:10).

This is also the case in this research. In biology discourse, it was noticed that some linguistic items collocate so often that after some instances, they become predictable. That is to say, given the occurrence of a particular word, the occurrence of other words is rendered more or less likely (e.g. in the biology corpus 'base' followed by 'pair').

5.4.4 Collocation in biology texts

> *The vocabulary of science text forms an interlocking system where, in de Saussure's words, '**tout se tient.'*** (Phillips, 1985:205)

The same idea is reflected in the following quotation which states that:

> *Language use seems to be 'basically a compositional' process, one of stitching together 'preassembled phrases into discourse.'* (Nattinger, 1980, in Carter and McCarthy, 1988:76)

These two quotations suggest that in science, writers follow a certain compositional process in which all the elements of a language need to be stitched together to make all the scientific information clear and explicit. In biology discourse, the striking feature is the way content words tend to 'chunk' or 'coagulate' (Skehan, 1992, in Knowles and Roe, 1995), thus forming a specific compositional process.

As described previously in Part 'Rationale behind the use of NCs in biology discourse', content words are sometimes formed of every day words (receptor, library, binding, etc.), which through their occurrence in biology context generate a new meaning. However, the complexity of these content words does not usually arise from the scientific terms themselves but "*from the complex relationships they have with one another*." (Halliday, 1993:71). As these content words are part of a more general language system, their value is determined by their position in this system and as such they have to be defined by reference to all the other parts. Halliday notes

> *Technical terms cannot be defined in isolation, each one has to be understood as part of a larger framework, and each one is defined by reference to all the others*. (Halliday, 1993:71)

In this research, we aim at raising our "*learners' conscious language awareness*" (Goodman 1990, in Stainton 1992), so that they become able to find out the meaning of content words without much difficulty and without referring to a dictionary. The latter is sometimes of no help as NCs (as described in 'Chapter Four /Functional Analysis of NCs') occur in a specific context and it is in that context that they get a certain meaning, the reason why the dictionary cannot supply that special meaning.

Glaser and Winter (1975) observe that "*compounds with more than 3 elements are rare in chemical dictionaries and conclude that* ***their use as terminological units is heavily text dependent",*** which is also the case of biology compounds. Therefore, a knowledge of collocations is seen nowadays vital for the successful mastery of English. However, before initiating learners into how to find out the meaning of content words, it is useful to give some examples of how content words in biology discourse tend to collocate.

The computational analysis is based on the list 'frequent non-grammar' and begins from the most frequent item (considered here as a node) to the least frequent one.

5.5 Computational Analysis of Content Words as Nodes

This part will be devoted to the computational analysis of the most salient occurrences.

The analysis will focus on the linguistic forms that arise in a specific content-speciality. To this end, Snow, Met and Genese (1989) distinguish between:

a) content-obligatory language that they define as the language that is required to learn a particular content; and
b) content-compatible language: the language that can be usefully taught within the context of a particular content domain. (in Ellis, 2003:235)

We have first identified the content-obligatory items (the 100 nodes) for 'collocation', and secondly, content-compatible linguistic items (tenses, voice, prepositions, etc.), 'colligation.' We are going to analyse 100 content word nodes, and see how they compound in spans of nouns, how they collocate in number of clusters, and how they are used in tenses.

So, each node will be presented as follows:

a) number of occurrences;
b) type of compounds;
c) tenses;
d) forms (or voices).

The list '**frequent non-grammar'** of the biology corpus starts with the expression 'et al' with 317 occurrences. This item is not dealt with, for it does not constitute a content-term. However, it is worth mentioning it here, as it shows how important it is for biologists to mention other biologists' research works. This high frequency of reference to other authors confirm Halliday's statement of a *"larger framework"* which constitutes a network of technical expressions that weave a highly specialised semantic web.

1) <u>Node 'RAT' : 189 occurrences</u>

The node 'rat' was not selected by the biologists, but because of its position in the list (second, after 'et al.') and its high frequency in the corpus (with 189 occurrences), and also because many nominal compounds are formed in biology including the word 'rat', it was decided to study it. 'Rat' is very frequently used because of its importance in biology experiments. it appears mainly in the 'Methods Section.'

It mainly appears in initial position. 'Rat' forms nominal compounds of :

- **Two nouns**: rat kidney, rat sera, rat sequence.
- **Two nouns separated by an adjective** (noun + adjective + noun) : rat genomic library, rat genomic DNA, rat submandibular gland .
- **Three nouns :** rat tissue kallikrein, rat liver expression,
- **Four nouns:** rat kallikrein gene family, rat D2 dopamine receptor.

The item 'rat α1-antitrypsin' forms 'a collocational framework' (Renouf and Sinclair, 1991). This pattern is often preceded by the participle 'purified' thus forming in Smadja's terms a 'predicative collocation' ('RAT' used in second position), e.g. 'purified rat α1-antitrypsin' (10 occurrences), where 'purified' is related to 'α1-antitrypsin' rather than to 'rat.'

There is also an overwhelming use of nominals with the node 'rat' summarised in the following: Hybridization / Purification / Regulation / Detection / Characterization of rat α1-antitrypsin **.**

Interaction / binding between rat α1-antitrypsin and rat tissue kallikrein .

2) <u>Node 'α' : 171 occurrences</u>

'α' appears in 171 instances. It is mainly used before 'antitrypsin' (46 instances) as in 'α1-antitrypsin.'

It is sometimes predicative as it is preceded by 'rat' (31 instances), 'mouse' (4 instances), 'human' (18 instances), and 'purified' (9 instances).

It forms nominal compounds of:

- **Two nouns** as in 'α1-antitrypsin precursor', 'mouse α1-antitrypsin', 'rat α1-antitrypsin.'
- **Three noun compounds** used with 'α' are the following:

rat α1-antitrypsin cDNA (6 instances), 'rat α1-antitrypsin mRNA' (6 instances), 'rat α1-antitrypsin gene' (2 instances), 'rat α1-antitrysin sequence' (2 instances) and 'mouse α1-antitrypsin' (2 instances).

- For the tense and voice of verbs used, the present tense in its active voice is dominant with this item, followed by the past tense in its passive voice.

- '**Is'** is mentioned 13 times in position -1 of the node. It is **either** used with an adjective : rat α1-antitrypsin **is** very l**ow**;

Rat α1-antitrypsin **is close** to 70% ; Rat α1-antirtypsin **is** significantly **similar;**
RAT α1-antitrypsin gene **is single copy**....; **or**
- '**Is'** is used with the passive voice : human α1-antitrypsin **is known to be**;
protein sequence of rat α1-antitrypsin **is aligned with** that;
probe showed that α1-antitrypsin **is expressed at** high.;
rat α1-antitrypsin **is found to be**;
Rat α1-antitrypsin **is highly conserved;**
protein sequences of rat α1-antitrypsin **are shown** in figure, etc.

- The present perfect/ passive voice:
the gene of human α1-antitrypsin **have been revealed**.;
cDNA clone encoding rat α1-antitrypsin **has been isolated from;**

- The simple present tense / Active voice:
data of postcomplex human α1-antitrypsin **reveal** the reaction mechanism;
in human and mouse α1-antitrypsin **renders** the inhibitor;
both human and rat α1-antitrypsin **have** a Met-Ser bond;
purified rat α1-antitrypsin **migrates** as a 54-KDa ;
analyses **show** that rat α1-antitrypsin mRNA levels **increase** during ;
the cDNA encoding rat α1-antitrypsin **shares** 70% to 80%;

- Passive voice / simple past :
purified α1-antitrypsin (5 ug) **was resolved**;
purified α1-antitrypsin (54000) **was recognized** by;
rat α1-antitrypsin **was purified** from pooled.../ from serum;
purified rat α1-antitrypsin **was resolved** in an.

The preposition ‘of’ (13 instances) precedes node ‘α1-antitrypsin.’ ‘of’ is fundamental to the construction of complex nominals.

The preposition that follows the node ‘α-antitrypsin’ is ‘in’ (7 instances).

It is mainly used with biochemical processes dealing with parts of the body (the synthesis of α1-antitrypsin in the lung, 4 instances) and in research processes (the detection of rat α1-antitrypsin in complex biological fluids; ... concentration of α1-antitrypsin in rat serum, Northern Blot analysis of α1-antitrypsin in male and female, etc.)

3) Node ‘ACID’ : 169 occurrences

The node ‘acid’ is used in 169 instances. it is mainly preceded by the adjective ‘amino’ and we thus get 85 instances of ‘amino acid.’ It is followed by the noun ‘sequence’ and ‘residues’ thus forming nominal compounds ‘acid sequence’ (22 instances) and ‘acid residues’ (19 instances).

‘Acid’ is very salient because of the collocational frameworks it forms. It often forms predicative collocations and thus is often preceded by adjectives as in:

amino acid (85 instances); lauric acid (17 instances); nucleic acid (9 instances); ascorbic acid (4 instances); acetic acid (3 instances) and fatty acid (3 instances).

The node ‘acid’ is often used to form Ncs of:

- **Two noun Ncs** preceded by 'amino' : amino acid sequence (22 instances); amino acid residues (19 instances).
- **Three noun Ncs:** casein amino acid mixture (8 instances);
- **Six noun Ncs** (2 instances) : on-line narrow-bore PTH-amino acid analyzer .

‘Amino acid’ also shows Nominal groups (Ngps in ‘Halliday’s terms’) i.e. nucleic acid and amino acid levels as in the following examples :

... in similarity at both **nucleic acid and amino acid levels;**

family at both, tissue kallikrein at **nucleic acid and amino acid levels.**

- Tense and Voice used with 'acid':

The overwhelming tense and voice used with this node is the simple past tense mainly in the passive voice:

12-hydroxylauric acid **was identified** as the...;

The appearance of 12-hydroxylauric acid **was monitored** in the;

corresponding to 12-hydroxylauric acid **was observed** for microsomes;

methanol containing 0.2% acetic acid **was applied** for ;

by feeding certain amino acid diets **was associated** with ;

dietary proteins and amino acid mixtures **was previously associated...**

Nucleic acid sequencing **was performed** by ;

Lauric acid **was incubated** with 50;

The simple present tense in the active voice is also used in a few instances as in :

- Present tense / active voice:

Its N-terminal amino acid sequence **matches** completely with;

protein whose derived amino acid sequence **matches** completely with;

with closely related amino acid sequences that **exhibit** differences.

4) <u>Node 'AMINO' : 151 occurrences</u>

Although 'amino' has been dealt with above in the study of the content word '**acid**', it is thought necessary to deal with it separately as it shows some other different patterns of collocation from those mentioned with 'acid.'

As it has been shown above 'amino' collocates significantly with 'acid' (85 instances). In the synoptic profile of 'amino' we can notice the high bonding between 'amino acid' (singular form / 87 instances), and 'amino acids' (plural form / 56 instances).

'Amino acid' is important in Biochemistry. An 'amino acid' is any of a group of organic compounds containing both the carboxyl (COOH) and amino [NH2] group, occurring naturally in plant and animal tissues and forming the basic constituents of proteins [amine + acid] (The Concise Oxford Dictionary, 1991).

The most striking feature of 'amino' is also the use of the adjective 'essential' (22 instances) preceding the word 'amino', where in the node 'acid' it is used only in two instances.

'Nonessential' is also used with 'amino' but it appears only in three instances and only with the plural form 'acids' in the following examples:

the levels of **nonessential amino acids** corresponded to;

by adding **nonessential amino acids.**

The other adjectives that also precede 'amino' are:

'derived' (4 instances); 'different' (3 instances); 'ketogenic' (3 instances); and 'predicted' and 'translated' (2 instances each).

The noun 'Casein' is also important as it is in the second position after the adjective 'essential.' 'Casein' precedes 'amino' in 10 instances. It is itself always preceded by a percentage i.e.

'... to a 45% casein amino acid diet' (3 instances); or

' a diet containing 30% casein amino acid mixture ' (2 instances), etc.

The third feature of the use of the word 'amino' is the right-hand collocates in position +2. The importance of these words resides in the fact that their frequency can be synonymous to significant collocation.

Therefore, 'sequence' is used in 21 instances (amino acid sequence); 'residues' (amino acid residues, 19 instances); 'sequences' (9 instances) as in 'the amino acid sequences', generally preceded by the participle 'predicted' thus giving 'The predicted amino acid sequences' (9 instances).

'diet' (8 instances) and 'diets' (6 instances) also follow the node 'amino.'

The overwhelming characteristic of the item 'amino' is the use of adjectives and their high frequency as mentioned above.

The other feature is the use of long nominal compounds. 'amino' is used with **nominal compounds of two nouns** : adjective(amino) + noun + noun: 'amino acid diet', 'amino acid levels', 'amino acid mixture', 'amino acid residues', 'amino acid sequence', etc.

Nominal compound including **three nouns**: noun + modifier (amino) + noun + noun , example 'casein amino acid diet', 'casein amino acid mixture', 'N-terminal amino acid sequence', etc.

As can be seen 'amino' sometimes appears in the initial position as in 'amino acid sequence.' It also sometimes appears in the middle of nominal compounds 'casein amino acid diet', 'N-terminal amino acid residues' or 'N-terminal amino acid sequence.'

‘amino’ is always followed by acid, thus showing a strong bonding between the two words forming the template ‘amino acid’ this being a fixed term used in biochemistry, chemistry and biology.

5) Node ‘SEQUENCE’ : 140 occurrences

The node ‘sequence’ appears in 140 instances in the whole corpus. It forms different kinds of Ncs . It is sometimes used in initial position as in some two noun Ncs e.g. sequence identity. It is mainly used in final position as in the following examples forming either a 2 noun Ncs or 3 noun Ncs.

- **2 noun Ncs** : nucleotide sequence (17 instances); cDNA sequence (7 instances); rat sequence; protein sequence; amino-acid sequence; α1-antitrypsin sequence.
- **3 noun Ncs** : rat α1-antitrypsin sequence;

‘Sequence’ also appears in the middle position in 3 noun Ncs and 4 noun NCS:

amino-acid sequence analysis (3 noun NC)

N-terminal amino-acid sequence analysis (4 noun NC)

As can be noticed, the item ‘sequence’ is salient as it is much used to form nominal compounds of different length (2 noun NCS, 3 noun Ncs and 4 noun NCS).

6) Node ‘PROTEIN’ : 138 occurrences

‘Protein’ is used in 138 instances in the whole corpus. A protein is a molecule formed by amino acids.

It is used in initial position and it mainly collocates with ‘sequence/sequences’, and ‘concentration’ thus giving: ‘protein sequence’ (3 instances), ‘protein sequences’ (2 instances), and ‘protein concentration’ (2 instances).

It is also used in final position as in:

- **Two noun Ncs** ‘soybean protein’, ‘binding protein’ ;
- **Four noun Ncs** as : ‘tissue kallikrein binding protein.’
- The tense and voice used are the simple past tense in the passive form :
  ‘purified **protein was concentrated and desalted’.**

7) <u>Node 'CELLS' : 137 occurrences</u>

The node 'CELLS' (used 137 times in the corpus) is preceded by scientific codes COS-1 (14 instances), CL 19A (13 instances), Ltk (4 instances), sf9 (3 instances), HRT-18 (2 instances), T-84 (2 instances) thus forming compound nouns of the form 'abbreviation or code + noun (as explained in 'Chapter Four', e.g. 'ltk cells' , 'HRT-18 cells', etc.

'Cells' forms compound nouns of the kind **'Noun+Noun'** as in :

ganglion cells (3 instances), receptor cells (2 instances) etc...

It also forms **3 noun compounds** as in 'Mouse Ltk cells'.

The striking features of 'cells' are that :

(i) The node 'cells' is mainly preceded by adjectives and participles.

The adjectives preceding 'cells' are : 'epithelial cells' (5 instances), 'endocrine cells' (4 instances), 'enterochromaffin cells' (2 instances), and 'eukaryotic cells' (2 instances).

The participles preceding 'cells' and acting as adjectives are:

'transfected cells' (4 instances), 'infected cells' (3 instances); 'cultured cells' (2 instances) and 'mock-transfected cells' (2 instances).

(ii) The use of simple past tense / Passive form :

There are 19 instances of 'cells' used in the simple past tense / passive voice with 'were' and 6 instances with 'was' in both affirmative and negative forms :

- was + past participle

PKC in cl 19A cells was measured after treatment;

activity in cl 19A cells was not significantly affected;

PKC activity of the cells was not affected by ...

- were + past participle

microsomes from COS-1 cells incubations **were performed** with;

By the 20th day, cells **were studied** on the ;

weekly immunoreactive ganglion cells **were also observed.**

- **The present tense** is also used.

- **Active form** : Cl 19A cells **mount** a chloride secretory;

in most or all cells that **express** it ;

T-84 cells **are** capable of electrogenic ;

microsomes prepared from untransfected cells which **exhibit** negligible capacity; One advantage of using cells in culture **is** that.

- **Passive form :**

Cells **are cultured** in the ;

The filter grown cells **are shown** in Table.

8) <u>Node 'DNA' : 122 occurrences</u>

DNA is a salient node in genetics. It means 'deoxyribonucleic acid' (i.e. deoxyribose nucleic acid) and is important in the determination of physical qualities like the eye colour, skin colour, etc.

Pollack defines 'DNA' as *" a large molecule assembled from the atoms of just 5 elements – carbon, phosphorous, nitrogen, nitrogen and oxygen."* (Pollack, 1994: 17)

'DNA' shows strong bonding when used with certain words e.g. 'template', 'polymerase.'

In our corpus, 'dna' forms **nominal compounds of two nouns**. It is used in initial position as in these instances : 'dna polymerase' (14 instances), 'dna strand' (13 instances), 'dna duplex', 'dna synthesis', 'dna replication', 'dna complex' and 'dna library.'

- 'DNA' is also used in final position as in : 'template dna' (14 instances),
- 'rat genomic dna' (6 instances), and 'sperm dna' (4 instances).

- The tense that is mostly used with 'dna' is **the simple past tense in the passive voice:**

10 ug of **rat dna was digested**;

molecular weight of **rat genomic dna was purified**;

the **elongated dna was separated,**....

9) <u>Node 'cDNA' : 118 occurrences</u>

'cDNA' appears 118 times in the corpus and it means 'complementary DNA'.

It forms NCS of two nouns: (Noun+noun) as in 'cDNA library' (10 instances); 'cDNA probe' (8 instances); 'cDNA clone' (7 instances); 'cDNA clones'(7 instances); 'cDNA sequence' (7 instances).

- **Three nouns** : rat α1-antitrypsin cDNA (5 instances), α1-antitrypsin cdNA probe (2 instances);

- Four nouns : rat brain cDNA library (10 instances);

rabbit kidney cDNA library (4 instances), and rat liver cDNA library (2 instances).

- The tenses used with this item are:

- **The simple past in the passive form :**

cDna **was examined;** cDNA **was inserted**; the cDNA **was expressed** in eukaryotic; the cDNA **was rescued**, etc .

- **The simple past / active voice** :

Cells transfered with each cDNA efficiently **catalyzed** the;

the screening of kallikrein-related cDNA clones, we **found** that.

The simple present tense :

- **Active voice :**

Nucleotide sequence of this cDNA clone **is** 1306 base pairs;

Reading frame in this cDNA **codes f**or a 415;

Enzymes encoded by each cDNA **can compete** effectively (modal auxiliary)

that rklk10 **is** the cDNA encoding protein k10 ;

- **Passive voice :** mRNA corresponding to this cDNA **is localized** in tissues.

10) <u>Node 'CHOLERA' : 96 occurrences</u>

'CHOLERA' is a disease and its importance in the corpus lies in the case studies done by biologists with people affected by this disease.

'Cholera' strongly collocates with 'toxin' in 68 instances, thus giving a **2 noun Ncs** (cholera toxin, 68 instances).

The other striking feature of this item is the use of the prepositions namely 'by / of / in / on / etc.' These are some of the examples that appear in the corpus:

'**BY'** ... induced **by** cholera toxin ; ... activated **by** cholera toxin; ... evoked **by** cholera toxin.
'IN' Secretory reflex **in** cholera toxin; Nervous involvement **in** cholera toxin
'OF' Inhibition **of** cholera toxin ; Action **of** cholera toxin **on** electrical parameters
Action **of** cholera toxin **on** intestinal secretion
Effects **of** cholera toxin **on** net fluid.

Actually, the use of these prepositions refers back to the nouns / verb with which 'cholera toxin' is used, as in English we normally say: 'The effects of', 'action of' to form nominal groups, but we say 'activated **by'** (because considered as a passive form).

11) Node 'TOXIN' : 77 occurrences

As is explained above 'toxin' highly collocates with 'cholera' in 68 instances. It thus forms **two noun NCs**. It also forms other nominal compounds (different from those mentioned in the node 'cholera'). These are: 'incubation' and 'response' : 'toxin incubation' (4 instances) and 'toxin response' (2 instances).
The tense used with 'toxin' is the simple past (both forms)
- **Active voice :**

Cholera toxin **stimulated** chloride secretion;
Isc induced by cholera toxin **was** significantly higher than.

- **The passive form (3 instances) as in** :

cholera toxin **was added.** (1 instance); cholera toxin **was studied** (2 instances).

12) Node 'DIARRHEA' : 74 occurrences

The node 'Diarrhea' appears 74 times in the corpus. 'Diarrhea' like 'cholera' are diseases.
'Diarrhea' does not form Ncs except in one case 'noncholera diarrhea' (7 instances).
'Diarrhea' is always preceded by adjectives which are:
acute diarrhea (4 instances), watery diarrhea (4 instances);
acute watery diarrhea (4 instances), acute noncholera diarrhea (4 instances);
persistent diarrhea (2 instances) whereby 'Diarrhea' shows a strong bonding with the use of adjectives.

- The tense used with the item 'diarrhea' is mainly **the simple past tense :**

Active voice : diarrhea **was** significantly greater**;**

diarrhea **was** significantly lower**.**

Passive voice : ... diarrhea **was defined** ;

347 diarrhea cases **were recorded.**

13) Node 'THYMINE' : 69 occurrences

The node 'thymine' is used 69 times in the corpus. 'Thymine' is the element that makes the difference between the DNA and RNA. 'Thymine' is present in the DNA whereas 'Uracile' is present in the RNA.

The specificity of the node 'thymine' is its use to form nominal groups (in Halliday's terms) instead of forming Nominal compounds. Thus we can find 'thymine' preceded by the prepositions 'of' (40 instances) as in the phrase 'Incorporation of thymine and cytosine' or 'for' as in 'Incorporation for thymine ...' (4 instances) where the preposition 'for' is usually preceded by units of measurement e.g. calculated Km values for thymine; ... and 26.4 uM for thymine; ... the observed Km for thymine; ... mol of enzyme for thymine.

- The tense used with 'thymine' is **the simple past in the passive voice** :

Cytosine and thymine **were observed;**

3'terminal thymine **was paired** ; cytosine and thymine **were incorporated.**

14) Node 'DOPAMINE' : 66 occurrences

'Dopamine' is used in 66 instances. It forms **two noun NCs** with right-hand collocates: dopamine receptor (17 instances); dopamine agonists (09 instances); dopamine receptors (09 instances); dopamine antagonists; dopamine agonist (4 instances); dopamine antagonist (2 instances) and dopamine neurons (2 instances).

- Left-hand collocates are abbreviations like 'D2', D2 dopamine (21 instances) or adjectives like 'various dopamine' (7 instances); 'indirect dopamine' (3 instances); 'postsynaptic dopamine' (3 instances); 'direct dopamine' (2 instances).

-The tenses used with the item ‘dopamine’ are: **The simple present tense** and **the present perfect tense in the passive voice.**

- The simple present tense :

Adenylate cyclase by dopamine **is reduced**; The D2 dopamine receptor **is found**;

The D2 dopamine receptor **is known** ; Dopamine receptors **are classified** into.

- The present perfect :

The D2 dopamine receptor **has been implicated .**

a rigid analogue of dopamine **has been suggested to .**

- The other tense used with 'dopamine' is **the simple present tense in the active form :**

The standard direct dopamine agonist **is** effective; The D2 dopamine receptor **interacts** with; Doses **stimulate** ; Postsynaptic dopamine receptors and **induce.**

- The gerund is also used with ‘dopamine’ as in these examples :

a series of dopamine antagonists **belonging to** various;

3 relative potencies of dopamine antagonists **regarding** various;

The only dopamine (DA) antagonists i**nhibiting** climbing;

The dopamine agonist **regarding** the gnawing;

Dose-response curves of various dopamine antagonists **regarding** climbing.

- The simple past (active and passive voice):

This dopamine binding **was** not responsive (Active form)

After the injection of dopamine agonists, mice **were introduced** (Passive voice).

15) <u>Node ‘CYTOSINE’ : 63 occurrences</u>

The node ‘cytosine’ is used in 63 instances in the corpus. ‘Cytosine’ is the element present in the DNA and the RNA.

‘Cytosine’ has been mentioned with ‘thymine’ as they collocate. ‘Cytosine’ does not form nominal compounds, but it forms nominal groups in Halliday’s terms. We thus have :

Thymine and cytosine (18 instances); Thymine or cytosine (12 instances);

Incorporation of thymine and cytosine (10 instances);

Incorporation of thymine or cytosine (12 instances).

The tenses used with the node 'cytosine' are :

- The simple present tense (Active / passive voice)

Active voice : Thymine and thymine **have** similar Km values;

with cytosine however, there **is** no;

pre-steady state burst with cytosine. This **indicates** that ;

Passive voice: although thymine and cytosine **are incorporated;**

These normal thymine and cytosine nucleotides **are replaced** by .

- The simple past in the passive voice: Cytosine and thymine **were incorporated.**

thymine or first second cytosine. The line **was generated.**

16) <u>Node 'HUMAN' : 63 occurrences</u>

Although 'human' is an adjective, and not a noun, it has been selected by biologists because it is important in the biology field. It is also selected because it is used in the middle of some Ncs e.g.

Activity in **post mortem human frontal cortex** or **post mortem human putamen.**

'Human' collocates mainly with 'Post mortem' on the left (4 instances). The right-hand collocates are mainly 'α1-antitrypsin' as in 'human α1-antitrypsin' (18 instances) and 'brain' as in 'human brain' (6 instances).

The tenses used are :

- T**he simple present tense / Active voice:** Identity between rat and human α1-antitrypsin **is** close to ; The soluble derivatives in human brain **are** potentially amyloidogenic .

Soluble APP derivatives in human brain **end** within the;

as well as its human counterpart **does** ;

data of postcomplex human α1-antitrypsin **reveal** the reaction.

- **The simple present tense / Passive voice :** Human α1-antitrysin **is shown** to ;

Soluble APP derivatives in human CSF **are labelled** ;

Binding activity in post mortem human putamen **is less pronounced.**

- The simple past / passive voice: Human PSA **was shown** to ;

matter grossly dissected from human cerebral cortex **was obtained.**

at 37°C only human α1-antitrypsin **was found** to.

17) Node 'NUCLEOTIDE' : 62 occurrences

'Nucleotide' exists in the protein structure. It is an important element in the biology domain as well as in our corpus as it forms different NCs. 'Nucleotide' is generally preceded by adjectives and as such its left collocates are adjectives like: single, correct, normal, paired, etc.

On the right, 'nucleotide' collocates with the following nouns and it forms **two noun NCs** :
Nucleotide sequence (16 instances); Nucleotide analogues (4 instances);
Nucleotide exchange (3 instances); Nucleotide numbering (3 instances) and Nucleotide sequences (3 instances).

- The tense used is mainly **the simple present in the passive voice :**

The nucleotide sequence **is bounded;** The nucleotide sequence **is numbered.**

- The gerund is also used with the node 'nucleotide':

Addition of a correct nucleotide **following** thymine in;
of the next correct nucleotide **following** the incorporation of .

18) Node 'RECEPTOR' : 59 occurrences

'Receptor' is a node that has been used 59 times in the present corpus. 'Receptor / receptors' is a term important in neurology. It forms nominal compounds mainly with left-hand collocates: 'D2 dopamine' (16 instances), and adjectives like 'B2-adrenergic' and 'α2-adrenergic.'

As right-hand collocates, 'receptor' is mainly followed by 'cells' and 'sites' thus forming

- two noun NCs : 'receptor cells' and 'receptor sites.'

The tenses used with the node 'receptor' are :

- **the simple present tense / active voice** as in these examples

The D2 dopamine receptor **interacts** with guanine-nucleotide-binding ;
a nicotinic receptor antagonist **blocks** the nerve.
Down-regulation of post-synaptic receptor density **is** apparent.

- **The simple present tense / passive voice :** The D2 dopamine receptor **is found** in ltk ;

The D2 dopamine receptor **is known to be**; The increase in muscarinic receptor stimulation **is relayed by .**

- **The simple past / passive voice :**

CI 201-678 on D2 receptor activity **was measured** according ;

a nicotinic receptor blocker, **was given** intravenous.

- **The present perfect / passive voice:** The D2 dopamine receptor **has been implicated i**n; This receptor **has been characterized** on .

19) <u>Node 'RECEPTORS' : 58 occurrences</u>

The node 'receptors' is used 58 times in the corpus. It is mainly preceded by 'DA' (standing for dopamine) 12 times and by 'dopamine' (written in full letters) 8 times. It is also preceded by 'G-protein coupled' as in 'G-protein coupled receptors' (4 instances).

'Receptors' appears mainly in the following tenses:

- **The simple present tense / active voice :** D4 subtypes of DA receptors **is** a plausible explanation.; The stimulation of these receptors **leads** to an increase.

Doses **stimulate** postsynaptic DA receptors and **induce** an apparent.

The simple present tense / passive voice: Dopamine receptors **are classified** into;

23390 to postsynaptic D1 receptors **is unchanged**; High density of NT receptors **is found** at the; Synaptic level since NT receptors **are found** on; DA Serotonergic receptors **are seen** at high; .of the family receptors that **are coupled to.**

- **The present perfect / passive voice :**

Sch 23390 to 5-HT2 receptors **has been described p**reviously;

NT receptors and DA receptors **has been demonstrated** by;

2 subtypes of DA receptors **has been suggested** by ;

Recently, VIP receptors **have also been characterised**.

20) <u>Node 'SECRETION' : 56 occurrences</u>

'Secretion' is used 56 times in the corpus. In position -1, 'secretion' collocates with some nouns forming **two noun NCs** as in:

'chloride secretion' (19 instances); 'cholera secretion' (2 instances); 'fluid secretion' (17 instances). It also collocates with adjectives as in 'intestinal secretion' (6 instances).

- The tenses used with 'secretion' are mainly :

- **The simple present tense / active form :** cholera toxin **influences** intestinal secretion ;

  Cholera toxin-induced net fluid secretion strongly **indicates** ;

  cholera toxin **induces** chloride secretion ; implying that VIP **induces** secretion;

  DbcAMP that **induces c**hloride secretion in CL19A cells;

  agents **induce** electrogenic chloride secretion.

- **The simple present tense / passive voice :**

  Cholera toxin on intestinal secretion **is principally mediated** through the activation;

  Inhibition of cholera secretion **is observed;**

  Cholera toxin stimulated chloride secretion **can be explained** predominantly (a modal auxiliary in the present tense).

- **The simple past / passive voice:** Cholera toxin-induced net fluid secretion **was studied.**

21) <u>Node 'BASE' : 53 occurrences</u>

The node 'base' is used 53 times in the corpus. It thus forms strong bonding with the words 'pair' and 'pairs.' There are 9 instances of 'Base pairs' in the corpus and 8 instances of 'base pair.'

'Base' is preceded by 'cytosine' (8 instances) and 'thymine' (4 instances). 'Base' forms NCs of the kind **'Noun + Noun'** as in 'base pair' and 'cytosine base.'

The dominant tense used with the item 'base' is the simple present tense in its both forms, and the gerund .

- **Simple present tense / active voice :**

  presence of an 06-alkylg:C base pair **perturbs t**he ;

  Cytosine in an 06-MeG:C base pair **is** slower ; because the 06-methylg:T base pair **retains** the ; The 06-alkylguanine:C base pair **adopts** a ;

  This cDNA **contains** 737 base pairs comprising the sequence.

- **simple present tense / passive voice:** nucleotide after a normal base pair **is limited** by.

- **Gerund :** Base pairs **comprising** the sequence;

  the structure of base pairs **involving** 06-alkylguanine;

22) Node 'CELL' : 50 occurrences

The node 'cell' appears 50 times in the corpus; whereas the node 'cells' appears 137 times in the corpus.

The term 'cell' in biology is a unit of each living organ. A group of cells forms a tissue and a group of tissues forms an organ.

The node 'Cell' appears in our corpus mainly in initial position. It forms **2 noun NCs** with right-hand collocates. It thus gives 'cell line' (13 instances), 'cell monolayers' (13 instances), 'cell culture' (6 instances), 'cell lines' (3 instances) and 'cell bodies' (2 instances).

'Cell' forms long NCs and is generally used in the middle of these words (NCs) as in the following :

The human colonic adenocarcinoma cell line; The human colon carcinoma cell line; Cloned human colonic carcinoma cell line; The Ltk mouse fibroblast cell line .

The myenteric ganglion cell bodies;

grown on Transwell porous-bottomed cell culture dishes Icostar .

The tenses used with this item are :

- The simple past / active voice:

Protein content of the cell monolayers **were** fairly comparable;

Cell monolayers **attained** stable electric; In total, CL19A cell monolayers **displayed** the characteristics.

- The simple past / passive voice: Cells **were treated.**

Cell monolayers **were treated with**; Cell lines **were not cloned;**

Electrical resistance of the cell monolayers **was not changed .**

- The gerund and **participles** are also used with the node 'cell.'

An example of gerund is : a cell line **expressing R**GB-2 .

- As participles we have : cell monolayers **obtained** on Transwell ;

The human colon carcinoma cell line HT-29 **grown** on.

However, with the cell lines **studied** previously, there; Cell monolayers **obtained** were electric ; Afferent fibers have their cell bodies **localized** in the.

23) Node 'mRNA' : 50 occurrences

The term 'mRNA' stands for 'RiboNucleic Acid messenger', a substance existing in the protein structure. The node 'mRNA' appears 50 times in the corpus.

'mRNA' is preceded by symbols like 'rKLK10' (13 instances), RGB-2 (2 instances) and 'α1-antitrypsin' (5 instances). It strongly collocates in position +1 with 'levels' (6 instances) and with 'level' (3 instances). It thus forms **2 noun NCs:** 'mRNA levels' (6 occurrences) and 'mRNA level' (3 occurrences).

- The main tense used with this item is **the simple past / passive voice as in:**

High level of RKLK10 mRNA **was found** in the;

The purity of the mRNA **was monitored** by agarose;

mRNA was examined ...; mRNA **was cloned** and sequenced;

In contrast, no RKLK10 mRNA **was detected** by this ;

While no detectable RKLK10 mRNA **was found** in; Expression of the RGB-2 mRNA **was verified** by Northern tissue distribution of the mRNA.

- **The Simple past tense / active voice:** (one instance)

We **used** mRNA from submandibular gland ...

- **The simple present tense** is also used with this item as in :

Active voice : It is evident that this mRNA **belongs t**o the kallikrein ;

as this mRNA **codes** for a protein;

Pattern of the RGB-2 mRNA **is** strikingly similar;

tissue distribution of the mRNA which **parallels** that of ;

- **Passive voice :** The precursor mRNA **is alternatively spliced.**

The RGB-2 mRNA **is expressed** at;

indicating that the α1-antitrypsin mRNA also **is synthesized** outside ;

The α1-antitrypsin mRNA levels **are found to** ;

α1-antitrypsin mRNA **can be detected** at (present / modal auxiliary)

24) Node 'GENE' : 47 occurrences

'Gene' appears 47 times in the corpus. It forms NCs of :

- **Two nouns** : 'gene family' (17 instances),
- **Three nouns :** 'tissue kallikrein gene' (15 instances),
- **Four nouns** : 'kallikrein gene family members' (2 instances) and
- **Five nouns** : 'rat tissue kallikrein gene family' (2 instances).

It is either put in final position as in 'tissue kallikrein gene' or in the middle of the NC as in 'kallikrein gene family members.'

The tenses used with the node 'gene' are :

- **The simple present tense / active voice:**

The hamster B2 AR gene which **contains** most of;

size of the Kallikrein gene family **varies** among different ;

The cAMP-binding site identified in the kallikrein gene **may serve as** a negative regulatory element of (present with a modal auxiliary)

The RGB-2 gene **has** at least one;

that the rat α1-antitrypsin gene **is** single-copy.

result **indicates** that RKLK10 gene expression is;

result suggests that RKLK10 belongs to the kallikrein gene family .

- **The simple present / passive voice:** Northern Blot and Southern Blot analysis showed that the RKLK10 gene **is expressed** at a high level in the;

Our results demonstrate that the rKLK10 gene **is expresed** at a high level in the;

result indicates that the RKLK10 gene expression **is negatively regulated** .

- **Simple past / Passive voice** : that gene **was labelled b**y nick-translation ;

The hamster B2 AR **was cloned** from a ;

and regulation of RKLK10 gene expression **were determined** by.

- **Present perfect / Passive voice :** expression of the α1-antitrypsin gene **has been well studied.**

25) <u>Node 'INTESTINAL' : 47 occurrences</u>

Although this is an adjective, and not a noun and although it does not form NCs, it was underlined by biologists as important in their field. This adjective modifies nouns as in 'intestinal secretion' (6 instances), 'intestinal wall' (6 instances), 'intestinal epithelium' (4

instances), 'intestinal mucosa' (4 instances), 'intestinal fluid' (3 instances) and 'intestinal segments' (2 instances).

- The tenses used with this node :

- **The simple past / passive voice :** Net intestinal fluid transport **was measured;**

  reflected as changes in intestinal weight **was continuously recorded ;**

  The myenteric plexus of intestinal segments **was subjected** to / The intestinal lumen **was filled** to minimize evaporation, the intestinal segment **was carefully covered.**

- Simple present tense / passive voice:

  of cholera toxin on intestinal secretion **is principally mediated;**

  The hypothesis that the intestinal secretory reflex **is conveyed** ;

  chamber technique using stripped intestinal preparation **cannot be used;**

- Simple past / active voice : whereas vasoactive intestinal polypeptide **elicited** a marked;

- A lot of **nominal groups** are used with the item 'intestinal' as in :

  through **protection of the intestinal mucosa** from the damage;

  cause **sensitization of the intestinal mucosa**;

  whereas **distention of the intestinal muscle layers** ;

  organization of the intestinal secretory reflex;

  Different **functions of the intestinal wall.**

26) Node 'KIDNEY' : 47 occurrences

'Kidney' is used 47 times in the corpus. The node 'kidney' has in its near environment –2 / +2 position the following most frequent collocates. It forms different Ncs of different length. It thus forms :

- **Two noun NCs** like 'kidney cDNAs' (4 instances) and 'kidney enzyme' (2 instances), 'kidney clones' (2 instances) and 'kidney prostate' (2 instances).
- **Three noun NCs** : rabbit kidney microsomes (2 instances); rabbit kidney cDNAs (2 instances)
- **Four noun NCs** : rabbit kidney cDNA library (4 instances).

The tenses used are : **The simple present tense / active voice:**

Rabbit kidney microsomes also **contain** a P-450 that catalyzes;

Each of the kidney cDNAs **encodes** an efficient;

The kidney enzyme **exhibits** relatively low w-hydrxylase activity.

All 3 kidney cDNAs **exhibit** long open-reading;

Sequences of the 3 kidney clones **diverge.**

- Simple present tense / Passive voice :

Rabbit kidney microsomes **are reported** to catalyze the w-hydroxylation.;

Kallikrein-related enzymes in the kidney **are thought** to play ;

- Simple past tense / Passive voice:

PCR product from rat kidney **was cloned into.**

27) <u>Node 'TEMPLATE' : 47 occurrences</u>

The node 'template' is used 47 times in the corpus. It collocates strongly with 'strand' and 'DNA strand' to form **two noun NCs** as in **'**template strand' (19 instances), and **three noun NCs** as in 'template DNA strand' (17 instances).

- **The simple past / passive voice** is the only tense used with this item :

20% excess of the template strand **was added**;

an adenosine in the template strand **was included;**

with adenine in the template strand **was included** as.

- Gerund and participles :

Study of human embryonic kidney 293 cells **transfected** with ;

A rabbit kidney cDNA library **constructed i**n ;

Rabbit kidney cDNAs **encoding** lauric acid (gerund)

28) <u>Node 'TISSUE' : 47 occurrences</u>

'Tissue' appears 47 times in the corpus. It collocates strongly with 'kallikrein' in position +1 to form 'tissue kallikrein' in 28 instances. It is used to form NCs of different length as :

- **Two noun NCs** : tissue kallikrein (28 instances)

- **Three noun NCs** : rat tissue kallikrein (5 instances) ; tissue kallikrein cDNA (4 instances)

- **Four noun NCs** : rat tissue kallikrein cDNA (3 instances)

The node 'tissue' is generally preceded by nouns forming nominal groups as in:

Structure and function of tissue kallikrein ;

The deletion of tissue kallikrein; The physiological function of tissue kallikrein

Specificity of tissue kallikrein and T-kininogenase; Interaction with tissue kallikrein; The best characterized member of this family is true tissue kallikrein (3 instances)

The tense used with the item 'tissue' is **the simple past** :

- **Active voice :**

Rat tissue kallikrein **was** detectable under ;

- **Passive voice :**

Three sections from every tissue sample **was evaluated** using ;

of rat submaxillary gland tissue **was homogenized** in 0.2 .

29) <u>Node 'BENZAMIDE' : 45 occurrences</u>

'Benzamide' appears 45 times in the corpus. 'Benzamide' collocates with 'derivatives' (38 instances), and it thus forms **two noun NCs**. It is often preceded by the adjective 'discriminant' to give 'discriminant benzamide derivatives' (16 instances).

- 'Benzamide' is also used in nominal groups as in :

Distribution of discriminant benzamide derivatives

Administration of discriminant benzamide derivatives

Unexpected **potentiation by** benzamide derivatives .

The tenses used with the item 'benzamide' are :

- **The simple present tense / Active voice :**

Benzamide derivatives **do not appear**;

Potentiation of APO by benzamide derivatives **seems to** reflect ;

Since discriminant benzamide derivatives **require r**elatively little;

Therefore,the discriminant benzamide derivatives **seem to** cross;

Potentiation seems that effective benzamide derivatives preferentially **block;**

Generally benzamide derivatives poorly **recognise.**

- **Simple present tense / Passive voice :**

Administration of discriminant benzamide derivatives **is followed**.

- **Simple past tense / Passive voice**

still found whether this benzamide derivative **was injected**.

- **Present perfect tense / Passive voice**

Disriminant benzamide derivatives **have been obtained.**

30) Node 'STRAND' : 45 occurrences

'Strand' is used 45 times in the corpus. In position -1, 'strand' collocates strongly with 'template', 'DNA' and 'primer'. It forms two noun NCs as in 'template strand'(19 instances), 'DNA strand' (16 instances), and 'primer strand' (4 instances).

In position +1, 'strand' is followed by 'was' (4 instances) thus showing the tense used with this node.

- **The simple past / Active voice:**

06-methylguanine in the template strand **was** roughly 4.7 times.

- **The Simple past / passive voice**

Excess of the template strand **was added;**

Adenosine in the template strand **was included** as;

adenine in the template strand **was included** as a.

31) Node 'PROBE' : 44 occurrences

The node 'probe' is used 44 times in the corpus. It forms NCs of different length as in the following examples :

- **Two noun NCs :** cDNA probe (8 instances); oligonucleotide probe (8 instances)

Hybridization probe (2 instances)

- **Three noun NCs** : Kallikrein cDNA probe (3 instances);

α1-antitrypsin cDNA probe (2 instances)

'Probe' is followed by two verbs in position + 1. These verbs are : 'showed' (4 instances) and 'was' (4 instances).

- **The simple Past / Active Voice**

The rat α1-antitrypsin cDNA probe **showed o**ne band;

Analysis using the cDNA probe **showed** that rat α1-antitrypsin;

with an RKLK10-specific oligonucleotide probe **showed** that its mRNA;

Using an RKLK10-specific oligonucleotide probe **showed** that the RKLK10.

- **The simple past / passive voice** :

The end-labelled oligonucleotide probe **was addded** to the;

Specificity of the RK10sp probe **was tested** against the;

When the RGB-2 cDNA **was used** as a hybridization probe in Northern Blot analysis.

- **'Probe' is also used in the corpus as a verb :**

The 0.8 kb EcoRI-Pst I fragment **was used to probe** a rat brain cDNA library;

in the hybridization solution from above **to probe** the filters overnight at 42°C;

The existence of these kidney w-hydroxylase clones will make it possible **to probe** the structure-function relationships among these closely related gene products.

32)Node 'BUFFER' : 39 occurrences

The node 'buffer' is used 39 times in the corpus.

A 'buffer' is a substance that maintains the hydrogen ion concentration of a solution when an acid or alkali is added (the Concise Oxford Dictionary, 1991).

It often collocates with 'phosphate' and 'wash as in: 'phosphate buffer' (7 instances) and 'wash buffer' (2 instances). As such 'buffer' forms **NCs of the kind 'Noun+Noun**.'

- 'Buffer' also collocates with 'loading' as in 'loading buffer' (3 instances).
- 'Buffer' is followed by a gerund 'containing (2 instances) + expressions of quantity:

buffer containing 3% bovine serum / buffer containing 5ug/ml concanavalin;

- It is also preceded by a present participle: The reaction was quenched into **denaturing gel loading buffer.**

33) Node 'SERUM : 36 occurrences

The node 'serum' is used 36 times in the corpus. It forms **two and three noun NCs.** It collocates with 'rat' in position –1, thus giving 'rat serum' (5 instances).

In position +1 and +2, 'serum' collocates mainly with :

- ‘total’ as in ‘serum total’ (8 instances);
- ‘cholesterol’ as in ‘serum cholesterol’ (5 instances);
- ‘cholesterol concentrations’ as in ‘serum cholesterol concentrations’ (3 instances).

‘Serum’ is also preceded by prepositions mainly ‘of’ and ‘in’ thus forming nominal groups instead of nominal compounds e.g.

produced significant **elevation of serum cholesterol;**

Elevation of serum total and LDL cholesterol;

observed **differences in serum cholesterol concentrations;**

Changes in serum total cholesterol .

The tense used is :

- **Simple past tense / passive voice :**

α1-antitrypsin **was purified** from serum to apparent homogeneity.

- **Present perfect / active voice :**

isolated rat α1-antitrypsin from serum and **have generated** antiserum.

34) <u>Node ‘LYSINE’ : 35 occurrences</u>

The node ‘Lysine’ does not form a nominal compound but is salient in the field of biology.

It is followed by names of substances e.g. ‘leucine’ (12 instances) and ‘methionine’ (5 instances).

‘Lysine’ is mostly preceded by nominal groups such as: **high dietary level of** lysine, **combinations of lysine** and methionine, **high levels o**f lysine, t**he excessive amounts** of lysine, **high dietary levels** of lysine, **a high concentration of** lysine, enhanced **the effect of** lysine... , etc.)

The importance of the use of the preposition ‘of’ in scientific English lies in the fact that it introduces ‘**quantitative expressions’ (**amount**)** and ‘**measurable activities’** (level, concentration, etc.) (Gledhill, 1996).

35) <u>Node ‘PENILE’ : 34 occurrences</u>

‘Penile’ is used 34 times in the corpus. It is selected for its content-specific. It appears as an adjective in texts testing biology experiments. It is selected for study here because of its use in

the middle of NCs as in 'apomorphine-induced penile erection' whereby the form 'Noun + Participle + adjective + Noun.'

'Penile' is an adjective that precedes the word 'erection' (18 instances) and 'erections' (12 instances).

It collocates mainly with 'apomorphine-induced yawning and penile erection.'

The tenses used with the node 'penile' are :

- Simple past / Passive voice:

Effect on yawning and penile erection **was shifted;**

The yawns and penile erections **were counted** during antagonism by;

The number of yawns and of penile erection **was determined** during;

a significant decrease in the number of penile erections **was also observed;**

Any yawn or penile erection **was observed**.

- Simple present / Active voice :

sniffing whereas yawning and penile erection **disappear** ;

agonists that **induce** yawning, penile erection are observed;

doses **induce** yawning and penile erection ;

36) <u>Node 'SITES' : 34 occurrences</u>

'Sites' is used 34 times in the corpus. In position -1, it is preceded by the terms 'binding' as in 'binding sites' (10 instances) and 'glycosylation' as in 'glycosylation sites' (8 instances). These terms preceding the node 'sites' are themselves preceded by adjectives as 'potential' and 'N-linked', thus reading 'potential glycosylation sites' (3 instances) and 'N-linked glycosylation sites' (2 instances).

- It also forms **two noun NCs** as the following:

'receptor sites' (2 instances) and 'restriction sites' (2 instances).

The tenses used with the node 'sites' are :

- The Simple present tense / active voice :

It exerts strong affinity to 3H-dopamine binding sites, **stimulates** adenylate cyclase.

protein kinase A phosphorylation sites **have** a line above.

- The Simple present / passive voice :

The possibility that classes of dopaminergic binding sites **is discussed;**

Postulated N-linked glycosylation sites **are indicated** by asterisks.

Shared restriction sites **are indicated** above.

- The Simple past / Passive voice :

These glycosylation sites **were identified** by comparison;

Five potential glycosylation sites **were identified** on the .

37) Node 'CLONES' : 33 occurrences

'Clones' as a salient node appears 33 times in the corpus.

'Clones' collocates mainly with 'cDNA' (10 instances), and 'kidney' (2 instances) and thus forms NCs of the kind **'Noun + Noun' e**.g. 'cDNA clones' and 'kidney clones.'

It is also preceded by one adjective 'positive' (6 instances).

The node 'clones' is mainly followed by 'were' (9 instances) to form **the simple past in the passive voice**. Examples of these are:

Positive clones **were further plaque-purified.**

Positive clones **were identified by ELISA.**

Positive clones **were sequenced;**

Positive clones **were subjected to another ;**

Positive clones **were subjected to hybridization .**

Two positive clones **were isolated.**

- The present simple tense / active voice:

of the three kidney clones **diverge** in the 3' ;

Identification of cDNA clones that **hybridize** with.

- The nominals used with 'clones' are :

The analysis of kallikrein-like clones;

Screening of the kallikrein-related clones.

Identification of cDNA clones that hybridize with .

38) Node 'PATIENTS' : 33 occurrences

The node 'Patients' is used 33 times in the corpus. It was selected by biologists because 'patients' are an important element in biology as they represent the subjects to be studied and / or treated.

The node 'Patients' does not form any nominal compounds. It is, however, preceded in position -1 by nominal groups: '**The number of patients',** '**Treatment of patients** with Parkinson's disease', etc.

It is often followed in position +1 by the preposition 'with' (8 instances) : **Patients with** Parkinson's disease; The brains of **patients with** Alzheimer's disease; **Patients with** this symptom blood.

In position +1, 'Patients' is also followed by 'were' (4 instances) to form **the simple past / passive voice :**

Patients **were examined** once a week .

Patients **were kept i**n the hospital until termination of diarrhea.

- Gerund and participles:

20% of patients **taking** placebo (gerund)

The number of patients **included** in the study (participle).

39) Node 'PCR' (Polymerase Chain Reaction) : 33 occurrences

Pollack (1994: 102) explains the polymerase chain reaction as follows,

> *When a cell is dividing and all its descendants are dividing in turn, the total number of cells will double and redouble each time a cycle of division is completed. This is a chain reaction.*
>
> *The polymerase chain reaction 'PCR' forces the two strands of a fragment of DNA into a similar chain reaction of replication.*

'PCR' is used mainly in initial position. It is followed by right-hand collocates i.e. products (8 occurrences), primers (5 occurrences), product (5 occurrences), amplification (2 occurrences) and reactions (2 occurrences).

It is generally preceded by nominal groups: 'analysis of' (4 occurrences), 'transcription of' (4 occurrences), 'pair of' (2 occurrences).

The tense used with 'PCR' is the simple past in its both forms (active and passive).

- Simple past / active voice, e.g.

Reverse transcription and **PCR reactions, they all showed;**

Southern Blot Analysis of **PCR products did not detect ;**

- **Simple past / passive form, e.g.**

Reverse transcription and **PCR reactions were carried out** as;

The confirmed **PCR products were eluted from ;**

a pair of **PCR primers were designed;**

The purified **PCR products were then ligated ;**

The PCR primers were designed according to ;

a possibility that **the PCR product was generated** from;

One-tenth of **the PCR products was subjected to** Southern Blot anlysis;

After electrophoresis, **the PCR products were denatured, neutralized;**

One-tenth of the **PCR products were directly loaded.**

40) <u>Node 'rna' : 33 occurrences</u>

'rna' stands for 'ribonucleic acid.' It is important in the field of biology. It forms nominal compounds and collocates only with the left item 'liver' (2 instances). In most sentences it is preceded by measurement units i.e 100 ug, 15 ug of total rna.

- The tense used with this node is **the past simple in its passive voice, e.**g.

For Northern Blotting **RNA was denatured** using glyoxal.

After electrophoresis **RNA was blotted** onto ;

The fractionated **RNA was transferred** by capillary;

The poly(A) + **RNA was eluted** with;

total RNA was loaded onto.

The **RNA concentration was determined** by ..

The **RNA was denatured** in form -aldelhyde, separated by 15 ug of total rna ;

Total RNA was extracted from rat;

Total RNA was resolved in 1.5%.

41) Node 'ANIMALS' : 32 occurrences

The node 'animals' has been selected by biologists as an important item in biology because animals are the focal element in the experiments of animal biology, biochemistry and microbiology.

The node 'animals' does not form nominal compounds. It appears mainly after nominal groups such as :

Operation of the following animals was performed ; The first series of animals;

Behaviour in the same animals; Behaviours in the same animals;

poor growth of the same animals; growth of the animals.

- The tense used is **the past simple / passive voice:**

Animals **were preselected.**

The animals **were anaesthesized,**

The animals **were given** free access to food and water.

The animals **were housed** individually in galvanized cages.

The animals **were kept** in separate cages.

Since 24-33 animals **were simultaneously tested;**

Operation of the following animals **was performed** without ...

- There are two instances of **the Present simple / active form**

generated in **the same animals indicate** that ;

and licking that **the animals display** together were quantified.

In the above examples, the present tense is used to report about the results of the experiments.

The node 'animals' is also much used with **the past participle** as adjective, mainly with the verb 'feed': **animals fed** lysine; animals **fed the control diet; animals fed** various combinations.

42) Node 'PEPTIDE' : 31 occurrences

'Peptide' forms **two noun NCs** with both left-hand and right-hand collocates. It thus yields the following examples:

- Left-hand collocate : Signal peptide (5 occurrences); site for the signal peptide;

part of the signal peptide ; majority of the signal peptide.

- Right-hand collocates: peptide bond (3 occurrences);

Peptide N-terminal (2 occurrences) ; peptide fractions (2 occurrences)

- 'Peptide' is also used with long nominal compounds as in: calcitonin gene-related peptide (CGRP) antiserum (4 NCs); mature peptide N-terminal Glu residue (4 NCs).
- The tense used with 'peptide' is **the simple past / passive voice:**

Purified RPLC peptide fractions **were analyzed** on;

Peptide fractions **were speed-vacuum-dried**;

but this peptide **could not be detected**. (modal auxiliary)

43) Node 'SECRETORY' : 31 occurrences

'Secretory' although it is an adjective (and not a noun), it is dealt with here because of its importance in the field of biology as it was highlighted by the biologists, and also because of its position in the middle of nominal compounds e.g. chloride **secretory** process;
chloride **secretory** cells; chloride **secretory** response

- It is also sometimes used in the middle of long nominal compounds as in:

-cholera toxin-evoked **secretory** reflex

-cholera toxin-induced **secretory** reflex

- The present tense is used with the node 'secretory' :

The cholera toxin-evoked secretory reflex **involves** (Active form)

The intestinal secretory reflex **is conveyed** only ...(Passive form)

- There are two instances with the modal 'could':

The secretory pathway **could however, play** an important role in AD.

The secretory pathway **could increase** production.

44) Node 'CHOLESTEROL' : 30 occurrences

'Cholesterol' is a term important in Biochemistry. It forms nominal compounds with both left-hand and right-hand collocates.

- Left-hand collocates / **two noun NCs:**

LDL cholesterol (15 occurrences)

serum cholesterol (5 occurrences)

Plasma cholesterol (2 occurrences)

- These are illustrated in these examples : elevation of **LDL cholesterol;**

moderately elevated **plasma cholesterol**; the highest **serum cholesterol**

the elevation of **serum cholesterol.**

- Right-hand collocates / **two noun NCs:** 'cholesterol' collocates only with 'concentrations' and 'concentration' :

Cholesterol concentrations (8 occurrences)

Cholesterol concentration (2 occurrences)

The node 'cholesterol' is often preceded by nominal groups such as : **elevation of low concentrations of,** etc.

- The tense used with this node is mainly **the simple past.** The simple present tense is used only in few instances .

- The simple past :

LDL cholesterol was significantly elevated (**passive voice)**

The highest serum cholesterol concentrations did not coincide (**active voice)**

- The simple present

Elevation of LDL cholesterol is produced by feeding **... (passive voice)**

The plasma cholesterol concentration can be altered ... (**modal auxiliary 'can' in the Present simple / passive voice).**

45)<u>Node 'PRIMER' : 30 occurrences</u>

The node 'primer' collocates with only one item 'strand.' It forms a nominal compound of two nouns only: **primer strand** (4 occurrences).

- The tense used with this node is **the simple past / passive form**:

The amount of primer strand elongated **was measured** as described below (2 instances).

The elongation of the primer **was measured** at different times.

46) <u>Node 'ENZYMES' : 28 occurrences</u>

'Enzymes' is important in Biochemistry. It is the protein acting as a catalyst in a specific biochemical reaction (The Concise Oxford Dictionary, 1991)

As far as collocation is concerned, 'enzymes' has only left collocates, and its use in collocational frameworks is hence predicative as is shown below:

The cytochrome P-450 enzymes (3 0ccurrences)

Restriction enzymes (2 occurrences)

Kallikrein-like enzymes and Kallikrein-related enzyme ... (2 occurrences)

The tenses used:

- **The present simple / Active voice**

The cytochrome P-450 enzymes **that catalyze** w-hydroxylation;

These enzymes **show** very divergent substrate specificity.

- **The Present simple / Passive voice**

These enzymes **are encoded** by a large number of closely related and tandemly arranged genes.

Some residues in these enzymes **are known** to be critical for the cleavage preference .

These enzymes **are potentially involved** in a plethora of physiological processes.

- **The simple past / Passive voice**

The cytochrome P-450 enzymes **were cloned** from a rabbit kidney cDNA library.

Restriction enzymes **were obtained** from commercial ...

47) <u>Node 'LIVER' : 28 occurrences</u>

'Liver' is highlighted by biologists because this item is important in experiments. It thus often collocates in our corpus with the word 'rat.' It forms collocates with nouns on both sides (left and right). So, the synoptic profile gives us the following results:

- Left-hand collocates : rat liver (8 instances)
- Right-hand collocates: liver cDNA (3 occurrences); Liver RNA (2 occurrences);

Liver microsomes (2 occurrences); Liver expression (2 occurrences)

It forms nominal compounds of two, four and five nouns e.g.

-**Two noun NCs** : liver microsomes (2 occurrences) ;

liver expression (2 occurrences) ; rat liver (8 occurrences), etc.

- **Four noun NCs** : rat liver cDNA library (2 occurrences)
- **Five noun NCs** : rat liver expression cDNA library (2 occurrences).

The tense used **is the simple past (Active / Passive voice):**

RNA levels in the liver **increased** by 2-fold after induction by acute-phase inflammation. (active)

Liver microsomes **were isolated**. (passive)

48) <u>Node 'PROTEINS' : 28 occurrences</u>

'Proteins' are important in biology because they are the source of life. The Concise Oxford Dictionary defines a 'protein' as "*any of a group of organic compounds composed of one or more chains of amino acids and forming an essential part of all living organisms.*"

In biology, we study the structure, the formation, etc. of the proteins.

In our corpus it is also frequently mentioned (28 occurrences). Its position as concerns collocations is always predicative as it follows the item it collocates with. These are : animal (4 occurrences) and plant (3 occurrences) as in the following **two noun NCs**:

Cholesterolemic effect of **animal proteins;**

Higher levels in **animal proteins;**

Acids compared with **plant proteins.**

'Proteins' is followed by one main verb 'to be' with 'were' (3 instances), and 'are' (2 instances) in the passive form e.g.

- The simple Past / Passive voice

Contaminating proteins **were eluted** with buffer

The pelleted proteins **were resuspended** in a minimum volume of buffer.

The proteins **were transferred** onto intracellulose.

- The simple present / Passive voice

No immunoreactive proteins **are detected** with this antibody.

49) <u>Node 'RABBIT' : 28 occurrences</u>

The importance of the node 'Rabbit' in biology is the same as that of 'mouse' and 'rat.' Rabbits and mice are used for experiments.

'Rabbit' has only right-hand collocates in position +1 and +2. It is an illustrative example of nominal compounds. It, thus, forms the following compounds:

- **Two noun NCs :** rabbit kidney (9 instances)
rabbit lung (3 instances)
rabbit ileum (2 instances)
- **Three noun NCs**: rabbit kidney cDNAS (2 instances)
rabbit kidney microsomes (2 instances)
- Four noun Ncs : rabbit kidney cDNA library (4 instances)

The tenses used with the node 'rabbit' are **:**

The simple present in both forms (Active and Passive) e.g.

- The simple present / active form:

Rabbit kidney microsomes also **contain** a P-450 that **catalyzes** the w-hydroxylation.

Rabbit kidney and liver **express** lauric acid w-hydroxylases.

- The simple present / Passive voice

Rabbit kidney microsomes **are reported** to catalyze the w-hydroxylation

- The simple past / Passive voice

cDNAs encoding 3 cytochrome P-450 enzymes were cloned from a rabbit kidney cDNA library.

the same as that of a rabbit kidney protein that was immunoaffinity purified.

50) <u>Node 'FOOD' : 26 occurrences</u>

'Food' is used in the biology field mainly in experiments where animals (rabbits, mice, etc.) are given certain food while checking certain hypotheses.

In our corpus, 'Food' is used in initial position and has only right-hand collocates. These are: Food consumption (6 instances) ; Food withdrawal (2 instances).

It forms **two noun Ncs only** as the examples above.

'Food' is used with the simple past in both forms:

- Simple past / Active voice:

Similarly food consumption **was** the lowest in the group fed lysine and methionine.

Lost weight and their food consumption **was** significantly lower.

Weight gain and food consumption **were** similar in animals fed the control diet.

- Simple past / Passive voice:

The animals **were given** free access to food and water.

Weight changes and food consumption **were monitored**.

Bodies **were measured** following food withdrawal.

However, food intakes **were not quantified.**

51) Node 'NUCLEOTIDES' : 26 occurrences

In biochemistry, a 'nucleotide' is an organic compound consisting of a nucleotide linked to a phosphate group (The Oxford Concise Dictionary, 1991). A 'Nucleotide' is important in the protein structure. For instance, Thymine + RNA + Phosphate = nucleotide.

It does not form nominal compounds. It is mainly preceded by adjectives as in:

normal nucleotides (4 instances)

regular nucleotides (3 instances)

mismatched nucleotides (2 instances)

- For right-hand collocates (position +1), we can notice the verb 'to be' in the present and the past tenses as in these examples:

- The Simple present / Active and Passive voice:

The competition with normal nucleotides **are indistinguishable** from those (Active) and that nucleotides **are incorporated** opposite one... (Passive)

Normal thymine and cytosine nucleotides **are replaced** by their respective phosphorothioate nucleotide analogues (passive)

- The Past tense / passive voice

The results with single nucleotides **were not taken** in total because…

52) Node 'ANTAGONIST' : 25 occurrences

'Antagonist' is any product that can inhibit any function.

An 'antagonist' "*is any substance or organ that partially or completely opposes the action of another.*" (The Oxford concise Dictionary).

In the synoptic profile, 'Antagonist' shows the following left collocates : DA (abbreviated form of 'Dopamine') (9 occurrences); Discriminant (4 occurrences);

Different (3 occurrences) ; Dopamine (2 occurrences).

Thus, 'antagonist' is preceded by one noun 'dopamine' (or 'DA') and two different adjectives ('discriminant' and 'different').

- As right-hand collocates, antagonist shows the following:

Antagonist property (8 occurrences); Antagonist efficacy (3 occurrences)

Antagonist efficacies (2 occurrences); Antagonist properties (2 occurrences)

'Antagonist' finally forms nominal compounds of the kind:

- **Two noun Ncs:** Antagonist efficacies (2 occurrences);

Dopamine Antagonist (2 occurrences) ;

- **Three noun Ncs** preceded by the adjective 'discriminant':

Discriminant DA antagonist property (4 occurrences)

Discriminant DA antagonist properties (2 occurrences)

The tenses used with this node are:

- **The simple past tense / Active voice**:

DA antagonist **was a**lso more effective in potentiating APO;

The intraventricular route **made** a discriminant antagonist property appear for haloperidol;

a dissociated antagonist efficacy **appeared f**or haloperidol;

the antagonist which **appeared** after its intraventricular administration **was** an artefact

- **The simple past / passive form:**

The aims of the present study were to check whether the discriminant antagonist property **was found** at different times ...

- **The conditional tense / Passive voice**

The discriminant DA antagonist **would be more marked.**

53) <u>Node 'BLOOD' : 25 occurrences</u>

'Blood' is an essential element in all biological experiments.

The lexis used in the near environment of 'blood' deals mainly with blood measurement e.g. Fall in blood pressure, improved blood pressure, the unwanted drop of blood pressure, measurement of blood pressure, the regulation of systemic blood pressure.

The node 'blood' has only right collocates :

-**Two noun NCs:** blood pressure (14 instances)
blood samples (02 instances)
blood flow (02 instances)

-**Three noun NCs:** blood pressure control
blood presure regulation

The tenses used are:

- **The simple past / Passive voice:**

Blood samples **were taken** from;

Bodies **were analyzed in blood samples** collected;

Blood pressure **was** frequently **noted;**

However, blood pressure **did not change** (Simple past /Active)

- **Simple present tense / Active voice**

At doses 2 mg/day blood pressure **increases** slightly ...

- **The present perfect / Passive voice**

Blood pressure **has been measured.**

54) <u>Node 'CLONE' : 25 occurrences</u>

In our corpus, the node 'clone' appears as a single noun e.g. **The clone** containing the longest insert.

'Clone' is also preceded by adjectives i.e. a single clone, the positive clone, a differentiated clone, an independent clone, the longest clone, etc...

From the synoptic profile, we can notice that 'clone' has left-hand collocates only and it is thus predicative.

'Clone' forms a **two noun NC** with 'cDNA': 'cDNA clone' (7 occurrences)

It also forms **three noun nominal compound**: lung 37b clone (3 occurrences)

The tense which is mostly used with 'clone' is the simple present tense in both forms:

- **The simple present / Active voice:**

Sequence of this cDNA clone **is** 1306 base pairs.; The cDNA clone **is** seven base pairs; Sequence from the cDNA clone **matches** completely with ;

The cDNA clone thus **covers** the complete.

-The simple present / Passive voice:

Sequence derived from lung clone 37b **is indicated** by (>) .

Sequence of the longest clone KDB3 **is dislpayed** in .

- The simple Past / Passive voice:

The cDNA clone **was established** by matching....

Coding region of this clone **was used** at 107....

In some concordance lines 'clone' collocates to the right with 'base pairs' in position +2.

55) <u>Node 'DEHYDRATION' : 25 occurrences</u>

'Dehydration' is treated in our analysis, although it does not form nominal compound, because it was underlined by the biologists as being relevant to the field of biology. This, at least, in relation to the diseases of cholera and diarrhea.

'Dehydration' is mainly used as one word as in e.g. 'Prevention and management of dehydration', 'the clinical signs of dehydration', 'status and degree of dehydration', 'to the prevention of diarrhea, vomiting and dehydration', etc.

In general 'Dehydration' is preceded by adjectives and it is followed by verbs in the past simple / passive form.

The adjectives that most often precede 'dehydration' are:

severe dehydration (7 occurrences)

mild dehydration (6 occurrences)

moderate dehydration(2 occurrences)

The adjective **'**mild' is preceded by the words 'correction' or 'correcting':

The correction of mild dehydration

In correcting mild dehydration and in maintaining hydration

The adjective 'moderate' is preceded by 'treatment' or 'correcting'

The treatment of moderate dehydration

Correcting mild to moderate dehydration.

As for the verbs, the following were often mentioned:

Mild dehydration **was observed**, and only; severe dehydration **was referred** immediately to; Mild, moderate or severe dehydration **were included** in …

of moderate or severe dehydration **were observed**.

56) Node 'FEEDING' : 25 occurrences

'Feeding' represents the same case as 'Dehydration' in that they both do not form nominal compounds, but are used in the study for their importance in biology.

'Feeding' is important in biology experiments where animals are fed certain diets to check certain hypotheses. It thus often collocates on the right with 'diet(s)' as in :

Feeding certain amino acid diets

Feeding cholesterol-free amino acid diets

Feeding the same diet containing....

We investigated whether feeding diets high in lysine.

Weight gains were associated with feeding unbalanced amino acid diets.

It is preceded by verbs in the passive voice together with the preposition 'by' (7 occurrences).

- The simple present tense / Passive voice

In rabbits, elevation of LDL Cholesterol **is produced by feeding** a cholesterol-free ...

In rabbits, serum and LDL cholesterol Concentrations **are elevated by feeding** low fat, cholesterol-free, semi-purified diet containing ...

- The simple past tense / Passive voice

a marked elevation of serum total and LDL cholesterol concentrations **was produced by feeding** cholesterol-free amino acid diets.

gains **were associated with feeding** unbalanced amino acid diets.

57) Node 'NEURONS' : 25 occurrences

A neuron is a specialised cell transmitting nerve impulses.

'Neurons' in the synoptic profile shows left-hand collocates that are mainly adjectives 'afferent neurons' (5 occurrences), 'submucosal neurons' (3 occurrences), and one noun 'dopamine neurons' (2 occurrences), sometimes in its abbreviated form 'DA neurons' (2 occurrences as well.)

It is followed by the simple present either active or passive.

- The Simple present / active voice

Mucosal afferent neurons only **make** direct contact.

neurons that **lack** synaptic input.

90% of the submucosal neurons **respond** to acetylcholine.

- The present Tense / passive voice

nature of the afferent neurons **is known**, as is;

We know that cholinergic neurons **are involved** in the ...

These neurons **have also been shown** (present perfect)

- One modal auxiliary in the passive voice is used for stating a hypothesis in:

have considered whether **dopamine neurons could be the target of** neurotensin activity.

58) <u>Node 'PLASMA' : 25 occurrences</u>

'Plasma' is a key word in biology as it deals with blood, one key element in biology experiments.

'Plasma' is defined by Longman Dictionary of Scientific Usage (1979) as *"the clear, water-like colourless liquid of blood and other body liquids."*

It has right-hand collocates. In position +1, it collocates with 'ketone' (5 occurrences), 'concentrations' (4 occurrences), 'desorption' (4 occurrences), 'cholesterol' (2 occurrences).

In position +2, we have: 'bodies' (5 occurrences), 'mass' (3 occurrences). We get then 'plasma' forming nominal compounds of:

-**Two noun NCs:** plasma concentrations (4 occurrences)

plasma cholesterol (2 occurrences)

-**Three noun NCs:** plasma ketone bodies (5 occurrences)

plasma desorption mass (3 occurrences)

'Plasma' is usually followed by verbs in **the simple past / passive voice** as in:

Plasma ketone bodies **were anayzed**.

Fresh plasma **was used** for determination of acetoacetate.

'Plasma' is often preceded by nominal groups as in :

Analysis of plasma ketone bodies.

Elevation of plasma ketone bodies.

Substantial changes in plasma levels of ketone bodies.

59) Node 'RESIDUE' : 25 occurrences

In chemistry, a 'residuum' is a substance left after combustion or evaporation (The Concise Oxford Dictionary, 1991).

'Residue' has left-hand collocates. In position –2, it has 'N-terminal' (2 occurrences) and the demonstrative pronoun 'this' (2 occurrences).

In position –1, it mainly collocates with 'Glu' (gluten) (3 occurrences), 'Gly' (glyceride) (3 occurrences) and 'Phe' (phenol) (3 occurrences).

'Residue' is used with **the simple present tense / Active voice**

This gly residue **serves** as the hinge residue for the stressed loop structure.

There **exists** a gly residue surrounded by the sequence Glu-x-Gly

The solid circle **indicates** the hinge= gly residue, and the solid triangle...

This Gly residue **serves** as the hinge residue for the stressed loop structure

- The simple Present and present perfect / Passive voice

A Phe residue **is found** in all three sequences. (**simple present)**

This Phe residue **has been found** in all of the known serpins . (**present perfect)**

60) Node 'PRE-STEADY STATE' : 25 occurrences

'Pre-steady state' is frequently used in article titles and in figure captions.

It has right-hand collocates such as:

The occurrence of pre-steady state **burst** indicates that... (7 occurrences)

The presence of pre-steady state **incorporation** of thymine (4 occurrences);

pre-steady state **kinetics** (2 occurrences)

The tense used with this node is **the present simple / active voice**

The occurrence of this pre-steady state burst **indicates** that the ;

The presence of the pre-steady state burst **shows** that the ;

To answer this question, one **needs to** carry out pre-steady state kinetic studies.

The elongation during the pre-steady-state burst **is** largely that of the DNA to which the enzyme is initially bound.

61) Node 'DIETS' : 23 occurrences

'Diets' is one of the essential elements in biology experiments where animals (mice, rabbits, etc...) are fed certain diets to test certain hypotheses.

It has left-hand collocates such as: 'amino acid' (6 occurrences) i.e. 'Feeding unbalanced amino acid diets', '37.2% casein amino acid diets', 'of different amino acid diets.'

'Cholesterol' (2 occurrences) as in 'high levels in cholesterol-free diets', etc.

It is also preceded by adjectives like **'semipurified'** (describing the diets) and '**experimental**' (describing the function of the diets).

The tense used with this node is : **the simple past / passive voice:**

> They **were randomized**, **transferred** to semi-purified diets.
>
> After 3 weeks of experimental diets consumption, plasma ketone bodies **were analyzed.**

62) Node 'GENES' : 23 occurrences

'Genes' are important in biology, especially in the field of Genetics.

'Genes' are *"a unit of heredity composed of DNA or RNA and forming part of a chromosome etc. that determines a particular characteristic of an individual."* (The Oxford Concise Dictionary).

The node 'genes' collocates with 'kallikrein-like' (3 occurrences) as in the example:

> This approach has the potential to clone all other **kallikrein-like genes** expressed in the rat.

It also collocates with 'kallikrein' (2 occurrences) as in :

> **Kallikrein genes** share a high sequence ...

The tense used with 'genes' is mainly **the present tense:**

> Kallikrein genes **share** a high sequence (**Simple present / Active voice)**
>
> but their corresponding genes **have not been identified (present perfect/Passive voice).**

63) Node 'HYBRIDIZATION' : 22 occurrences

"*A hybrid (as a noun or adjective) means the offspring of two plants or animals of different species or varieties."* (The Concise Oxford Dictionary, 1991).

‘Hybridization’ as a noun forms left- and right-hand collocates in position -1 and +1.

On the left, ‘hybridization’ collocates with ‘Blot’ (3 occurrences) as in: Southern Blot hybridization; and ‘colony’ (3 occurrences) as in: Another cycle of colony hybridization with oligonucleotide probes.

On the right side, it collocates with the preposition ‘of’ (5 occurrences) forming thus nominal groups e.g. Southern Blot **hybridization of rat** genomic DNA.

‘conditions’ is the second right collocate with 4 occurrences, e.g.

Under high-stringency hybridization conditions, ...screened under low-stringency hybridization conditions..., 2P-labelled probe under these hybridization conditions.

‘probe’ (2 instances) as in : receptor gene as a hybridization probe;

was screened as a hybridization probe in Northern blot analysis.

The tense used is **the simple past / passive voice :**

Where the RGB-2 cDNA **was used** as a **hybridization probe** in Northern blot analysis of mRNA.

1000 white colonies **were screened** by **colony hybridization**, of which 84 colonies were recognized by rat tissue kallikrein cDNA.

64) <u>Node ‘ANTAGONISTS’ : 21 occurrences</u>

‘Antagonists’ in its singular form ‘ANTAGONIST’ has already been studied as ‘node 52.’

As its singular form, ‘Antagonists’ has the same left collocates i.e.dopamine (7 occurrences), DA (3 occurrences) together with apomorphine (3 occurrences) and dopaminergic (3 occurrences).

- It forms **two noun NCs**: Dopamine antagonists, apomorphine antagonists, etc.
- The right collocates of ‘antagonist’ are different from those of ‘antagonists.’

‘Antagonist’ has nouns as right collocates i.e. property, efficacy etc., whereas ‘Antagonists’ is followed by the conjunction ‘and’ (2 occurrences) and the preposition ‘of’ (2 occurrences). It is also followed by the gerund (‘belonging’ and ‘regarding’/ 2 occurrences each), examples:

Effects of various **dopamine antagonists regarding** climbing, sniffing, etc.

Curves of **various dopamine antagonists regarding** various stereotyped behaviours.

A series of **dopamine antagonists belonging** to various chemicals.

Other **dopaminergic antagonists of several stereotyped behaviours belonging** to various chemicals.

- The tense used with this node is mostly **the simple past/ Passive voice**

0.75 mg/kg apomorphine antagonists **were administered** i.c.v 20 min. before APO;

apomorphine antagonists **were injected** i.p. either 95 min.

65) <u>Node 'DUPLEX' : 21 occurrences</u>

'Duplex' was underlined as a main item by biologists. 'Duplex' means 'having two elements.'

- It collocates on the left with 'DNA' (7 occurrences).
- Its right-hand collocates are symbols represented in the letters (A, B, C, D), and the tense used with it is **the simple past/Passive voice**

Kexo for each **DNA duplex was measured** as described;

fragment and 0.1 uM **duplex A were incubated** with …

66) <u>Node 'REFLEX' : 21 occurrences</u>

'Reflex as a noun scores quite high in our corpus because it deals with the reflex of the animals in biology experiments.

'Reflex' can be both a noun and a modifier. In our corpus, it generally refers to *'an automatic response to the stimulation of a nerve.'*

- Its left collocates are : 'secretory' (11 occurrences) and 'axon' (2 occurrences).
- Its right collocates are: the preposition 'in' (4 occurrences);

It occurs with a verb in its passive form 'activated' (3 occurrences), and a noun 'arrangement' (2 occurrences) .

It forms nominal compound of **two nouns**: **axon reflex** (toxin is an axon reflex , ... arranged as an axon reflex in extrinsic neurons) and **reflex arrangement.**

- It also forms **long nominal compounds** as in :

The cholera toxin-induced secretory reflex

The cholera toxin-evoked secretory reflex

The tense used is t**he simple present tense :**

Active voice: Toxin **is** an axon reflex;

This observation **indicates (active voice**) that the secretory reflex activated by cholera toxin **is not arranged** as an axon reflex (**passive voice)**;

That the intestinal secretory reflex **is conveyed** only through the myenteric plexus (**Simple present/ Passive voice)**

67) <u>Node 'STEADY-STATE' : 21 occurrences</u>

'Steady-state' shows right-hand collocates. These are :

experiments	4 occurrences
incorporation	4 occurrences
conditions	2 occurrences
phase	2 occurrences

As they appear in the following examples :

The actual results of **steady-state experiments**;

taken with related **steady-state experiments**;

Studies conducted under **steady-state conditions**;

while the second apparently **steady-state phase** reflects the conversion.

It has one collocate common to 'the node 59' / 'pre-steady-state' which is: 'incorporation'of thymine and / or cytosine.

68) <u>Node 'ANTISERUM' : 20 occurrences</u>

According to the 'Concise Oxford Dictionary', an 'antiserum' is *"a blood serum containing antibodies against specific antigens, injected to treat or protect against specific diseases."*

'Antiserum**'** yields the following collocates in our corpus:

- Left: in position -2 , 'anti-rat' (2 occurrences) and in position -1 'α1-antitrypsin' (2 occurrences).

- It has one right-hand collocate 'against' with 4 occurrences,

Some instances of the use of 'antiserum'+ against:

and have generated **antiserum against** the purified protein;

A **rabbit antiserum against** AB 1-40 generously;

The antigen-overlay method using **antiserum against** purified α1-antitrypsin;

Blots were incubated with **antiserum against** α1-antitrypsin.

- 'Antiserum' forms nominal compounds of the following kinds:
- **Two noun NCs: 'rabbit antiserum** '; and '**rat antiserum**.'
- It also forms **long nominal compounds** :

phosphatase-coupled goat anti-rabbit antiserum...(**4 nouns)**

citonin gene-related peptide (CGRP) antiserum (**4 nouns)**

rabbit anti-rat α1-antitrypsin antiserum (**4 nouns**)

swine antirabbit immunoglobulin antiserum **(4 nouns)**

69) <u>Node 'CLONED' : 20 occurrences</u>

'Cloned' has been highlighted by biologists because of its relevance to the field of biology in general and, to that of genetics in particular.

'Cloned' is used in our corpus as a participle acting as adjective followed by nominal compounds i.e. a **cloned cell line**; a **cloned human colonic carcinoma cell**; a **cloned human intestinal cell line.**

- It is also used as a past participle in the passive form (without the auxiliary 'be') e.g.

rKlK10 cDNA **cloned** from the submandibular gland ...

rklk10 cDNAs **cloned** from the kidney and ....

- It is also used as a past participle thus illustrating the use of different tenses i.e. the simple past / and present perfect

The RGB-2 **was cloned** (**simple past / passive voice);**

We **have cloned** and **determined** the nucleotide sequence (**present perfect / Active**)

The RGB-2 **has been cloned (present perfect / Passive)**

70) <u>Node 'DRUG' : 20 occurrences</u>

'Drug' is a medicinal substance essential in treatments.

It collocates on the left with such nominal groups as: 'the concentration of drug in each area'; 'choice of a particular drug'; 'Diffusion gradient of the drug'; 'The clearance of the drug in each area'; etc.

- The node 'DRUG' forms **two noun NCs** and as such it has right-hand collocates as in the following examples :

'drug intake' (2 occurrences) / 'drug administration' / 'drug addiction' / 'drug regime' / 'drug treatments', etc.

71) <u>Node 'DRUGS' : 20 occurrences</u>

The node 'drugs' does not show any synoptic profile (unknown reasons for such a technical problem), but it does have concordance lines. The most important of these are :

... harmful anti-diarrheal drugs;

- The tense used with 'drugs' is **the present tense (present simple and present perfect**) as is shown by the highlighted verbs.

Parkinson's disease treated with usual anti-parkinsonian drugs **indicate** improvement of neurological symptoms;

The use of antibiotics and antimotility and antisecretory drugs **has been** ineffective...

In fact, increasing doses of these drugs **modulate** yawning in a biphasic manner.

The node 'DRUGS' does not form any nominal compound but it is much used with nominal groups as in: 'The use of drugs'; '...by a number of drugs'; 'The administration of drugs'; 'the apparent affinity of drugs'; 'inhibition by several drugs'.

72) <u>Node 'MICROSOMES' : 20 occurrences</u>

In the present corpus, 'Microsomes' form one main collocate as a **three noun nominal compound :** rabbit kidney microsomes (2 occurrences).

- The tense used with this collocate is **the present simple** e.g.

rabbit kidney microsomes also **contain** a P-450...(Active form)

rabbit kidney microsomes **are reported** to catalyze the w-hydroxylation **(PAssive form)**

'Microsomes', when used on its own (not as a nominal compound) and preceded by a preposition, it is often followed by the past participle acting as adjective:

The catalytic activity **of microsomes isolated** from COS-1;

The rates of each enzyme are simlar and are very high when compared to those **of microsomes prepared** from untransfected cells;

lauric acid **with microsomes prepared** from COS-1 cells;

acid was observed **for microsomes isolated** from transfected cells.

- **The past tense in its passive voice** is also used with this node:

Microsomes **were isolated** from COS-1 cells;

for 5 minutes the microsomes **were pelleted** from the supernatant by centrifugation.

Liver microsomes **were isolated** from a rat treated with ...

73) Node 'MOLECULAR' : 20 occurrences

'Molecular' as an adjective appears mainly in the titles of the articles under study i.e 'Department of Biochemistry and Molecular Biology.'

For the synoptic profile, it shows some collocates in the near environment. On the left, 'molecular' shows mainly adjectives such as 'calculated' (3 occurrences), 'same' (3 occurrences) and 'relative' (2 occurrences).

The right collocates are : 'weight' (10 instances) and 'mass' (4 instances)

It is often used in the middle of long nominal compounds as in:

Protein **molecular** weight markers;

- The tenses used with this node are: the simple past tense in the Active and passive voice:

Human PSA **was shown** to cleave **high molecular weight seminal vesicle protein;**

The RK10 protein **showed** the same **molecular mass** as deglycosylated k10 protein

- **The simple present tense / Active voice:**

Rat tissue kallikrein which specifically **cleaves low molecular weight kininogen** at two.

74) Node 'MOUSE' : 20 occurrences

'Mouse', like 'rat' and 'rabbit', is essential in biology experiments.

It is preceded in position -2 by 'human' (4 occurrences) and in position -1 by the conjunction 'and' (6 occurrences). It has mainly nouns as right collocates in position +1 'α1-antitrypsin'

(4 occurrences), 'counterparts' (2 occurrences) and 'fibroblast' (2 occurrences). In position +2, it collocates with 'cells' (2 occurrences).

- It forms nominal compounds of:

- **Two nouns** : mouse counterparts (2 Occurrences)
  mouse α1-antitrypsin (4 occurrences)
- **Three nouns:** mouse fibroblast cells (2 occurrences)

- 'Mouse' is used in Nominal compounds with more than three nouns:

  Ltk mouse fibroblast cell line; mouse epidermal growth factor-binding proteins.

75) <u>Node 'PHOSPHATE' : 20 occurrences</u>

'Phosphate' is a substance which is important in the structure of the DNA and RNA.

It collocates with 'potassium' and 'sodium' to form **two noun nominal compounds:**

'potassium phosphate' and 'sodium phosphate' (always preceded by units such as 100 mM, 50 mM, etc.) where the position of phosphate is predicative.

Phosphate also collocates with 'buffer' (5 instances) as in: 'phosphate buffer.'

76) <u>Node 'POLYMERASE' : 20 occurrences</u>

The noun 'polymerization' means 'having many parts.'

The synoptic profile of the noun 'polymerase' shows that it has only left-hand collocates which are distributed as follows. In position -2, it collocates with 'T7'(8 occurrences), 'coli' (3 occurrences), 'Taq' (2 occurrencse, and 'Change' (2 occurrences).

In position –1, it only collocates with DNA in 'DNA polymerase' (15 occurrences).

These give us the following structure :

T7 DNA polymerase (8 occurrences)
Taq DNA polymerase (2 occurrences)

- 'Polymerase' is mainly used in **the simple present / Active voice:**

  T7 DNA polymerase **involves** five kinetically discernable steps.

  Transcriptase and Taq DNA polymerase **lack** proof-reading activity.

77) <u>Node 'RABBITS': 20 occurrences</u>

'Rabbits' are used in biology for experiments. It does not form any nominal compound. Its synoptic profile shows just the preposition 'in' (9 occurrences) as a left-hand collocate in position -1. In position +1 'rabbits' is followed by the preposition 'by' (3 occurrences) and by the past participle 'fed' (3 occurrences). We can thus get the following examples:

hypercholesterolemia **in rabbits;**

Studies demonstrated that **in rabbits** a marked elevation of serum total and LDL cholesterol.

The elevation of LDL cholesterol concentrations **in rabbits** are presented in Table.

LDL cholesterol produced in rabbits by dietary casein or by amino acid mixtures.

Variability of **cholesterolemic responses in rabbits** fed increased amounts of selected essential amino acids (EAA)

The hypercholesterolemic response in rabbits was similar when casein was replaced with an equivalent level.

- The tense used with 'rabbits' is generally **the simple past** :

The hypercholesterolemic response in rabbits **was** similar when casein. **(Active voice)**

New Zealand white male rabbits **were injected** with 50-100 (2 occurrences) (**Passive voice)**

and correlations **were determined** for individual rabbits by regression analysis. **(Passive voice)**

- There is one instance of the use of **the future / passive voice:**

P-450 genes in rabbits **will be identified.**

78) <u>Node 'CASEIN' : 19 occurrences</u>

'Casein' is "*the main protein in milk, especially in a coagulated form as in cheese.*" (The Concise Oxford Dictionary, 1991).

'Casein' shows a strong bonding with the terms 'amino acid(s).' It thus collocates 10 times with these items in positions +1 and +2 respectively. It is generally preceded by percentage numbers or by the adjective 'essential' e.g. 'levels of all essential casein amino acids except arginine.'

It forms **nominal compounds of two nouns** (casein amino acid) and **three nouns** as in: 'casein amino acid diet' or 'mixture'. Some examples of these would be:

'... corresponding to a 45% casein amino acid diet' (4 occurrences);

'diet containing 30% casein amino acid mixture' (3 occurrences);

- **'Casein'** is mostly used in **the simple past / passive voice**:

acid mixtures derived from casein **was associated** with down-regulation of the hepatic LDL receptors.

The effect of a casein amino acid mixture, like that of casein, **was shown** to be dose related.

rabbits was similar when casein **was replaced** with an…

In mixtures 2-8 selected essential amino acids **were raised** to levels corresponding to a 45% casein amino acid diet.

79) <u>Node 'DISEASE' : 19 occurrences</u>

The node 'disease', in position -1 in our corpus, collocates quite strongly with 'Parkinson's' (11 occurrences) and 'Alzheimer's' (2 occurrences), thus giving 'Alzheimer's disease' and 'Parkinson's disease'.

In position –3, it collocates with 'patients'; and in position –2, it collocates with the preposition 'with', e.g. 'Patients with Parkinson's / Alzheimer's disease.'

As right collocates, in position +1, it gets the verb 'to be' **in the present tense,** third person singular **'is.'**

The node 'disease' does not form nominal compounds as such, but it is part of the possessive case 'Parkinson**'**s / Alzheimer's disease'.

- The tense used with this node is **the simple present tense in its two forms:**

Active voice :

Parkinson's disease **is** only moderate.

to patients with Parkinson's disease **shows** a significant...

Our studies with CI 201 678 in Parkinson's disease treated with usual anti-parkinsonian drugs **indicate** improvement of neurological symptoms.

Administration of the non-ergot compound CI 201-678 to patients with Parkinson's disease **shows** a significant and beneficial effect on global disability.

- Present simple / Passive voice :

of patients with Alzheimer's disease **is encoded** as ...

Alzheimer's disease **is characterized** by deposition

- Simple past tense / passive voice :

A rigid analogue of dopamine with a non-ergot structure, **was administered** to 24 patients with Parkinson's disease; 24 patients with Parkinson's disease **were treated**.

80) <u>Node 'N-terminal' : 19 occurrences</u>

In its synoptic profile 'N-terminal' is preceded and followed by nouns, adjectives, articles and prepositions.

In position -2 , 'N-terminal' is used with the article 'a' (2 occurrences), the adjective 'mature' (2 occurrences) and the preposition 'with' (2 occurrences).

In position -1**,** it is preceded by the article 'the' (6 occurrences), the nouns 'peptide' and 'residues' with 2 occurrences each.

In position +1, it is used with 'Glu' and 'amino' (4 occurrences each), together with the preposition 'from' (2 occurrences).

In position +2, it mainly collocates with 'acid' (4 occurrences), 'amino' (4 occurrences) and 'residue' (2 occurrences).

- 'N-terminal' forms nominal compounds of :

Two nouns: N-terminal Glu; residues N-terminal; N-terminal sequence.

Three nouns : The Ala residue N-terminal; N-terminal 26 amino acid sequence; N-terminal 33 amino acid residues...

Four nouns: The mature peptide N-terminal Glu residue**;** N-terminal amino acid sequence analysis.

- The tense used is the simple present / Active voice :

which **has** an N-terminal Glu;

N-terminal amino acid sequence analysis **reveals** a single N-terminal Glu;

Although the N-terminal amino acid sequence of ... **differs** from a partial N-terminal 26 amino acid sequence; The translated mature protein sequence from the cDnA clone **matches** completely with the N-terminal 33 amino acids.

81) Node 'POLYMERIZATION' : 18 occurrences

'Polymerization' forms nominal groups on the left side of the synoptic profile as in '**The rate of** polymerization' (5 occurrences).

It is also preceded by 'change after' (2 occurrences) as in 'the conformational **change after** polymerisation.'

It has one main right-hand collocate which is 'step' (4 occurrences), and it thus forms one kind of nominal compound i.e **plymerisation step** (4 occurrences) as in the following examples :

The ratio of the flux through the **polymerisation step**;

The change in the rate constant for the **polymerisation step;**

The rate of **polymerisation step** is the same in both cases .

Integrated flux through the **polymerisation step** over that period...

- The tense used is **the simple present tense (active and passive voice)**:

The ratio of the flux through the polymerisation step **falls** to only 37 within;

The curve **is strongly influenced** by the rate of polymerisation.

82) Node 'MUCOSAL' : 17 occurrences

'Mucosal' is an adjective that mostly collocates with the term 'addition' (3 occurrences) as in 'mucosal addition of cholera toxin.' It is also followed by 'nerves', 'afferent neurons', 'blood flow', and 'the serosal surfaces', as in the example: 'access to both the mucosal and the serosal surfaces.'

As left-hand collocates it is mainly preceded by units like '0.5 mM mucosal phlorizin after glucose; 10-5 M mucosal amiloride had no effect, etc.'

The simple past / active and passive voice is used with the node 'mucosal' :

Mucosal addition of 0.5 ug/ml amphotericin B also **enhanced t**he Isc. (Active voice)

Cells **were treated** by the mucosal addition of cholera toxin (Passive voice)

83) Node 'SUBMUCOSAL' : 17 occurrences

The synoptic profile of the node 'Submucosal' shows different collocates from that of the node 'mucosal.'

‘Submucosal’ has two right-hand collocates. These are ‘plexus’ (12 occurrences) and ‘neurons’ (3 occurrences). ‘Neurons’ always appears with ‘plexus’ (in 3 instances) as shown in these examples:

Myenteric plexus controls the submucosal neurons;

Sensory neurons in the submucosal plexus;

Secretory neurons in the submucosal plexus.

‘Submucosal’ is used in **‘the passive voice’ of the simple present’** and **‘the simple past tenses**: All ganglionic transmission **is localized** to the submucosal plexus (present)

The function of the submucosal plexus **was not performed** (past)

84) Node ‘SUBSTRATE’ : 17 occurrences

A ‘substrate’ is defined in the ‘Concise Oxford Dictionary’ as “*the substance upon which an enzyme acts.”*

The synoptic profile of the node ‘substrate’ shows one right-hand collocate only i.e. ‘specificity’ (7 occurrences) and with which it forms a two noun nominal compound ‘substrate specificity’ as in these examples:

Key residues contributing to substrate specificity;

Important determinants for the substrate specificity of tissue kallikrein;

‘Substrate’ is used in **the simple present tense / active and passive voice:**

- Active voice :

Enzymes **show** very divergent substrate specificity at both primary and extended residues;

K10 protein **has** substrate specificity different from that of true tissue kallikrein.

- Passive voice:

key residues determining substrate specificity **are indicated by** solid squares.

Tissue kallikrein **are characterized** by limited substrate specificity .

85) Node ‘AGONIST’ : 16 occurrences

‘Agonist’ is used in biology to mean literally ‘something similar to something else.’

In our corpus, it often collocates on the left side with:

Dopamine (4 occurrences), DA (3 occurrences) and the letter 'B' (2 occurrences).

Its right-hand collocate is 'isoproterenol' (2 occurrences). This could be illustrated in the following examples:

Administration of **the B agonist isoproterenol** (2 occurrences)

Effect of **the DA agonist** mediated by D-4 sites.

Metabolism of **the DA agonist** resulting in increased brain levels;

Test-doses of **the dopamine agonist**;

Irresponsiveness to **the dopamine agonist** regarding the gnawing behaviour.

The simple past / passive voice is used :

Binding of [3H] spiperone **was inhibited** by the agonist dopamine.

86) <u>Node 'AGONISTS' : 16 occurrences</u>

Like 'AGONIST', 'AGONISTS' collocate mainly with 'dopamine' and its abbreviated form 'DA'.

Its left-hand collocates in position -1 are: 'dopamine' (9 occurrences), 'DA' (3 occurrences) and 'D2' (2 occurrences), e.g.

Mice treated with **various dopamine agonists**; a variety of **DA agonists** of different chemical structures; high doses of **D2 agonists**, etc.

In position -2 'agonists' collocates with adjectives like: 'various' (3 occurrences), and 'indirect' (2 occurrences) e.g.

Doses of **direct dopamine agonists**; by either **indirect dopamine agonists**; mice receiving **various dopamine agonists; direct and indirect dopamine agonists**; mice receiving **other dopamine agonists** in subthreshold dosages, etc.

'Agonists' is used with the present **simple tense/active voice** as in :

Within the same range of doses of D2 agonists that **induce** yawning, penile erection, are observed .

Low doses of D2 DA agonists **stimulate** DA autoreceptors and …

- The nominals used with 'agonists' are:

The potentiation of dopamine agonists; Behaviours elicited by dopamine agonists in mice; The injection of dopamine agonists.

87) Node 'CLONING': 16 occurrences

'Cloning' left collocates show two items:

PCR-based (4 occurrences) and the preposition 'For' (2 occurrences).

'Cloning' is used in our corpus as a noun as well as a verb.

When used with 'PCR-based', cloning is a noun, e.g. '**PCR-based cloning'**;

When it is used with 'for', cloning is a verb, e.g. a useful technique **for cloning** kallikrein-like genes.

This approach ... is also suitable **for cloning** other genes belonging to multiple gene family...

Its right-hand collocates in position +1 are the noun 'strategy' and 'method' with which it forms nominal compounds as in:

The PCR-based cloning method thus provided a useful technique for cloning kallikrein-like genes. / **The PCR-based cloning strategy.**.

- The simple past / active voice is used with this node:

The PCR-based cloning method thus **provided** a ...

The PCR-based cloning strategy **provided** an efficient and reliable way to potentially identify..

88) Node 'ENCODING' : 16 occurrences

'Encoding', with 16 occurrences in our corpus, is used as a participle / gerund. In position -1, it collocates with cDNA (4 occurrences), cDNAs (2 occurrences), clones (2 occurrences) and sequence (2 occurrences).

On the right, in position +1, it is followed by the articles 'the' (3 occurrences) and 'a' (2 occurrences), as well as with nouns 'rat' (2 instances) and 'α1-antitrypsin' (2 instances). This could be shown in the following examples :

RKLK10 is **the cDNA encoding** protein K10, T-kininogenase;

a cDNA encoding T-kininogenase from rat submandibular;

a cDNA clone encoding rat α1-antitrypsin;

Isolation of **cDNA clones encoding rat α1-antitrypsin.**

Pairs comprising **the sequence encoding** a mature protein of...

The 5'primer corresponds to **the sequence encoding** signal peptide

From the examples above, we can notice that 'encoding' is often introduced in the middle of nominal compounds thus forming nominal compounds of the type mentioned in 'Chapter Four': NOUN + participle + NOUN e.g. cDNA encoding protein K10;

The sequence encoding signal peptide, etc.

or NOUN + NOUN + participle + NOUN e.g. cDNA clones encoding rat α1-antitrypsin

It is also often followed by names of substances such as:

T-kininogenase, lauric acid, enzymes, α1-antitrypsin, etc.

89) <u>Node 'HYPERCHOLESTEROLEMIC' : 16 occurrences</u>

'Hypercholesterolemic' with 16 occurrences in the corpus is used in cases where an increase of cholesterol is noted.

It is an adjective and as such it often precedes nouns. In position +1 these nouns are 'amino diets' (2 occurrences), 'properties' (2 occurrences) and 'response' (2 occurrences).

Concentrations of **hypercholesterolemic amino acids** and/or their break-down products.

Concentrations in several **hypercholesterolemic diets;**

The hypercholesterolemic response was similar;

In position -1, it is preceded by the adverb 'moderately' (2 occurrences) as in :

To have moderately hypercholesterolemic properties when fed at high concentrations .

'Hypercholesterolemic', as an adjective, is used to qualify nouns such as :

responses / properties / diets / amino acids/ and effect .

The tense used with this node is often **the simple past :**

Earlier studies **showed** that the hypercholesterolemic response in rabbits **was similar** when casein **was replaced;**

The hypercholesterolemic response **was** greater when…

90) Node 'SPECIFICITY' : 16 occurrences

As explained before, 'specificity' collocates with 'substrate' (node number 84) in 7 instances giving 'substrate specificity'.

'Cleavage' is another collocate that appears with 'specificity' in position -1 with 2 occurrences. 'Kallikrein' also is mentioned but in position -2.

'Specificity' forms nominal compounds with the nouns it collocates with as in the following:

- Two noun nominal compound:

Its distinct **cleavage specificity;**

Residues contributing to **substrate specificity;**

Determinants for the **substrate specificity** of tissue kallikrein;

- Four noun nominal compound :

The tissue kallikrein substrate specificity, etc.

'specificity' is used with **the simple present tense** :

- Active voice :

K10 protein **has** substrate specificity different from that of true tissue kallikrein ;

The enzymes **show** very divergent substrate specificity;

- Passive voice :

The tissue kallikrein substrate specificity **are conserved** in the protein encoded;

key residues determining substrate specificity **are indicated** by solid squares.

91) Node 'BODY' : 15 occurrences

From its synoptic profile, 'body' collocates strongly with 'weight(s)', and sometimes with the abbreviated form of this latter i.e. wt.

'Weight' has 7 occurrences, wt (5 occurrences) and 'weights' (2 occurrences).

As a whole, 'body' gives a collocational framework of either units mg/kg, ug/kg etc. preceding the collocation 'body weight', or in position -2 we get the noun 'changes' thus it gives 'changes in body weight(s)'.

'Body' followed by 'weight' forms nominal compounds of two nouns:

Body weight was maintained up to the end of the fourth day of treatment.

Changes in **body weight** showed significant inverse correlation.

50 mg/kg **body weight**;

was given intravenously at a dose of 10 mg/kg **body wt** after 3 hours in some experiments .

The node 'Body' is followed by **the simple past tense:**

- Active voice:

Changes in body weight **showed** significant inverse correlation .

Our results suggest that observed changes in weights **were** not responsible for observed differences in serum cholesterol concentrations.

- Passive voice :

Body weight **was maintained up** to the end of the fourth day of treatment.

92) <u>Node 'ELECTROPHORESIS : 15 occurrences</u>

'Electrophoresis' is a term used mainly in chemistry and physics. It is defined by the 'Concise Oxford Dictionary' as *"the movement of colloidal particles in a fluid under the influence of an electric field."*

'Electrophoresis' shows collocates only with the left-hand side in position -1 'gel' (7 occurrences) and in position –2 'agarose' (3 occurrences).

It forms nominal compounds of two nouns 'gel electrophoresis' (7 instances), and of three nouns 'agarose gel electrophoresis' (3 instances).

Examples :

…was measured by **gel electrophoresis** as described above.

... monitored by **agarose gel electrophoresis**;

were hybridized with the rk10sp probe after **agarose gel electrophoresis**, etc.

- The tense used with 'electrophoresis' is t**he simple past / passive voice:**

After **electrophoresis, RNA was blotted onto...**

The total elongation **was measured by gel electrophoresis** as described above ...

... as determined **by agarose gel electrophoresis was subcloned** into the M 13 mp 19 sequencing vector.

93) Node 'FATTY' : 15 occurrences

In chemistry, *"fatty acids refer to any class of organic compounds..., especially those occurring as constituents of lipids."* (The Concise Oxford Dictionary).

'Fatty' shows a synoptic profile of the following.

In position -1, it collocates with the adjective 'free' (11 occurrences).

In position +1, it collocates with 'acids' (12 occurrences) and 'acid' (3 occurrences).

In position +2, it collocates with 'were' (5 occurrences) mainly to form the passive voice and with 'substrates' in two instances. This would yield :

free fatty acids	(11 occurrences)
fatty acid substrates	(2 occurrences)

'Fatty' precedes noun 'acids', e.g. fatty acids, or nominal compounds as in 'fatty acid substrate'.

When used with its collocates, 'fatty' is followed by instances of **the past simple /passive voice:**

Free fatty acids **were analyzed** only one hour after a meal.

Free fatty acids **were analyzed** postprandially one hour after a meal .

Free fatty acids **were log-transformed** before the analysis of 5 min. duration.

Free fatty acids **were measured** (2 occurrences)

94) Node 'RESPONSES' : 15 occurrences

The node 'Responses' has been considered as a main item in biology because, in experiments, animal responses are important for the scientist to determine / confirm some hypotheses. *"A 'response' is a feeling, movement or change caused by a stimulus or influence"* (The concise Oxford Dictionary).

'Responses' does not form nominal compounds, but it is preceded by adjectives.

In position –2, we have 'various' (3 occurrences), and in position -1 we have 'behavioural' (5 occurrences), e.g. 'for various behavioural responses studied' (3 occurrences).

In position +1, 'responses' collocate with 'were' (2 occurrences) which is used to form the passive voice of the simple past:

Growth rates and ketogenic responses **were not generally correlated** with hypercholesterolemia.

Hypercholesterolemic responses **were not related to** poor weight gains or weight losses.

95) Node 'BASAL: 14 occurrences

The node 'BASAL' is an adjective. The synoptic profile shows that 'basal' in our corpus has the following collocates:

In position -1, it has the definite article 'the' (4 occurrences) and the adjective 'comparable' (2 occurrences).

In position +1, it collocates with:

electrical	5 occurrences
ganglia	3 occurrences
tonic	2 occurrences

In position +2, it collocates with 'Parameters' (5 occurrences) and 'level' (2 occurrences).

We thus get the following examples:

Monolayers **with comparable basal electrical parameters** can be utilised to study ...

posterior **basal ganglia**, anterior **basal ganglia**;

Which restores **the basal tonic level** of DA transmission;

The stimulation of postsynaptic DA receptors beyond the **basal tonic level** …

While stimulation of unchanged **basal activity** of adenylade cyclase by dopamine is reduced.

96) Node 'FIBERS' : 14 occurrences

'Fibers' appears 14 times in the corpus. It strongly collocates with the word 'nerve' (6 instances) and forms a NC of the kind 'Noun+Noun' : nerve fibers.

Two adjectives are used to describe the node 'fibers.' These are:

'Afferent fibers' (2 instances) and 'immunoreactive fibers (2 instances).

In position +1, 'fibers' is followed by the preposition 'in' e.g.

The number of fluorescing **nerve fibers in** three sections from.

The tense that is used with 'FIBERS' is **the simple past / passive voice:**

The density of immunoreactive nerve fibers **was estimated** on a three plus scale .

A significant decrease in the number of nerve fibers **was observed** in the myenteric plexus.

97) Node 'GROWTH' : 14 occurrences

The item 'growth' is used 14 times in the corpus.

It forms **nominal compounds of two and three nouns**. In position -1, it collocates with 'nerve' 'nerve growth' (2 instances).

In position +1, it collocates with 'rates' and 'reduction', with two instances each e.g.

Differences in **growth rates** and ketogenic responses were not generally correlated.

Growth reduction induced by a selectively high dietary level of lysine...

It also forms **NCs of three nouns**: 'nerve growth factor' (2 instances),

'human growth hormone gene.'

'Growth' forms one instance of a long NC (more than three nouns):

mouse epidermal growth factor-binding proteins.

98) Node 'NERVE' : 14 occurrences

'Nerve' like 'growth' and 'fibers' is used 14 times in the corpus and collocates strongly with these terms. 'Nerve' forms nominal compounds of the kind 'Noun + Noun' as in :

'nerve fibers' (6 instances) and 'nerve growth' (2 instances).

It also forms a **three noun Nc :**e.g. rat tonin, a subunit of **nerve growth factor.**

There are also two instances of Nc of the kind 'Noun + Gerund +Noun':

nerve blocking agents.

- The tense used with this node is **the simple past / passive voice:**

The density of immunoreactive nerve fibers **was estimated** on a three plus scale.

The number of nerve fibers **was observed.**

The simple past / Active voice :

In contrast, a mixture of… **was** also moderately hypercholesterolemic... but **did not affect** growth rates.

but this **was** not concomittant with significant growth reduction.

99) Node 'OLIGONUCLEOTIDE' : 14 occurrences

The item 'oligonucleotide' appears 14 times in the corpus.

In position +1, this item collocates with 'probe' (8 instances), 'probes' (3 instances) and 'primers' (2 instances).

'Oligonucleotide' forms NCs of:

- **Two nouns**: oligonucleotide probe / oligonucleotide probes / oligonucleotide primers.

In position -1, it is mainly preceded by specific adjectives like 'rklk10-specific' (3 instances), 'S2-specific' (2 instances) and 'end-labeled' (2 instances).

The tenses used with the node 'oligonucleotide' are:

- **The simple past / Active voice:**

  Northern Blot Analysis with an rklk10-specific oligonucleotide probe **showed** that its mRNA level in the submandibular gland ...

  The clones which **hybridized** to the tissue kallikrein cDNA probe but not the PS-and S2-specifc oligonucleotide probes were grown in ...

- **The simple past / Passive voice:**

  The end-labeled oligonucleotide probe **was added** to the prehybridization solution and then incubated at 60°C overnight.

  The filter **was subjected** to hybridization with an end-labeled oligonucleotide probe.

  The PS- and S2-specific oligonucleotide primers **were synthesized** according to the sequences of Ex3PS ...

  S2-specific oligonucleotide probes **were grown** in 2ml of LB medium.

  DNA sequencing oligonucleotide primers **were synthesized** using the ABI Model 380B DNA synthesizer.

100) Node 'POSTSYNAPTIC': 14 occurrences

The item 'postsynaptic', although it is an adjective, it has been highlighted by the biologists and it is used 14 times in the corpus.

It is preceded in position -1 by the prepositions 'of' (4 instances) and 'on' (3 instances).

In position +1, it is followed by DA (abbreviation for 'dopamine', 3 instances) and 'dopamine' (in full letters) (3 instances).
In position +2, 'postsynaptic' collocates with 'receptors' whereby we get 'postsynaptic dopamine receptors', and it thus precedes a nominal compound 'dopamine receptors'.

The tenses used are :

- **The simple present tense / Active voice :**

  Therefore, at the start of administration, CI 201-678 **acts** preferentially at denervated postsynaptic dopamine receptors...

  Higher doses **stimulate p**ostsynaptic dopamine receptors.

- **The present perfect tense / Passive voice :**

  The abeorphine CI 201-678 ... **has been suggested t**o act on pre- and postsynaptic Da receptors.

  This **has been interpreted** as the involvement of DA receptors for low doses ...

- **The gerund :**

  On the contrary, high doses of D2 agonists, **by stimulating postsynaptic** DA receptors, would restore a high apparent level of the dopaminergic transmission.

- The nominals used with 'postsynaptic' are :

  Introduction of postsynaptic acting dopamine DA agonists;

  in **the stimulation o**f postsynaptic DA receptors beyond the basal tonic level.

5.6 Discussion and Results

In the preceding part (the Computational Analysis) a concordance-based study of the most salient biology terms (represented in the form of Ncs) and their collocates is given. We tried to look at which words combine with each of the 100 selected nodes.
The main objective of this study is to make students discover the meaning of words by making them "*aware of representative patterns of language use and the selected grammatical structures.*" (Thurstun and Candlin, 1998:270)
The analysis shows the lexis and the grammatical structures used in the near environment of the NCs as nodes.

In ‘Chapter Three’, through the students’ answers to the questionnaire, it was found out that NCs (because of their structure i.e. long strings of nouns) cause a real comprehension problem as students sometimes could not find out what refers to what, for “*word combination conveys an important load of information concentrated in ‘one nominal group’ as this determines efficiency in the subject field*.” (Thouvenin, 1996:33)

In ‘Chapter Four’, we tried to show aspects of complexity of the biology discourse represented in lexical density in Halliday’s terms, and how this is expressed. We then counted and categorised the different NCs used in the fifteen journal articles that form the corpus (see Tables in ‘Chapter Four’).

In Chapter Five / Section 5.5, NCs were studied in their near environment. We provided a representative sample of the use of the word in context. The words preceding and following the node were examined. The analysis implemented is quantitative, as it shows the number of occurrences of certain patterns with the node, and qualitative as it shows what items exactly collocate with the node.

The overall analysis based on discovering both biology lexis and the grammatical structures used with this lexis, yielded the following findings.

Out of the one hundred nodes underlined by the biologists,

- Seventy seven were nouns related to the biology field, e.g.

  amino acid,
  protein,
  sequence,
  dna, cDNA,
  thymine,
  dopamine,
  Nucleotide,
  Gene,
  peptide, etc.

- Ten were adjectives related to the biology field:

  human,
  intestinal
  secretory
  penile
  molecular
  mucosal
  submucosal,
  hypercholesterolemic,
  fatty,
  postsynaptic.

- Ten were general English nouns dealing with animals, parts of the body, etc. but were related to the biology field as some of these nouns were used in the experiments, or for treatment, etc.:

  Rat, Rabbit, Mouse, Animals;
  Kidney, Liver;
  Disease, Patients;
  Food, Diets.

- Three were gerunds:

  Feeding
  Cloning
  Encoding

This information is represented in a table below.

100 nodes				
Number	77	10	10	03
Word classes	Nouns related to the biology field	Adjectives related to the biology field	General English nouns	Gerunds related to the biology field

Table 5.1: Classification of the 100 nodes in word classes

The general English words were kept for analysis because:

i) They were underlined by the biologists as being important in biology;

ii) These items are used to form NCs perceived by students as impediments in their text comprehension for according to Halliday, and this is also true for Algerian undergraduate learners,

> R*eplacing clausal patterns by nominal ones induces loss of information, which itself sets in a gap between the author's ideas and the reader's reception of it. The discourse tends to rely much on metaphorical induction, introducing more imagination and apprehension in the understanding, rather than clear rational comprehension.* (Halliday, 1994:353)

We are aware that the undergraduate biology students are uninitiated as they have not joined the target discourse community yet.

Therefore, the analysis falls in two main parts:

5.6.1 a study of specific features of biology lexis (collocation); and

5.6.2 a study of the grammatical structures frequently used with the 100 nodes selected from the corpus (colligation).

5.6.1 The specific features of biology lexis

An exhaustive study of the biology corpus has been done with the aim to come up with those elements that would be specific to biology discourse. The main feature related to biology lexis is the way this latter is expressed.

Biology lexis from the present corpus is frequently introduced in the form of nominal compounds (a string of nouns, one noun after the other) where each noun has its specific function.

The nominal compound form of the 100 items under study varies from two nouns to six nouns in length, giving the following figures.

Form and Length of NCs	2N N + N	3N N+N+ N	4N N+N+ N+N	2N N+adj+ N	2N + symbol /Abbrev + N	5N N+N+ N+N+ N	3N N+adj+ N+N	6N N+N+ NN+N +N
Total	64	28	14	08	07	05	02	04

Table 5.2: Form and length of the NCs used in the 100 nodes

The table above shows that two-word lexical items were the most frequently-occurring multi-word units in this corpus. these were either:

- Two nouns (B of A, 64%, e.g. cell bodies, nerve fibers, etc.); or
- noun + adjective + noun (8%, e.g. rat genomic dna); or
- abbreviation / symbol + noun as opposed to their lengthy alternative (7%, e.g. ‘DA agonist’ for ‘Dopamine agonist’).

The frequency of occurrence of NCs, of three nouns and more, is quite representative in this analysis.

- Three noun NCs represent 28%;

- Four noun NCs represent 14%;
- Five noun NCs 5%; and
- Six noun NCs 04%. Although the latter does not show a high percentage, it is worth mentioning it here because as Willis, D.(1990:43) suggests while talking about complex noun phrase *"... even though research suggests that roughly one noun phrase in eight has this kind of multiple modification, a proportion of one in eight certainly justifies thorough pedagogic treatment."*

Two-word lexical items seem to be of no great difficulty to non-native students as most of these items appear to be biology terms that are frequently used (64%), and which the students are familiar with. The difficulty for both native and non-native students of English arises when the NC is formed of 3 nouns and more. As our learners are non-native speakers of English, and they are undergraduate university students still getting initiated to the scientific discourse, this proves doubly problematic because they have not acquired yet:

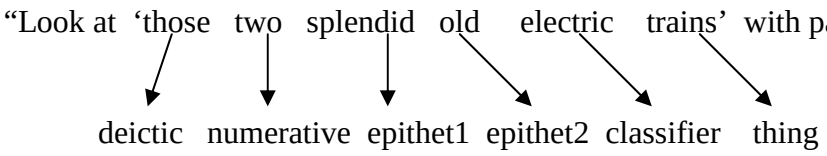
a) the technical skills
b) the terminology of the texts
c) the linguistic skills of native speakers (Thouvenin, 1996:8).

These three characteristics of language learning are sine qua non steps before acquisition of a '*minimum form syntactic information.*' (Ibid)

The difficulty in deciphering NCs efficiently lies in the fact that the learners ignore the different functions of the components of the NC.

A functional analysis of grammar describes a property of most nominal groups as being headed as is the case of our examples in this study.
All the words underlined by the biologists are nouns acting either as 'classifier' or as 'headwords' (Halliday, 1994:185) in the formation of the NCs as biology terms.

To explain what is meant by 'classifier' and 'headword' in a NC, we refer to Halliday's example spoken by a three-year old child:



where a ‘classifier’ can be either an ‘adjective’ or a ‘noun’, and a ‘headword’ has to be a noun.

The difficulty in such a combination rises from the inadequacy, new and casual in this context between the usual identification of the grammatical function with the semantic function of the noun. Indeed the noun performs a grammatical function as a classifier, i.e. adjective, which it is not used to do, and which learners consider as a transgression of a grammatical rule. Being trained to the foreign language through a strict prescriptive type of grammar, they feel inhibited to scrap these principles in a specialist’s context.
Let us see how dual functions are explained from a syntactic point of view.

5.6.1.1 Grammatical Functions of nouns in NCs

i) The noun as ‘head’:

According to Richards, Platt and Weber (1985:128), ‘head’ in grammar “*is the central part of a phrase. Other elements in the phrase are in some grammatical or semantic relationship to the head”*.

Richards et al. gave the following example for explanation. In the noun phrase: “*The fat lady in the floral dress”*, the noun ‘lady’ is the ‘head’ of the phrase. The other elements that Richards et al. refer to, are those items that appear in the near environment of the ‘head’. These elements are considered by some grammarians as ‘classifier’ or ‘modifier’.

Katamba’s description (1994:73) cited by Thouvenin (1996:10) is more wholesome. Katamba states that in English,

> *The head is normally the item on the right hand side of the compound. ...The syntactic properties are passed on to the entire compound. The syntactic head is also the semantic head*

of the compound. The non-head element in the compound specifies more narrowly some characteristic of the head.

These 'non-head elements' used to modify the headword are exactly the elements causing problems for non-native learners as their order 'what modifies what' is not easy to predict.
If we consider our corpus, in **Node 1**, the NC example 'purified rat α1-antitrypsin' (where 'purified' relates to 'α1-antitrypsin' rather than to 'rat') might cause a comprehension problem for our non-native students of English.

The other difficulty of NCs is the use of a string of nouns (2 nouns and more following each other). Nouns used in a NC can have different functions. The first one is mentioned above, a noun used as a 'headword' e.g. cell bodies, cDNA probe, rat α1-antitrypsin cDNA, ... where bodies, probe and cDNA are all 'headwords.'

ii) The noun as 'classifier:

The second function of a noun in a NC is that of a 'classifier' or 'modifier'.
In systemic linguistics, a 'classifier' is *"a word in a noun phrase which shows the sub-class to which a person or thing belongs."* (Richards, Platt and Weber, 1985:39).
A classifier can be either 'a noun' or 'an adjective' as is shown in the example by Richards et al. 1985,

classifier		**noun classified**
electric (adjective)	⟶	trains
steam (noun)	⟶	trains

In the corpus, we come across the example of : 'cDNA probe' where 'cDNA' is a 'classifier' and 'probe' is 'head.'
This point would lead us to discuss the position of the different nouns that are used to form the NCs that appear in our corpus.

5.6.1.2 Position and Grammatical Functions of Nouns used in NCs

(i) Position of nouns used in NCs

In the corpus, we can also notice another difficulty which consists of showing the selected node in different positions (initial, middle and final) playing thus different functions.

For example the node 'cDNA' plays different functions when used in different positions. 'cDNA' is sometimes used in **initial position** with the function of **a classifier** e.g. **cDNA** probe.

It is also used in **final position** as '**a headword'** in 'rat α1-antitrypsin **cDNA**', and as **a classifier in the middle of a NC** as in 'Rat brain **cDNA** library.'

The above explanation of the functions of nouns used in Ncs would lead us to another degree of difficulty which is the inflected forms of nouns (singular / plural).

(ii) Grammatical functions of nouns used in NCs

- Singular / Plural Relationships of Nouns used in NCs:

In the corpus list of "Frequent non-grammar" words, singular and plural forms of nouns were listed separately. Thus, certain nodes listed under both forms (singular / plural) show the following:

- Some examples show simply that a certain node is used in its singular or plural form as the case of **the node 85 'agonist'** and **the node 86 'agonists'** in the examples 'dopamine agonist' and 'Dopamine agonists.' The nodes 'agonist' and 'agonists' are both preceded by 'Dopamine' or its abbreviated form 'DA' e.g. 'Effect of the DA agonist...'; 'Test-doses of the dopamine agonist'; 'The potentiation of dopamine agonists'; 'The injection of dopamine agonists.'

The node 'agonists' (in the plural form) is always preceded by adjectives e.g. various / direct / indirect **dopamine agonists.**

The difficulty that arises in the singular / plural relationships of the nouns used to form the NCs is related to the function(s) of the nouns used to form the NCs. That is to say, if we consider the nodes 'cell / cells', 'antagonist / antagonists', 'receptor / receptors' as examples, we can come up with the following explanation.

- The function of the nodes ‘cells’ (137 occurrences), ‘antagonists’ (21 occurrences) and ‘receptors’ (58 occurrences) in the plural form are used as ‘headwords’: they are the final words of the NCs. Examples : Ganglion **cells**, receptor **cells**, and mouse Ltk **cells;** Dopamine / DA **antagonists,** apomorphine **antagonists**;
Dopamine / DA **receptors**, family **receptors** and NT **receptors.**

- The nodes ‘Cell’, ‘antagonist’ and ‘receptor’ used in their singular form, do not mean ‘one’, but their function in the corpus is that of a ‘classifier’ as in these examples: ‘cell line’ and ‘cell culture.’
In ‘cell lines’, ‘cell monolayers’ and ‘cell bodies’, ‘cell’ acts as a ‘classifier’. ‘line(s)’, ‘monolayers’ and ‘bodies’ act as ‘headwords.’

The same is true for ‘antagonist property(ies)’, ‘antagonist efficacy(ies)’ where ‘antagonist’ is ‘a classifier’ and ‘property(ies)’ and ‘efficacy(ies)’ are ‘headwords’; ‘receptor cells’ and ‘receptor sites’, where ‘receptor’ is used as ‘a classifier’ , and ‘cells’ and ‘sites’ as ‘headwords’.

- The nodes ‘drug / drugs’ :

The use of the nodes ‘drug / drugs’ is somewhat different. The use of the singular form of the node ‘drug’ is three-fold:

The word ‘drug’ is used in the corpus to mean:

a**. 'one’**: to show choice of a particular drug;

b. **the class ‘drug’** in its broadest sense - drug in general. This is mainly used in nominal groups e.g. The clearance of the drug in each area; The concentration of drug in each area ...

c. as part of a NC, **‘drug’** is used as **‘a classifier’** : drug intake; drug treatments; drug addiction, etc.

The plural form of the node ‘drugs’ is used as a ‘headword’ in nominal groups e.g. The use of drugs; inhibition by several drugs; The administration of drugs; etc.

The other striking features characterising the selected 100 nodes of the biology discourse corpus are abbreviated forms and acronyms of some terms.

In biology, besides compounding, further economy in the expression of lexical items is achieved through abbreviations and acronymy.

In the present research corpus, the use of acronyms, symbols and other abbreviations as the most obvious economy of expression is overwhelming.

5.6.1.3 Acronyms, Abbreviations and other Symbols

In biology discourse, abbreviations, e.g. 'DA' standing for 'Dopamine' and acronymy, e.g. 'DNA' standing for 'DeoxyRibonucleic Acid', are sometimes introduced without any prior mention of the full noun because they are assumed to be known by the readers.

However, Abbreviations and acronyms are sometimes introduced between brackets, after the noun has been introduced in full letters. This practice seems more convenient and helpful for our learners (biology undergraduates) for two reasons :

- They are still being initiated in the subject-speciality (they are not biologists yet); and
- They are non-native speakers of English and do not fully master yet these concepts.

i) Acronyms

Sager (1990) states that,

> *Special languages can also create new forms by various forms of compression of existing expression forms. The most common form is that of acronyms formed from initial letters of longer words.* (1990:76)

In biology journal articles, acronyms like DNA, RNA, etc. are rarely written in full letters because they are assumed to be known.

In the corpus under study, they are written in abbreviated forms because the corpus texts are composed of articles taken from scientific journals which address specialised readers (professionals).

ii) Abbreviations

Abbreviated forms, like acronyms, are also widely used in biology discourse. These are used either between brackets after the first mention of the noun, or are introduced without any prior mention of the full noun because they are assumed to be known.

The most used abbreviated forms in this corpus are :
'DA' agonists ('DA' for 'Dopamine'); 'Glu' (Gluten) residue; 'Gly' (Glyceride) residue; 'Ph.' (Phenol) residue.

According to Sager (1990:95) in biology

> *a name is not immutable in its form; it can be reduced or expanded according to its type and textual use. When the context is unambiguous, it is possible to abbreviate the generic attribution to species to the initial letter of the genus,* [e.g. Geranium ibericum = G.ibericum]

iii) Symbols and Greek Letters

In the biology corpus, we can also notice the use of symbols and Greek letters to form specific lexis. According to Sager (1990) "*ordering in regular patterns of designation was done for medicine and biology through the creation of a new quasi-artificial language based on Greek and Latin"* (Sager, 1990:239).
In biology, Greek letters such as 'α' and 'β' are used with substances e.g. 'α1-antitrypsin', 'β amyloid' , etc. The other special way of referring to biology terms is the use of symbols in the form of letters or numbers preceding a noun written in full letters such as 'Ltk cells, HRT-18 cells, COS-1 cells' (node 7), or preceding an acronym such as 'RKLK10 mRNA, RGB-2 mRNA' (node 23).

iv) Lower Case letters in front of Acronyms

In our corpus, we find the letters 'm' and 'c' respectively used with 'RNA' and 'DNA' ,e.g. 'mRNA' = RiboNucleic Acid messenger, and 'cDNA' = complementary DeoxyRibonucleic Acid. These letters (in lower case) seem to be used to describe and specify (to give more details) the acronym with which it is used (complementary, messenger, etc.)

Other characteristic features have also been noticed in the biology discourse corpus. These have been well-explained by Sager *et al.* (1980) when they state that,

> *The binomial code of biology provides a means of identifying uniquely thousands of species, subspecies and even varieties. It can be enlarged to indicate sources of designation, erroneous designations and their correction by the use of capital letters, abbreviations, brackets and other punctuation signs.*
> (Sager *et al.*, 1980:319)

v) Capital letters

As stated above by Sager et al., in biology the use of capital letters before nouns show "*erroneous designations and their correction by the use of capital letters*", e.g. N-terminal amino acid sequence, NT receptors ...

vi) Numbers in the middle of the NC

Numbers are used in the middle of the NC , mainly to specify it, e.g. the cytochrome P-450 enzymes, lung 37b clone, etc.

vii) Punctuation signs

Punctuation signs such as colons and hyphens are used with some terms mainly those dealing with chemistry, e.g. 06-methylg:T base pair; 06-MeG:C base pair, etc.

viii) Affixes and Prefixes

According to Sager *et al.* *"English has both analytic and synthetic means of word formation and these are used extensively in special languages."* (Sager *et al.*, 1980: 243).

For Sager *et al*. "*analytic means are the combination of independent lexical items into larger unit.". (Ibid).* In this research, this means the formation of NCs, the point already referred to in the preceding parts. "*Synthetic means are the modification of items by means of affixes*."

For Sager *et al.*, "*affixes are lexically meaningful items which cannot stand on their own and which sometimes have no meaning of their own but acquire one of several possible meanings in combination of a stem, e.g. -dom, -hood, etc*." (Sager *et al.*,1980: 243).

In English grammar, affixes can be added to words, and can thus change their meaning (kind / unkind), or their function (Kind as an adjective / kindness as a noun).

An affix can be added to the beginning of a word (un + kind) and this is called a '*prefix'*.

It can also be added to the end of a word (kind + ness) and this is called a '*suffix.'*

In our corpus, it is noticed that only prefixes are used. These are: 'sub / anti / hyper / pre / post', as in the following examples:

'sub + mucosal', 'ant i+ serum', 'hyper + cholesterolemic', 'post + synaptic', 'pre + steady', etc.

The use of affixes is worth mentioning in this research as they might be a source of difficulty for Algerian arabised undergraduate students who do not probably know their meaning.

For biology students in the Francophone biology department, these would not be of much difficulty because mainly the same prefixes (pre-, anti-, hyper-, post-, etc.) are used in French as well. Because biology uses a lot of Latin and Greek words, so these affixes are common in both languages : French and English.

Figure 5.3 sums up the different word classes used with the NCs in the biology corpus.

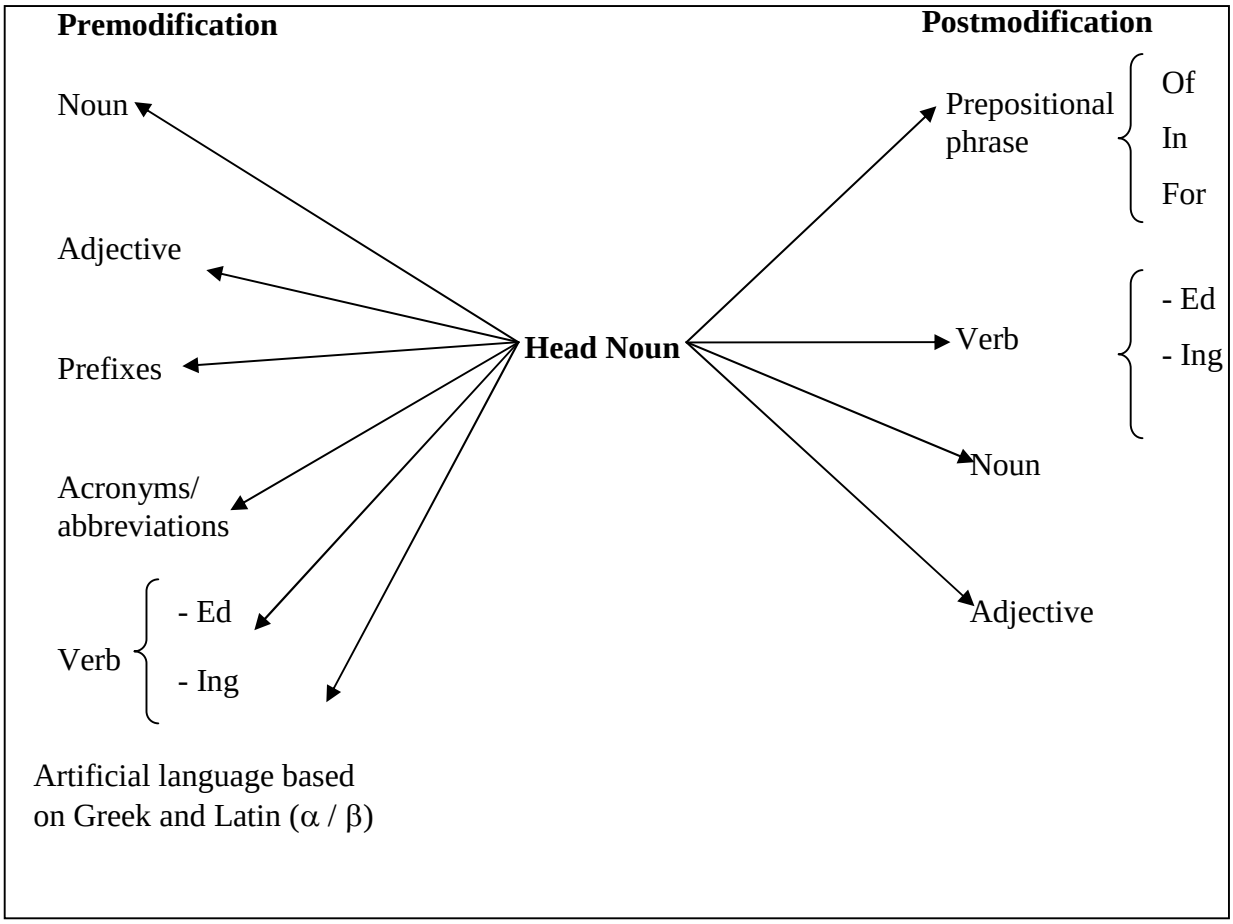


Figure 5.3 : An Expanded Pedagogic Model of NCS in Biology Discourse

(adapted from Latorre and Pons (1986) re-adapted from Quirk et al., 1979)

5.6.2 Analysis of the Grammatical Structures

The second step in the computational analysis was to identify the different grammatical structures that happen in the near environment of the node represented in NCs.

A concordance-based study was undertaken in order to view these contextualised grammatical structures.

The aim from such an approach was to find out which grammatical structures are used with the NCs as nodes (i.e. the grammar used in the near environment of the content-words as NCs), in order to develop later teaching materials focusing on the vocabulary and the

grammar that are used in the fifteen journal article corpus. Once the concordance lines were printed out, a categorisation and frequency count of the different structures were done. This gave the following results.

From the concordance line study of each node from the one hundred nodes selected, the following grammatical structures were found to be mostly used:

- **tenses,**
- **prepositions**,
- **specific adjectives**,
- **participials,** and to some extent
- **modals.**

A manual counting of these structures was done to find out their frequency. Each time one of the above items was used with the node (as a NC), it was ticked and counted. The numbers are given in percentages as they were calculated out of the 100 nodes under study.

5.6.2.1 Tenses

During recent years, English has increasingly become an important means for teaching and learning scientific subjects. To reach such an end, English language practitioners (sometimes together with a subject-specialist or a specialist informant) have often examined the different characteristics of modern scientific / technical English.

Different studies in ESP / EST have dealt with the most frequent use of tenses in different disciplines.

- Barber (1962) in a study entitled “Some Measurable Characteristics of Modern Scientific Prose” is a contribution in English syntax and phonology,
- Herbert, A.J., (1965) “The structure of Technical English”,
- Lackstrom, Selinker and Trimble (1972) “Grammar and Technical English”
- Tarone, Dwyer, Gillette and Icke (1981) “On the Use of the Passive in Two Astrophysics Journal papers”;

And the most recently published work as the revised and updated version of this last article,

- "On the Use of the passive and Active voice in Astrophysics Journal Papers: With Extensions to other Languages and other Fields" (1998) from the same authors Tarone *et al.*

The 1981 article and the 1998 updated article are almost the same except for the three additions:

a) the 1998 article title is different from the 1981 article in that Tarone *et al.* deal with both the Active and the Passive voice and they also added the mention: 'With extensions to other languages and other fields';

b) in the 1998 article the authors added a passage called 'The Rhetorical Shape of Research Articles' where they explain the difference in rhetorical shape between the field of logical arguments such as Astrophysics and Mathematics, and experimental research such as that in biology, etc.;

c) the last difference resides in the addition by the authors of a paragraph at the end of the 1998 article entitled 'Extensions of the Findings to Other languages and Other Fields' where the authors Tarone *et al.* find the same results (compare their 1981 findings with the findings of an astrophysics research paper (the same field and the same genre) written in Russian.

Other recent works dealing with tense study include those focusing mainly on the use of a substantial corpus (e.g. Bank of English corpus at COBUILD) together with computational tools (different computer programmes for frequency analysis, etc.) such as :

a) Williams, I.A., (1996) "A contextual Study of Lexical Verbs in Two Types of Medical Research Report: Clinical and Experimental."

and another work which will be of reference in this research is:

b) Hunston, S. and Francis, G. (1998) "Verbs Observed: A Corpus-driven Pedagogic Grammar."

From the above selected works, Verbs and tenses seem to be given enough importance in the field of teaching E.S.P., considering the number of publications.

In this research material (analysis of the 100 nodes), the following tenses were found to be frequently used .

Simple Present tense		**Simple past tense**		**Present perfect tense**	
Active	Passive	Active	Passive	Active	Passive
44%	27%	14%	56%	03%	11%

Table 5.3 : Tenses and voice used in the 100 nodes

(**Table 5.3** should read that the 'simple present, active voice' is used in 44 nodes out of the 100 nodes under study and this is used as a percentage i.e. 44%). As this table shows, the simple present tense ranks first with 71 occurrences (44 actives and 27 passives), followed by the simple past (70 occurrences (14 actives and 56 passives) and in last position we have the present perfect with 03 occurrences in the active voice and 11 occurrences in the passive voice.

The active voice yielded 61 occurrences (44 present, 14 past and 03 present perfect) and the passive voice 94 occurrences (27 present, 56 past and 11 present perfect).

The sum of these figures shows that the passive voice is widely used in the 100 nodes, as is the case with most scientific discourse.

This figure (94 occurrences / passive voice and 61 occurrences / active voice) is however different from the figure found in the analysis of the whole biology corpus in 'Chapter Four / Section 4.7.1.' which shows more use of the Active voice (1594 Active verb forms) than the passive voice (1433 Passive verb forms).

In this research, we aim at describing the main features of the biology discourse i.e. features which are frequently used so as to include them in the biology course as it has been noticed that tenses, although acquired early by science students, they are often

subsequently replaced with over-generalised verb forms, e.g. gave- gived; and that the passive voice is rarely found in the learners' oral or written production.

Our intention is not as much to offer explanations about why such things happen as to leave it to teachers (in their classrooms) and subject-specific practitioners to explain why such things do happen. Our aim is to base a teaching methodology on these features and conceive ways: tricks, techniques, rules or didactic principles for the learner to anticipate comprehension hurdles.

After having examined the use of tenses and voice, other word classes preceding and following the key words are examined.

5.6.2. 2 Prepositions

The other aspect that has been highlighted in the biology discourse is 'the use of prepositions' in twenty nodes.

Prepositions contribute heavily to the structure of the noun group, that is why they need to be brought to the attention of the students. They can either follow or precede the node as can be shown in these examples from the corpus:

*'Incorporation **of** thymine' as opposed to 'Incorporation f**or** thymine';*

*Cholera toxin **on** intestinal secretion*

*Cell monolayers **obtained on** Transwell ;*

*The human colon carcinoma cell line HT-29 **grown on;***

***through** protection of the intestinal mucosa **from** the damage;*

*Study of human embyonic kidney 293 cells t**ransfected with;***

*A rabbit kidney cDNA library **constructed in** ;*

***Interaction with** tissue kallikrein;*

*The reaction was **quenched into** denaturing;*

*Contaminating proteins were **eluted with** buffer.*

5.6.2.3 Modals

Modals (can, may and could) are used with 04 nodes. 'Could' is used (11) eleven times.

Examples from the corpus:

*α1-antitrypsin mRNA **can** be detected ...*

*The cAMP-binding site identified in the kallikrein gene **may** serve as a negative regulatory element;*

*but this peptide **could** not be detected ;*

*The secretory pathway **could,** however, play an important ;*

*The secretory pathway **could** increase production.*

5.6.2.4 Adjectives

'Adjectives' also appeared with noticeable frequency in the biology corpus and so need to be mentioned here.

Reference to the use of adjectives in the near environment of the 100 nodes has been made in the first part of 'Chapter Five'.

The position of adjectives with the different nodes often varies.

- The position of adjectives:

a) Adjectives are sometimes used attributively i.e. they precede the node as in these examples:

Lauric acid / **nucleic** acid / f**atty** acid

Dietary proteins / **direct** dopamine

Infected cells / **cultured** cells / **epithelial** cells.

b) Adjectives are also used in the middle of the NC as a node:

Rat **genomic** DNA / rat **submandibular** gland /

Protein **molecular** weight markers

apomorphine-**induced penile** erection

cholesterol-**free** diets / cholera **toxin-induced secretory** reflex.

c) Adjectives are sometimes used predicatively i.e. they follow the node as in : substrate specificity **different** from that of true tissue kallikrein.

5.6.2.5 Participials

Participials in both forms (-ed and –ing) are also quite used in the biology discourse.

a) **Past participles used as adjectives**, or what is sometimes called "*a lexical passive to highlight its adjectival properties."'* (Thouvenin, 1996:19), e.g.
kallikrein-**related** cDNA clones / a **cloned** cell line / **infected** cells
N-linked glycosylation sites / kallikrein-**related** enzyme
Cholera toxin-i**nduced** net fluid secretion .
The human colon carcinoma cell line HT-29 **grown** on.

b) **Present participial (-ing form) used as adjectives:**
Fluorescing nerve fibers / nerve-**blocking** agents /
loading buffer / **Binding** protein
PCR-based **cloning** strategy / **cDNA encoding** protein K10 ...

As a whole, the computer text analysis included fewer verbal participial premodifiers (i.e. –ing and -ed) than adjectival premodifiers (e.g. nucleic, secretory, molecular, epithelial, etc.). This is so in NC form because according to Quirk et al. (1979:904) *"Adjectival and nominal premodifiers tend to signify permanent characteristics; while temporary characteristics appear associated with –ing and –ed premodifiers."* (in Latorre and Pons, 1986:101). Biology lexis appears to lean heavily in favour of permanent characteristics.

The lexical and grammatical words used with the 100 nodes mentioned above are of such frequent occurrence, and so central to the meaning and structure of the language of biology, that they need to be brought to the attention of both teachers and students.

The most important feature of the discussion of the computational analysis results is that it suggests an order of presentation in class. The order for the first modifiers of NCs are nouns (very frequently used as shown in all the nodes) and adjectives. Participials (-ed and –ing) come in third, and are also important as they account for important premodifiers and postmodifiers. These items will be covered in the following chapter 'Pedagogical Implications'.

5.7 Conclusion

'Chapter Five' dealt with the computational analysis of 100 selected content words (used as nodes) from the biology corpus. The analysis using ATA programme showed the frequency of occurrence of these nodes together with the lexis and the grammatical structures that appear in the near environment.

A concordance-based approach was undertaken in order to view the specific features of biology lexis (showing its forms and functions) followed by the identification of the different grammatical structures.

The results of the manual analysis (in Chapter Four) are almost similar to those of the computational analysis in this chapter, in terms of lexis (frequency of occurrence of NCs, their forms and functions) and the grammatical structures, with one specific difference. The passive voice is found to be more widely used in the 100 nodes (with 94 occurrences as compared to the active voice with 61 occurrences only) than in the analysis of the whole biology corpus in Chapter Four / Section 4.6.1 where active verbs exceed passive verbs in most texts. The active voice is used in 09 texts out of 15 texts, with 1594 active verbs and 1433 passive verbs.

The use of the passive voice more than the active voice in 'Chapter Five' confirms, in a way, the fact that in science the passive voice is more frequent as we are not interested in 'who' does / did the action. The other difference between the manual analysis and the computational analysis is that the latter, through the concordance-based approach unveiled lexical and grammatical features that could not be done otherwise, e.g. affixes, acronyms and abbreviations, artificial language based on Greek letters (α / β), and other different symbols.

Chapter Six

Pedagogical Implications

CHAPTER SIX

PEDAGOGICAL IMPLICATIONS

6.1 Introduction

In 'Chapter Three', through the students' questionnaires and informal interviews, we found out that, in the English course (in Biology Department), reading is the most needed skill for students. The aim is to prepare them for the final year of study where reading biology literature in English is necessary for the writing-up project (mémoire). We also framed the kind of literature the students need to read i.e. journal articles. A description of this type of literature is given in order to identify the content.

In 'Chapters Four and Five', a corpus of biology journal articles has been selected and a thorough study (manual and computational) of this corpus has been implemented. The study shows some characteristic features of biology journal articles necessary for designing an appropriate teaching methodology for the target learners.

As the aim of this research work is to improve the English language course quality for undergraduate biology students, in this chapter, we shall firstly suggest solutions to the administrative and pedagogic problems (as perceived by students and teachers in 'Chapter Three'), and secondly present a teaching methodology together with some teaching tasks / activities that would stimulate the learning of English with a hoped reasonable efficiency.

6.2 Solutions to the administrative / organisational problems

In reference to the problems mentioned by the students in 'Chapter Three' in terms of the English course organisation, the following suggestions can be made for the four-year degree (D.E.S) currently in practice in biology, and for the LMD system, for the future (as it has been introduced only recently, 2005-2006):

- **English should be taught during the four years of biology degree (D.E.S) and the three years of the LMD programme**

English should be introduced from the first year of the biology degree to the fourth year to avoid the two-year void currently in practice for the D.E.S. programme, and from the first year to the third year of the LMD programme to avoid one year void.

- **The number of teaching hours per week**

In the biology Department, English is currently taught for one and a half hours (one session) in the third year during the two semesters, and one hour a week in the fourth year during the first semester only in the D.E.S programme. For the LMD programme, it is taught in the second and third years only. To relate this point (the number of teaching hours) to the first point (teach English throughout the years of biology study), we can consider the following timetable:

- Suggestions for the D.E.S. programme

First Year: one session of two hours during the two semesters;

Second Year: one session of two hours during the two semesters;

Third Year: two sessions of one and a half hours each (i.e. three hours a week) during the two semesters;

Fourth Year: one session of two hours during the first semester only, as in the second semester, students are busy preparing their projects (mémoires).

- Suggestions for the LMD programme

First Year: one session of two hours during the two semesters;

Second Year: two sessions of one and a half hours each (i.e. three hours a week) during the two semesters;

Third Year: one session of two hours during the first semester only, as in the second semester, students are busy preparing their projects.

In the D.E.U.A. (Diplome d'Etudes Universitaires Approfondies) programme, English is taught from the first year to the third year of instruction.

Rationale for the suggested timetable

The introduction of the teaching of English in the first and second years of the D.E.S programme, and the first year of the LMD programme would serve as :

- A revision of the different language notions and functions for the students who had already studied English at the secondary school;
- An initiation to the English language for those students who had studied other foreign languages i.e. German, Spanish, Italian, etc.;
- An introduction to the E.S.P. course (which will be developed in the third and fourth years) to prepare them for more specific teaching / learning points related to their subject-speciality.

At this level, in the first and second years, teachers of English can use 'E.S.P. concocted texts' (texts produced for teaching purposes, as opposed to 'authentic texts') to illustrate the different points (lexical or grammatical) of the language they have to teach.

- **Recruitment of full-time teachers of English**

It has been noticed that teachers of English in Science Institutes are generally recruited as part-time teachers only. The instability of the teaching staff may lead to lack of efficiency in language teaching.

- **The number of students per class (Group Size)**

In Science Institutes, English is generally taught (as a formal lecture) to the students of the whole year i.e. to 200 students put together in a large amphitheatre. Dealing with groups of this size transforms the 'English course' as a wholesale business, it becomes nothing but a matter of filling in a timetable to justify the 'needed mark' for the pass result by the end of the academic year. English then becomes a side-interest matter, and with such a casual status, in the students' mind it never leaves its 'tourist conceived content', that owes more to pre-conceived views (schemata) on foreign language learning.

The suggestions put forward above, to be implemented efficiently, would require much smaller groups of 30 to 40 students in order to give more emphasis to listening and speaking, since some of the learners will be involved in exposés and communications where the oral expression is fundamental. Such a pedagogical disposition requires administrative steps i.e. recruitment of a more important number of staff which itself asks for more budget, and everybody knows that financing in Higher Education suffers from this terrible paradox whereby, those who take administrative decisions are not trained for that purpose (academic objectives), and those who must take them are not trained for the management of educational institutions.

It is hoped that with the advent of the LMD system, the problem of group size would be obsolete as the teaching of the different components would be done in groups relatively smaller.

- **Provision of books, scientific journals and textbooks**

The biology library should be provided with more books, scientific journals and textbooks for learning English to help teachers prepare their teaching materials accordingly, and students to work independently of their teachers during their free-time.

It is hoped that by solving the administrative / organisational problems (time allocated to the English course, group size, recruitment of full-time teachers and provision of more books), we can help in improving pedagogically the quality of the English course through a suggested methodology.

Therefore, the following part will refer to the different types of syllabuses, approaches and methods that have been used in teaching / learning the English language. It will then deal with methodology.

6.3 Types of syllabuses in English language teaching / learning

Throughout the world, the English language teaching / learning has moved from one specific kind of syllabus to another, focusing on one teaching approach or another. The most common ones are:

- The structural syllabus from the 1930s to the 1960s in which the teaching of the language was based on the selection of the grammatical items and structures (tenses, sentence patterns, etc.) was regarded as synonymous with language teaching / learning.
- The 1970s witnessed the communicative approach which introduced functions and notions in teaching/learning English. This approach was widespread in E.S.P, both in the written and spoken language. Textbooks in E.S.P. (Nucleus series, English in Focus series) were geared towards the use of these functions and notions. This approach is still used nowadays in language teaching.

The two approaches listed above have been adapted and widely used in Algeria, either at secondary school level, or university level.

- Competency-based language teaching which sees learning outcomes in terms of competencies was mostly used in 1970s and 1980s (Richards, 2001:128).

This language teaching provides courses tailored specifically for a group of learners in Great Britain, Australia and the United States of America (in English-speaking countries).

For Richards, competencies refer to "*observable behaviours that are necessary for the successful completion of real–world activities"* (2001:129). As an example, Richards referred to the case of all the refugees in Australia who, in 1986, had to enroll in a competency-based programme if they wished to receive federal assistance. Australia is considered as the world's largest providers of language training to immigrants using a competency-based approach. The learning outcomes of such language training are specified in terms of work-related competencies i.e. "*to enable learners to participate effectively in society."* (Richards, 2001:131). A similar course called 'ESOL Jobcentre Plus Course' is given in Leeds for people who have just settled down in Britain and need to improve their English for employment.

However, a shortcoming of all the syllabuses listed above is that "*they tend to isolate, divide and sub-divide ..."* (Lewis, 1993:108). Instead, we should adopt a more holistic view of language.

With the advent of computational linguistics, the 1990s saw the emergence of the Lexical Syllabus and the Lexical Approach. According to Lewis (1993:109), *"Sinclair and Willis largely equate the Lexical Syllabus with a word-based syllabus."* This means that a lexical syllabus *"would specify words, then meanings and the common phrases in which they were used"*, (Willis, in Lewis, 1993:109), that is to say the lexical item and its co-text.

The most distinctive feature of the Lexical Approach is its attitude to the treatment of text. Lewis (1993:106) summarises the Lexical Approach as follows:

- Firstly, it is conspicious of de-contextualised language, recognising the importance of co-text, and therefore preferring extended text or discourse.
- Secondly, it proposes a range of awareness-raising activities directing students' attention to the chunks of which text is composed.

However, the development of different teaching approaches has sometimes relegated the teaching of content of language to a secondary role. Carter and McCarthy state that *"in extreme cases, such an approach to methodology denies the relevance of a content syllabus – in task-based learning, for example, or in the communicative approach. Such confidence in method renders syllabuses unnecessary."(1988: 45)*

In this research, we would not go to such an extreme, but would consider such a point of view. We do not as well aim to design a syllabus as such, but to suggest a teaching / learning methodology for students in biology departments because as it is stated by Lewis (1993:108) *"a shortcoming of task-based approaches is that they make it difficult to specify syllabus content."*

Another claim made by Ellis (2003:205) is that the distinction between 'design' and 'methodology' is not relevant in task-based teaching. He further explains this point by referring to Nunan (1989) who argues that in this kind of teaching the focus shifts from "*the outcomes of instruction", i.e. the linguistic knowledge or skills to be mastered, towards "the processes of learning", i.e. what learners need to do in order to learn.* Thus, Nunan claims, *"the 'what' and the 'how' of teaching are merged."*

Therefore, my own concern is to look at the contribution which lexical items can make in determining content.

In ‘Chapter Five’, we specified the different lexical items (i.e. content words in the form of NCs) and we explained their meanings and the phrases in which they were used (their co-text). As multi-word lexical items are under-exploited in language teaching/ learning, the Lexical Approach I propose (not a lexical syllabus) recognises these multi-word lexical items for their collocational power (their frequency in the corpus and their importance in the field of biology) as requiring special pedagogical treatment. This special treatment will also be embedded in the development of the teaching methodology that would specify the kinds of tasks that students need to do in order to learn the language.

The term ‘methodology’ has been mentioned several times in this research without specifying its meaning. In the following part, it is necessary to:

- specify the meaning of methodology, and then
- determine the kind of teaching methodology to be adopted in this research.

6.4 Methodology in language teaching

The spread of the English language throughout the world in different disciplines, and for different purposes, has brought new views about its teaching. In E.S.P, the development of the English language teaching profession (in recent years) has seen the rise of methodology to a dominating position. This can be seen in the number of published books and conferences held around the world e.g. In U.K., the British Council conference in 1984, the TDTR (Teachers Develop Teachers Research) Conference / Aston University in 1992, and other conferences held in U.S.A, Malaysia, etc.

6.4.1 Methodology in the literature

To Carter (1993:50) “methodology refers to the range of teaching methods, procedures and strategies adopted by a teacher in order to help learners to acquire a language.” Carter also states that the more teachers understand about the structure and organisation of language, the more principled and systematic can be their choices and design of language learning activities. Therefore, language teaching / learning is seen in terms of approach, method and technique.

Richards *et al.*'s explanation of methodology is more wholesome and more appropriate to the learning context. According to Richards *et al.* (1985:177) methodology is, on the one hand, "*the study of the practices and procedures used in teaching, and the principles and beliefs that underline them.*"

It includes:

a) *the study of the nature of language skills (reading , writing, speaking, listening) and procedures for teaching them;*
b) *the study of the preparation of lesson plans, materials and textbooks for teaching language skills;*
c) *the evaluation and comparison of language teaching methods."*

On the other hand, "*methodology refers to such practices, procedures, principles and beliefs themselves. One can criticize or praize the methodology of a particular language course."* (Ibid). However, in language teaching the problem that most teachers face is to decide which approach and method to adopt, in order to find the appropriate techniques to be successfully implemented in class.

The development of the many and different teaching approaches, methods and techniques shows that no single approach and method can guarantee successful results.
Carter and McCarthy found that

> *no one method of organization is adequate for a balanced and comprehensive course because language has many facets, and corresponds partially to many different patterns of organization. It can be represented substantially as a set of structures, and a list of words; it can also be seen as performing a variety of functions. Or the learner can be monitored through the skills or through a set of tasks.* (Carter and McCarthy, 1988: 144)

They, then, carried on by suggesting that some teaching from each of these points of view is necessary to make the teaching effective and efficient.

Adopting such a view would thus lead us to some sort of eclecticism.

6.4.2 Methodology: a proposal

The present methodology would focus on:

i) Specification of the language skill

Based on the students' answers in the questionnaire in 'Chapter Three', the methodology used in this research will focus on the reading skill. We will not aim at developing the reading strategies or reading skills as such - the problems are seen as lying within the grammar and vocabulary of the discourse, the reason why we have used a computational analysis in 'Chapter Five'.

ii) Use of the Target Language

The method to be used in this research involves the use of the target language. As E.S.P. courses are necessarily subject-matter oriented, decisions about the choice of language content relate both to subject-matter and linguistic matter. In this direction, Rodgers and Richards pointed out that "*One makes decisions about what to talk about (subject matter) and how to talk about it (linguistic matter).*" (1986:21)

In order to deal with methodology, it is important to refer to certain conditions necessary in the language learning process.

6.5 Conditions for Language Teaching / Learning

Willis (1996:11) summarises the conditions for effective language learning into four (three essential and one desirable).

For Willis (1996) the three essential conditions for language learning are:

- Exposure to the target language,
- Use of the target language, and
- Motivation.

Willis states that one without the other, or even two without the third, will not be sufficient.

- The fourth condition for Willis is 'Instruction' which although not totally essential, is highly desirable.

For our learners (as non-native speakers of English) the four conditions are weighted equally i.e. the four conditions are essential for their language learning.

a) Exposure to the Target Language

The learners should be exposed to the target language in use. In our case, students should be exposed to texts written in English from biology source materials e.g. journal articles.

For a reading course to be most cost-effective for the students, a solution is to bring their own texts for study in class.

The target texts should be authentic, and should not be simplified for two main reasons:

- Simplification such as removing certain scientific terms and / or processes may deprive the learners from getting the necessary information about their discipline from the text.

- According to Willis, "*simplification such as removing certain features of a text, for example by rewriting complex noun groups or breaking up grammatically complex sentences into a series of two or more simple ones, deprives learners of the opportunity to become familiar with the original forms, which may occur frequently in the target language.*" (Willis, 1996:12)

b) Use of the Target Language

Once the texts for study have been selected, we should give students opportunities to use the target language. By asking the students to do specific tasks / activities, we aim at making them use the target language in class. The other important aim is to help students develop some kind of language awareness so as to use the target language out of class (on an individual basis, independently of a teacher) in the long run.

c) Motivation

According to Ellis (2003:235) "*content-based courses are premised on the assumptions that learner will best learn language while they are engaged in learning subject-matter.*"
Therefore, by providing the students with the kind of literature (journal articles) that they need to read from the target literature (biology) as requested by them in the questionnaires in 'Chapter Three', it is hoped that the students become motivated. This exposure becomes real input and should help language acquisition. In this case, the students' motivation is purely 'instrumental', (whereby Mitchell's codification of social networks in Chapter Two / Section 2.2.2), and they see the target language as "*a means to an end.*" (Willis, 1996:14)

Motivation is not necessarily understood as stimulating whether in the case of 'a means to an end language' or 'language as an end' i.e. a culture oriented acquisition. English is largely apprehended as the medium of an ideology. This a priori about foreign languages, as essentially determined by ideological purposes, fosters doubts and hesitations concerning the possibility of reducing a language to a strict instrument of professional mediation, or a constructive apparatus that could make two cultures (the learner's and the foreign one) meet on equal grounds of exchange and mutual knowledge.

Teachers will have to devote a special attention to 'de-ideologisation' of languages, specially in the Algerian context where for historical reasons, the use and acknowledgement of Arabic as a national language also meant the riddance of colonial = 'foreign' values prior to the recovery of an Algerian language and identity.

d) Instruction

For Willis (1996) instruction (the fourth condition of language learning) is desirable. In our case, 'Instruction' is the essential condition for language learning.
According to Willis (1996:16) "*instruction can certainly help students notice specific features of the target language. It can give students the opportunity to process grammatical and lexical patterns, and to form hypotheses about their use and meaning.*"
Also for Willis "*learners are more likely to recognise these features occurring in the input they are exposed to.*"*(Ibid)*

Therefore, in ‘Chapter Five’, the analysis of the most salient (frequent and significant) items used in the biology discourse and their concordance study is done to show the students some examples of what grammatical and lexical patterns constitute the biology discourse, and also to examine evidence of when these are used.

The main objective from this study is, on the one hand, to provide teaching materials that promote language awareness. It makes the students aware and conscious of particular biology language features occurring naturally in their reading, and how these function in particular instances. The teaching materials should include tasks and activities that would “*drive the students’ language development forward.”* (Willis, 1996: 16). On the other hand, students should also be given a chance to ask about other features they notice in the discourse under study, and that are not dealt with in the teaching / learning programmes.

6.6 Suggested Methodology for Teaching English to Undergraduate Biology Students

It was stated in ‘Chapter One’ that the objectives of teaching English to Algerian university students is a desirable end, especially in the situation where pre-university education is largely arabised while science and technology in the world is largely anglicised.

However, in ‘Chapter One’, while introducing the research problem we stated that the current model of education in Algeria may be described as consisting of objectives and practices. We, then stated that practices consist of means that are assumed realised ends and that it is in these means that the inefficiencies attendant to the current model are located.

In this part, we have tried to remedy some of these inefficiencies by providing ‘what we think to be an appropriate methodology’ for teaching English to undergraduate biology students who need to practise processing written texts quickly and efficiently.

The methodology aimed at, what Kumaravidelu (1991:98) calls “*a pedagogic perestroika*” is “*a fundamental restructuring of the relationship between teaching input and learning outcome; between curricular content and classroom procedures; and, most of all, between teacher and learner roles in generating classroom discourse.”*

This suggested methodology will be based on Willis's four conditions for effective language learning.

6.6.1 Learning conditions

6.6.1.1 Exposure to the target language

In the questionnaires' answers in 'Chapter Three', it was found out that teachers of English in the biology department used 'E.S.P. concocted texts' instead of 'authentic texts'. These 'ESP concocted texts' are generally simplified versions of the original texts, and are produced for teaching purposes only. 'ESP concocted texts' may be of some help in class only. Outside the classroom, biology students need to read biology texts from specialised books and journals.

Therefore, for the biology reading course to be most effective, we have to provide students with the original version of texts, or ask them to bring their own texts to use in class. As such we can help them save time as they will read in class what they must read anyway, and at the same time we can help them to acquire the language of their specific discipline. This could be done by giving students a full description of the characteristic features that constitute the biology discourse. In this research, through a corpus-based analysis, a list of the different features (lexical and grammatical) forming the biology discourse is given in 'Chapters Four and Five.'

Using the information we have got from the corpus text analysis, we can provide evidence of typical language use, thus leading us to the part 'instruction'.

6.6.1.2 Instruction

In this research, 'instruction' is the second essential condition for language learning, as compared to Willis's model where it is ranked in the fourth position.

The research purpose is to raise the students' language awareness by helping them find out the specific patterns of the target language of biology discourse through 'instruction.'

'Instruction' is an essential condition for the target language learning where the students are taught those language features (be they grammatical or lexical) occurring naturally in the input (journal articles) they have to read. As such, we suggest to raise the students'

language awareness by providing tasks / activities that would drive the students' language learning forward in order to help them become independent readers. Therefore, in this chapter, tasks and activities to meet the above objectives will be developed in order to help learners use the target language.

6.6.1.3 Use of the target language

Once conditions one and two have been fulfilled, that is, once the students have been exposed to the target language of the target literature, and the different features (lexical and grammatical) of the target language have been dealt with, the students are asked to do tasks and activities related to those features that mostly cause problems. By doing communicative tasks and activities, it is hoped to develop the students' learning so that they get experience in the language. It is only when the non-native student learns the language in context, and not in a traditional teaching course, that s/he manages to master linguistic items such as determinants. These are elements that contribute to the acquisition of idiomatic formulas through "live situations" i.e. communication, rather than in "artefact situations" such as formal learning about grammar. These observations tend to become a fact today as the result of what Johns calls "*extensive experience.*" (1991:28).

Also when the learners' practice is geared towards his expectations i.e. towards providing him/her with those efficient elements that would facilitate his / her text comprehension, it points out his professional needs and it helps in accumulating more information i.e. increase in data acquisition. This gain, being nothing but an extension of the corpus, will directly stimulate his / her motivation in providing new questioning, perspectives and consequently new needs.

6.6.1.4 Motivation

Motivation, in general terms, refers to the different factors that affect the person's desire / wants to do something, and this also applies to motivation of learners in language learning. Richards et al. distinguish between two types of motivation:

a) *Instrumental motivation: wanting to learn a language because it will be useful for certain 'instrumental goals';*

b) Integrative motivation: wanting to learn a language in order to communicate with people of another culture who speak it.(Richards, Platt and Weber, 1985: 185)

The latter type of motivation is needed in Algeria at a much later stage of education i.e. at postgraduate and professional levels.

According to the questionnaire answers in 'Chapter Three', the learners' motivation is purely 'instrumental'. As undergraduate students, they want to learn English because it will be useful in reading their specialised journal articles to get the necessary information to write-up their 'mémoire' in order to pass and get their university degree. Therefore 'motivation' is seen in this research as the fourth essential condition for language learning. In our context, 'motivation' could take place by providing the learners with 'authentic' texts i.e. journal articles that they are most likely to read for their 'mémoire' writing, this being related to the first essential condition for language learning 'Exposure to the target language.'

'Authentic texts' should be provided because they show how the language is used by the speakers or writers in natural situations. Extensive reading of authentic material covering all sorts of scientific documentation in this field will certainly contribute to put them in the mental frame of a "*greater naturalness in their own language use."* (Fox, 1987:149). Such a disposition for improving their capacity for assimilation requires a training in reading skills specifically adapted to their urgent needs in the handling of documentation. Indeed, these needs are divided in two types:

a) The necessity to construct a highly targeted lexicon to process their research questions i.e. the building of a conceptual apparatus. In this case they need to select the parts, passages, chapters of books immediately concerned by this question and penetrate their depth through 'scanning.'
b) The necessity to locate this lexicon in a peripheral terminology that would show their scientific curiosity, i.e. their reading background, and which requires from them casual readings through 'skimming'.

A good bibliography at the end of the research work, will distinguish these two types of sources, one referential providing direct inspiration, and the other one complementary identifying the discourse field.

For effective language learning, provision of authentic texts alone would not bring all the necessary motivation for language learning. Thus, "instruction" (the second essential condition) represented in:

a) the teacher's role to uncover and explain all the most salient grammatical and lexical features that form the present discourse under study; and
b) giving the learners the opportunity to process these lexical and grammatical features and raise their awareness,

helps them to form "*hypotheses about these items use and meaning*." (Willis, 1996:16). This could be done by developing tasks and activities to do in class (and outside the classroom) to develop their learning by using the target language.

To reach such a stage, students need to realise the following learning outcomes.

6.6.2 Learning outcomes

Some examples of these learning outcomes that the suggested methodology may propose are given below.

- By the end of the course learners will be able to process written texts quickly and efficiently by:

a) Developing the reading competence of the students by promoting discovery of meaning and by making students aware of representative patterns of language use and selected grammatical structures;
b) Helping learners to experience language in use and discover new knowledge for themselves rather than being spoon-fed i.e. to help learners become independent readers;
c) Encouraging the underlying competence of the students to deal with any kind of text to achieve a reasonable level of reading competence in this language;
d) Stimulating the need to value 'How to read' more than 'What to read' (as our learners already know what they need to read).

To reach the above objectives, it was thought to develop 'consciousness-raising' ('C-R') in students' learning, or develop students' language awareness. Willis and Willis explain C-R as the fact "*to provide learners with activities which encourage them to think about samples of language and to draw their own conclusions about how the language works.*" (Willis and Willis, 1996:63).

Among the characteristics of C-R, Willis and Willis list:

- *The attempt to isolate a specific linguistic feature for focused attention;*
- *The provision of data which illustrate the targeted feature.*
- *The requirement that learners utilize intellectual effort to understand the targeted feature (to involve the learner in hypothesizing about the data and encourage hypothesis testing)* (1996:64)

In other words, "*learners are encouraged to notice particular features of the language, to draw conclusions from what they notice and to organize their view of language in the light of the conclusions they have drawn.*" (Ibid). However, we should avoid putting too much emphasis on target language analysis because it is often said that "*too much self-consciousness can restrict opportunities for language acquisition and can inhibit the learner.*" *(*McCarthy and Carter, 1994:160).

We should, instead, develop communicative teaching practices which would help learners to analyse the language and to experience it in use, as in general "*Learners begin with a holistic experience of language in use. They end with a closer look at some of the features naturally occurring in that language.*" (Willis and Willis, 1996:52). Therefore, in this research, we aim at reaching the students' consciousness-raising by suggesting a task-based teaching.

In this work, we are taking Willis and Willis's definition of the word '*task*' to mean "*a goal-oriented activity in which learners use language to achieve a real outcome.*" (1996:53), that is to say "*activities where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome.*" (Willis, 1996:3).

The other reason for suggesting a task-based teaching / learning is because "task" *is the ideal unit for specifying the content of specific purpose courses because it most closely reflects what learners need to do with the language." (Long,* in Ellis, 2003:209).

As the learning skill needed by students has been specified by the students themselves, their teachers and the administration as being the reading skill, therefore, we should consider the following criteria for determining the kinds of tasks to include in the English reading course.

Although, as Ellis noted, tasks do not need to be graded with the same level of precision as linguistic content (2003:220), in this research we have tried to establish some criteria for the selection of tasks.

For Ellis (2003:210), in task-based syllabuses, we should consider:

- How to classify tasks;
- What the thematic content of the tasks should be;
- How to sequence tasks in accordance with explicit grading criteria.

In this research these criteria have been respected as they apply to the situation. The selected tasks to be done by the students cover:

- Classification and it includes: gap-tasks, information transfer, listing, comparing, problem-solving, project-writing, etc.;
- The tasks are built around themes related to the students' subject-speciality (biology) to be done in English. The suggested tasks focus on form and content;
- Sequencing tasks: the tasks are sequenced in terms of complexity to match the learners' level of development. This means that the tasks are planned and presented in different phases (pre-task, during task and post-task) and as such they start from easy tasks (predicting, listing, information transfer, etc.), to more difficult ones (comparing, problem-solving, project-writing), bearing in mind that the notion of 'difficult tasks' and 'easy tasks' is really debatable in terms of language learning.

The sequencing and content of tasks are meant to create opportunities for language learning and skill development through collaborative knowledge building.

The suggested task-based pedagogy will also be based on the use of concordance output to create language-learning activities. The value of concordancing lies in the development of teaching materials focusing on vocabulary and grammar. According to Tribble and Jones (1990:35), two generalizations can be made about applications of concordance output:

a) *Most of them favour discovery learning i.e. to enable learners to discover new knowledge for themselves, rather than being spoon-fed;*

b) *They do this by providing examples of authentic language.*

The fact that the source material for exercises is drawn from authentic texts, rather than being concocted by teachers, increases motivation as it gives learners immediate contact with the target language in use. In the same line of thought, Thurstun and Candlin find that "*using concordance-based materials is that they can help develop this ability to guess the meaning and use of unknown words from context."* (1998:270)

The task-based activities using concordancing "*are intended to familiarize learners with the use and meaning of the key words, helping them to develop insights into the collocations and grammatical structures with which the key words are associated."* (Thurstun and candlin, 1998:271).
The main features to be studied have been singled out through computational analysis (Chapter Five), and a teaching methodology to help learners process texts quickly and efficiently have been specified above.

In the following part, we suggest a general teaching framework of the reading course together with some tasks. These tasks will be designed to develop the students' language awareness so that they could cope with their text processing for information necessary for the writing-up of their final year project .

diagram

6.7.1 Pre-task Phase

The activity in this phase is used for brainstorming and could last between five to fifteen minutes. Skehan (1998a) suggests ten minutes is optimal. (in Ellis, 2003:248). The pre-task phase introduces the class to the topic and the task. Teachers explain the task and clarify the intended outcome. In this phase, as far as the biology field under study is concerned, we can include the following tasks:

- To give the title or part of the text and ask the students to read it and predict the rest of the text;
- As biology deals with plants and animals, to give the students pictures of different plants and animals and ask them to discuss the relation that could exist between these pictures and the content of the text.
- To give them tabularised information representing results of an experiment and ask them to guess what the experiment was about and how it was done. The students begin by discussing the points under study, and try to identify how the experiment was done. The students are then given the text.

Sometimes anecdotes or events are linked to the experiment or the scientist who published it. Historical data and situational framing in cultures in which it occurred are excellent mnemonics to sediment knowledge and helps understand how sciences progress through time and across languages. Anecdotic revisiting of an ‘experimental event’ gives stuff to recounting which constitutes the best form of exposure through narration. All human knowledge unveils through narration, because it is one of the human being’s fundamental functions to tell the world as he experiences it.

6.7.2 Task-cycle

In this phase, the students read the text to confirm their predictions made in the pre-task phase. They, then, do the task on their own or in pairs / groups, using whatever language they want, in order to solve a problem, or explain an experiment, etc.

The students will have in the end to report the task they did to the whole class in the target language i.e. English, and compare their results.

6.7.3 Language Focus

Willis (1996:38) divides this cycle in two parts: 'Analysis' and 'Practice.'

- In 'Analysis', the students examine and discuss specific features of the text.
- In 'Practice', the teacher conducts practice of new words, phrases and patterns occurring in the data, either during or after the analysis.

In this research, 'Language Focus' is the most important phase of the framework and it represents the second condition for language learning i.e. 'Instruction' mentioned in this chapter.

This phase of the framework is relatively new in language teaching. It is used to raise the students' awareness by showing the different language features occurring naturally in the text. It is also narrowly linked to the research work done in 'Chapter Five' and which provided a solid foundation for the description of language use in biology.

The computational analysis in 'Chapter Five' was done with the purpose to be used in this phase. For Willis and Willis, the '*Language Focus phase*' helps students to explore language, to develop an awareness of aspects of syntax, collocation and lexis, to help them systemise what they have observed about certain features of language, to clarify concepts and to notice new things. (1996:58)

The computational analysis in 'Chapter Five / Section 5.5' unveiled the different language features often used in biology texts. These features are:

- Nominal Compounds,
- Tenses (mainly The past simple, the perfect tenses, the simple present, etc.),
- Prepositions,
- Adjectives (including participles),
- Affixes and suffixes.

These items should be focused upon because, not only were they found to be frequently used in biology texts (through computational analysis), but the students also stressed the importance of these items in the questionnaire answers.

Li and Pemberton (1994:184) in (Thurstun and Candlin, 1998:268) point out that tertiary students do not necessarily find discipline-specific technical vocabulary difficult, "*rather, it is the vocabulary with a middle frequency of occurrence across texts of various disciplines that students find most problematic.*" Nation (1990) refers to this range of vocabulary items as *'academic vocabulary.'*

In the 'Practice' part of the 'Language Focus' phase, the teacher sets some language-focused tasks, based on the text that the students have read. Examples of tasks to be included in this phase would be :

a) Read the text and find NCs of the type N + adjective + N, N+N+N, etc., and say or explain what the NC means (i.e. use of disambiguation)
b) Find all the verbs in the simple past form. Say which refer to past time and which do not.
c) In this research, 'Practice' will be mainly based on 'concordancing' because the latter gives you a way of seeing patterns in language in use that would remain hidden under other circumstances.

Tasks should be developed with the underlying principles that:

- *Students need to be made aware of why they are being asked to perform tasks;*
- *Ensure that students adopt an active role in task-based lessons;*
- *Encourage students to take risks;*
- *Ensure that students are primarily focused on meaning when they perform tasks;*
- *Provide opportunities for focusing on form;*
- *Require students to evaluate their performance and progress*.(Ellis, 2003:276-278)

We may say that the above teaching framework has been initiated and exposed in order to be used later in a suggested teaching lesson for biology students.

Thus, and following Willis's teaching framework, we suggest the following way of exploiting the selected teaching text in class.

Students are divided in pairs or groups of three or four. The teacher will organise a kind of competition, in time and number of items found, between these groups. When it happens often enough, it could be evaluated in the form of a 'continuous assessment' that would be

more efficient and learning made more rewarding than the traditional term paper for the routine pass mark unique administrative obsession.

6.8 Willis's Teaching Framework to be implemented in class

6.8.1 Pre-task phase

The teacher gives the students **the title of the article only**, in our case we suggest one of the text of the corpus under the title '*Hypercholesterolemic Responses in Rabbits to Selected groups of Dietary Essential Amino Acids'* (by Elzbieta M.Kurowska and Kenneth K.Carroll, 1994) together **with table 2** below (as published in the original text / Appendix E)

TABLE 2

Serum total and LDL cholesterol concentrations and growth performance of rabbits fed experimental diets[1]

Diet		Serum total cholesterol	LDL cholesterol	Initial weight	Weight gain	Food consumption
		mmol/L		*g*	*g/d*	
Experiment 1						
14.7% amino acids	(6)	2.28 ± 0.26[a]	1.26 ± 0.18[a]	1630 ± 45	12 ± 2[b]	68 ± 4[b]
45%[2] Lys Leu Ile Met Thr	(6)	5.51 ± 0.26[b]	4.40 ± 0.31[b]	1610 ± 70	3 ± 1[a]	43 ± 3[a]
45% Lys Leu Ile Met Val	(5)	5.95 ± 1.16[b]	4.91 ± 1.14[b]	1670 ± 40	−5 ± 2[a]	42 ± 3[a]
45% Lys Leu Met	(5)	5.28 ± 0.62[b]	4.47 ± 0.54[b]	1630 ± 30	0 ± 2[a]	37 ± 4[a]
45% Thr His Phe Trp Gly	(6)	3.65 ± 0.41[a]	2.97 ± 0.31[b]	1640 ± 50	13 ± 2[b]	64 ± 2[b]
Experiment 2						
14.7% amino acids	(5)	3.00 ± 0.41[a]	2.04 ± 0.36[a]	1580 ± 30	11 ± 3[b]	71 ± 7[b]
45% Lys Met	(6)	7.73 ± 1.11[b]	6.36 ± 0.91[b]	1600 ± 25	−4 ± 2[a]	32 ± 7[a]
45% Lys Leu	(5)	4.60 ± 0.70[a]	3.65 ± 0.65[a]	1600 ± 40	4 ± 3[b]	49 ± 4[b]
45% Leu Met	(5)	4.66 ± 0.36[a]	3.57 ± 0.39[a]	1610 ± 60	7 ± 2[b]	41 ± 4[b]

[1]Values are given as means ± SEM. Number of rabbits per group is shown in parentheses. Results were obtained after 3 wk of feeding. Values for each experiment in a column with different superscripts are significantly different ($P < 0.05$).
[2]Amino acids included at levels corresponding to a 45% casein amino acid diet.

and asks the students to:

a. Predict the content of the text.

The pre-task phase is used as the introduction to the topic and task, where students discuss among themselves (in pairs or groups as stated before) to predict the content of the whole article.

Students predict or attempt to reconstruct the content on the basis of given clues from part of the text, without having read or seen the whole.

b. Write 4 or 5 questions and ask themselves if they are likely to be answered in the full text. Later, they see how many of their questions are answered.

c. Match texts to visuals: when a description of an experiment is accompanied by diagrams, pictures, etc., the teacher could jumble either the text or the visuals.

In this phase, we can use some 'traditional' exercise types for brainstorming and bring them into harmony by introducing them into the classroom as exploratory and collaborative tasks, as it is done above.

6.8.2 Task-Cycle

In this phase, tasks are also set by the teacher. The students are given one task or more depending on the purpose to be achieved.

After giving 'the title of the text together with Table 2' in the pre-task phase, ask pairs or groups of students to do the following suggested tasks :

***TASK 1* :** Write down 5 questions you would like answers to.

- Teachers should give students a few minutes to write their questions.
- Students then plan a brief report to the whole class to compare their different predictions.
- Pairs (or groups) tell each other the questions they thought of and discuss possible answers.

After the report phase (as long as the teacher does not give away the correct answers), the class will naturally want to read the piece again to see who remembered the best and whose first impressions were the most accurate (or the most inadequate, sometimes even strange). Strange answers may make classroom situations very enthusiastic, because of funny or weird propositions. It may even become a game whereby the teacher may encourage improvisation, role-playing and all forms of creative activities that usually students appreciate because it appeals to their capacity to initiate original propositions.

Being science students, these activities reconcile them with the scientific principle that discovery and invention are always preceded by intuition and dream that is imaginary

venture. *"Telling stories, about ourselves and about others, to ourselves and to others, is the most natural and the earliest way in which we organise our experience and our knowledge."* (Geertz, 2000: 193).

After the report stage, students are given the whole text to read once or twice, to see how close / strange their predictions were. Students read again the whole text in order to find how many of their questions got answered (5 out of 5, 4/5, 3/5, etc. or none.)
The teacher gives general feedback about the performance of students.

Task 2: **Memory Challenge**
The teacher sets the second task which is 'memory challenge.' S/He also has to divide the class in two groups :
*GROUP 1***:** Pairs turn the text over and are asked to list the five important stages of 'Materials and Methods Section' of the text under study (which are: Animals, Diets, cholesterol analysis, Analysis of plasma Ketone Bodies and free fatty acids, and Statistical Analysis).

GROUP 2 **:** will have to report the different results got from the experiments held in the section 'Materials and Methods.'

Planning and Report
Once the students finish the task, they report the answers to the whole class. The teacher encourages but does not reveal solutions. Finally all students read the text again to check what they have written and see which pairs / groups remembered most. The teacher gives general feedback about the students' performance in task 2.

As it is noticed above, it is the learners who should be doing the work. The teacher's role is that of a facilitator or guide, to help students process the text for themselves in order to do the task. When there is competition or contradiction between the groups, the teacher becomes a moderator or / and a referee.

After the students finish reporting task 2, the teacher gives a summing-up or evaluation session about the whole students' performance in *TASK 1* and *TASK 2* before focusing on language.

The final phase in the task framework is *'Language Focus'* with its two components *'Analysis and Practice.'*

'Language Focus' gives learners the opportunity to have a closer look at the different language forms used in the text and also to see how language functions in context.

McCarthy and Carter (1994:194) see "*task-based approaches as lending themselves best to engagement with the discourse process."* Therefore, we believe that the '*task-cycle*' is important for our learners (as non-native speakers of English) because it helps them to process texts quickly and efficiently:

a) quickly: because learners should focus on important information instead of a word-for-word reading, and
b) efficiently: because the information should be targeted i.e. restricted to the immediate scope of the theme necessary for our learners to construct their reasoning process.

Through these tasks, learners use the target language in order to achieve an outcome i.e. to process texts efficiently. They do the task, exchange meanings and ideas, write these ideas, present them orally to the class and discuss them among themselves. Each pair / group tries to explain HOW or / and WHY they have included such ideas.

It is only by going through this whole process that we can call the above exercises as 'tasks' considering the definition of 'task' given by Willis and Willis (1996:53) and which was hence used in the present research, i.e. a "*task being a goal-oriented activity in which learners use language to achieve a real outcome."*

6.8.3 Language Focus

As stated previously (in 'Section 6.7.3 / *Language Focus'*), this phase is important because it brings to the fore the lexical and grammatical features that constitute the whole text i.e. the discourse of biology.

The teacher should do a text analysis beforehand, to find out what lexical and grammatical items are widely used. S/He should, then, develop different tasks based on these used features. This way the content of the lesson will be based on empirical evidence rather than on intuition - as was done by teachers previously.

To come back to the text under study *"Hypercholesterolemic Responses in Rabbits to Selected groups of dietary Essential Amino Acids, 1994"*, the analysis done manually revealed the following features.

The text is divided into (and so are the other texts of the corpus used in 'Chapters Four' and 'Five') into these parts: Title - Name(s) of author(s)- Abstract - Introduction - Materials and Methods - Results and Discussion .

The overwhelming features used in the text under study are:

- NCs forming collocations in most cases,
- simple past tense,
- participles,
- prepositions, and
- modals.

6.8.3.1 Nominal Compounds

The most striking feature is the use of biology terms in the form of NCs, and these in turn form collocations most of the time. This result confirms the first findings of what constitute biology discourse referred to in 'Chapter Five.'

The manual counting shows that the length of the NCs, used in the text under study, varies between 2 noun NCs and 5 noun NCs. The whole text contains:

- 34 examples of 2 noun NCs (base pair, plasma cholesterol, blood samples, animal proteins, plant proteins, ketone bodies, etc.);
- 12 examples of 3 noun NCs (plasma ketone bodies, blood cholesterol concentrations, plasma cholesterol concentrations, etc .);
- 5 examples of 4 noun NCs (casein amino acid mixture, casein amino acid diet, etc.)

The above examples are not only NCs as such, but they form collocations where for example ‘base’ always collocates with ‘pair’, ‘cholesterol’ collocates with ‘plasma’ or ‘blood’, ‘casein’ with ‘amino acid’, etc.

- 1 example of 5 nouns Nc (soybean protein amino acids, where ‘soybean’ is itself a NC formed of 2 nouns ‘Soya’ and ‘bean’)

The figures above show clearly the much higher use of NCs in biology discourse as professional reading i.e. journal article. This is seen as a problem for an Algerian undergraduate L3 / L4 reader trying to read the literature of his specialist field, who has been used to reading only general and simple English texts. This problem is mainly due to the fact that the linguistic form of NCs is rarely used in Algerian readers’L1 (Arabic/Berber) and if used, it reads differently from English, and is absent in the Algerian readers’L2 (French). For example, in French, we say ‘concentration de cholesterol dans le sang’ for ‘blood cholesterol concentrations’, and ‘échantillons de sang’ for ‘blood samples’, ‘flux de gènes’ for ‘gene flow’ etc., where ‘**de**’ meaning ‘of’ is always used.

In Arabic, we also use prepositions, and we say / ارتفاع الكولسترول في الدم / for ‘blood cholesterol concentrations.’

For reasons explained above, it is important to deal with NCs in the biology course by developing different tasks to help learners understand how the English language functions in context. Unlike the native English reader, the non-native reader, in this case, the Algerian undergraduate student has not internalised the set of recovery procedures that would help him, once transferred, operate on English NCs (Williams, 1984:146).

The other features present in the text under study are grammatical structures.

6.8.3.2 Grammatical structures

- The simple past (Active and Passive voice), participles and modal verbs are all interrelated areas of English grammar. Their introduction in the course is not based on tradition or intuition, but rather on empirical evidence (as these were found out in the text and were counted). The high frequency of such items in speciality texts is important to some extent, but what is more important is rather to find out what there might be interesting. This could be answered by the different tasks to be done by the learners.

In the text under study, the figures of the use of the simple past are as follows:

- In the section ‘Abstract’, there are 2 examples of the simple past passive and one example of the active form;
- In the ‘Introduction”, there are 18 examples of simple past active and 10 examples in the passive voice;
- In the section “Materials and Methods”, there are 26 examples of simple past passive and 5 examples of simple past active;
- In the ‘Results’ section, there are 16 examples of the simple past tense/active voice together with 1 example of the past perfect tense/active voice e.g. animals fed the lysine... diet **had already produced** significant elevation of serum cholesterol, and 6 examples of the simple past passive;
- In the “Discussion” section, there are 18 examples of the simple past / passive and 18 examples of the simple past / active.

The total use of the simple past in the text under study is 62 examples of the simple past tense/passive voice , and 56 examples of the simple past /active voice, with a higher use of the simple past tense/passive in the section ‘Materials and Methods”, as scientists are more concerned with ‘what an experiment consisted of’ and ‘how it was done’, rather than with ‘who did it’.

The present simple tense is also used but with less frequency. It is used once in the passive voice in the ‘Abstract”.

In the ‘Introduction’, there are two examples of the present tense / active voice, and two examples of the present tense / passive voice, used in the first paragraph only.

In the ‘Materials and Methods” section, only the simple past in both forms (Active and Passive) is used.

In the section ‘Results’, there are two examples of the present tense / passive voice, and one example of the present active. These are used only in the first sentence of the section.

In the last section 'Discussion', there are ten examples of the present tense / active voice represented mainly by the verbs (show, indicate, etc.), and three examples of the present tense / passive voice .
The figures above give the total use of 8 examples of the present tense / passive voice, and 13 examples of the present tense / active voice in the whole text under study.

The other grammatical feature in the text is the use of 'modal auxiliaries'. Modals constitute one of grammatical problem areas in Teaching English as a Foreign Language.

- **The modals** used are mainly 'can' in its active voice (1 example i.e. On the basis of our present and previous results we **can,** therefore, attempt to rank essential amino acids...), and

5 examples of 'can /could' in the passive voice as illustrated below:

- Plasma cholesterol concentration **can be altered b**y changing...;
- The hypercholesterolemic effect of animal proteins **can be attributed** to their generally higher content of essential amino acids;
- Our results show that an even more pronounced hypercholesterolemia **can be obtained** in rabbits by...;
- This rather weak ketotic response **could** simply **be induced by** an increase in the catabolism.

There is also one case of the modal 'may' in the active voice e.g. This **may occur** directly via high postpandrial plasma concentrations.

These findings correspond to Butler's. Butler quantified the uses of the modal verbs in each of the IMRD sections of 12 texts (consisting of 2 journal articles and 2 textbook extracts) from the subject areas of Botany, Animal physiology and Physics, and found differing levels of density. This, he theorised, could be related to function:

- in *the introduction*, the writer informs the reader of what can be expected in the rest of the text, and therefore *will/shall* are frequently used;
- in *the results sections*, the use of *could* refers to what was or was not possible to do in the experiments;

- in *the discussion sections*, claims are made about what is legitimate to conclude, what *may* or *must* be the case.(Butler, 1990 in Thompson, 2000:148).

Last, but not least, is the use of participles in their two forms (present and past).

- Participles (or participial adjectives)

Another major sub-class of adjectives can be distinguished by its endings –ed or –ing, i.e. participial adjectives. A number of participles in –ed and –ing are frequently used as adjectives, both in general English and E.S.P. These could be a one-word adjective, e.g. amazing/amazed, interesting/interested, or compound adjectives, e.g. a hard-working student, a left-handed student, time-consuming, time-saving.

We should also refer to the function of these participles and notice the difference in meaning between such pairs as:

> "The thing that excites you is exciting (i.e. the active participle)"; and "The result is that you become excited (i.e. the passive participle)" (Allsop, 1988:74).

Another example: "This story-teller is so fascinating that anybody listening to him grows fascinated." (Malavieille et al., 1999:76).

Participles (or participial adjectives) are important to deal with in the biology class because they are overwhelmingly used:

1. with NCs , e.g.

remaining essential amino acids (present participle);

unbalanced amino acid diets (past participle);

2. to describe different scientific phenomena, e.g.:

- a **marked** elevation of serum total ...;
- **observed** differences / **observed** changes;
- **semi-purified** diets;
- various combinations of amino acids **compared** with plant proteins;…

Participles in their two forms need to be introduced in class and explained to Algerian students who, as non-native speakers of English, cannot draw the difference between the -ing and -ed forms of participles, e.g.: 'diets **containing** mixtures 2-8', as compared to, 'mixtures 2-8 **contained** in diets...'
These forms are actively used in the speech and writing of the learners, but are often used incorrectly.

The above features, drawn from the authentic text under study, represent impediments for scientific text processing to Algerian students. Therefore, they need to be introduced and explained in class through different tasks. In the language focus phase, the teacher should develop tasks that deal with the items mentioned above (be they lexical or grammatical) to explore their meaning and interpretation. For the purpose of tasks at this phase, it was thought useful to use 'concordancing'.

6.8.4 Computer Concordancing

Computer concordancing is used in this research for materials production, and it is introduced as a "*way to help students develop their own skills for interpreting texts*." (Kowitz and Carroll, 1991:135).
Besides the main reason given above, other reasons are taken into consideration for using concordancing in the classroom. These reasons, taken from Kowitz and Carroll, 1991:137, are summarised as follows:

a) Concordancing makes a text more accessible, by focusing attention on specific elements, rather than confronting the student with the whole text. The main features for study are selected for them (by the teacher) without the student having to extract them from the whole mass of text;

b) Using concordancing encourages pair / group work as the objective is to take a set of concordance output and divide it up between groups or individuals.
The fact that each student has a specific task encourages an investigative approach and promotes development of students' interpretative skills.

c) As concordancing lends itself easily to subdivision, thus creating a set of individual or small group assignments, it is thought to be the appropriate methodology for large classes. It gives each student the chance to participate in class as the given individual tasks in each lesson encourage individual participation.

Having given the reasons for using concordancing, in the following part, we shall describe the preparation of the materials for concordancing.

6.8.4.1 Steps for the preparation of concordance materials

The preparation of materials for concordancing follows necessary steps after having scanned and edited the texts in a computer-readable form (as explained in 'Chapter Five / Section 5.2).

These steps are:

- Selecting the items to concordance;
- producing the concordance; and
- editing the output.

i) Selecting items to concordance

It is the teacher's duty (after the students' consultation) to find out what items need to be focused on and selected. The selection varies from text to text, and from actually what the students need to know i.e. What items, Which principle, etc. needs to be made clear and explicit.

In this research, the main items to be focused on are those elements already listed in this chapter, e.g. NCs, participles, prepositions, etc. which were found to cause problems to students for text comprehension and which were also very frequently used in biology texts. The concordance lines do not themselves solve the problem of difficulty, but it is through the tasks that the students are asked to do, that we hope to bring explanation and make the difficult items clearer.

ii) Producing the concordance

The items to be concordanced are selected by the teacher focusing on whatever item s/he believes is relevant to whatever point s/he is teaching. These items could be key lexical items, or features of a structural or rhetorical patterns (in this research, different examples of concordance of NCs are given in 'Chapter Five'). Once the items to be concordanced have been selected, the teacher uses a concordancer to extract all the occurrences of that item or pattern and develop tasks accordingly, so that students can explore meaning and interpretation.

Example 1 below gives a few concordance lines of the node 'blood' from the biology corpus.

	***blood**	
study, the protocol for	**blood**	sampling was similar, but
bodies were analyzed in	**blood**	samples collected in heparinized
the regulation of local	**blood**	flow and electrolyte transport
have variable effects on	**blood**	cholesterol concentrations. In ...

The concordance lines of the selected items are of the '**Key Word In Context'** (KWIC) type. It consists of the key word in the centre of each line (represented by an asterisk) with context to each side of it. With ATA programme, there are four words preceding and four words following the KWIC, as shown above.

Also, if concordance lines are not clear enough to help students explore the meaning of the item under study, students can look at the surrounding text - the larger context (that is the sentence or the paragraph where the item is used) on the computer by clicking on '**view context'** from the ATA menu.

iii) Editing the output

After the selection and extraction of different occurrences of that item are done, we should edit the output by putting off unnecessary items (hapax legomina i.e. one occurrence in each 10,000 words).

The steps described above are necessary for the language lesson production as they form the basis for the development of tasks. Therefore, as part of the 'language focus phase' (described previously), the following tasks can be included in the lessons.

Tasks for the three items (NCs, participles and prepositions), that are frequently used in the corpus, and that present impediments for the students' text comprehension, are developed below.

6.8.4.2 Procedure

After reading the whole text (suggested for study and which title is given in the pre-task phase), the following tasks are proposed:

- **Nominal Compounds**:

The teacher divides the article into different paragraphs and give each pair / group a paragraph and ask them:

a) To locate each of the NCs in the paragraph and turn the compound into a paraphrase, or,
b) to recover the meaning of each NC circled (by the teacher) in the paragraph by writing its expansion in the margin (using word order as a means of disambiguating i.e. to use the principle of right-to-left tracking).
c) Students can also be asked to form NCs.

Teacher gives the students phrases and asks them to turn them into compounds with the following instruction:

"Turn the following phrases into compounds" , examples:

The concentrations of cholesterol in blood;

A diet based on amino acid of casein; etc.

d) A translation exercise

Give the students noun groups in Arabic and / or French and ask them to translate them in NCs in English. The aim from such an exercise is to check whether the students have internalised the process by which we form and read NCs in English.

e) Concordancing:

The teacher gives concordance lines and asks students to find out the meaning of the feature by looking at the surrounding text (the co-text).

For this task, we need two or three computers in the classroom because it has often been stated by ministers, Heads of different institutions, etc. that "*to teach foreign languages, we need adequate teaching materials" (Pour enseigner les langues étrangères, il faut disposer de moyens pédagogiques adéquats).* Because English is seen by the Algerian authorities as the necessary foreign language (see 'Chapter One'), the government should therefore provide the necessary resources.

Another possibility would be that the teachers can simply bring concordance printouts (in the form of handouts) ready to be used in the classroom. The students will first analyse lexical and / or grammatical patterns in examples of the items in a set of concordance lines, and they then complete tasks related to these items.

Some sample tasks using concordances in teaching English could be of the types:

- Sentences with gaps;
- Put in order;
- Find the rule, i.e. deductive teaching;
- Explore differences between forms and /or functions of words;
- Look at synonyms and see how much overlap there is;
- Examining texts for cohesion, etc.

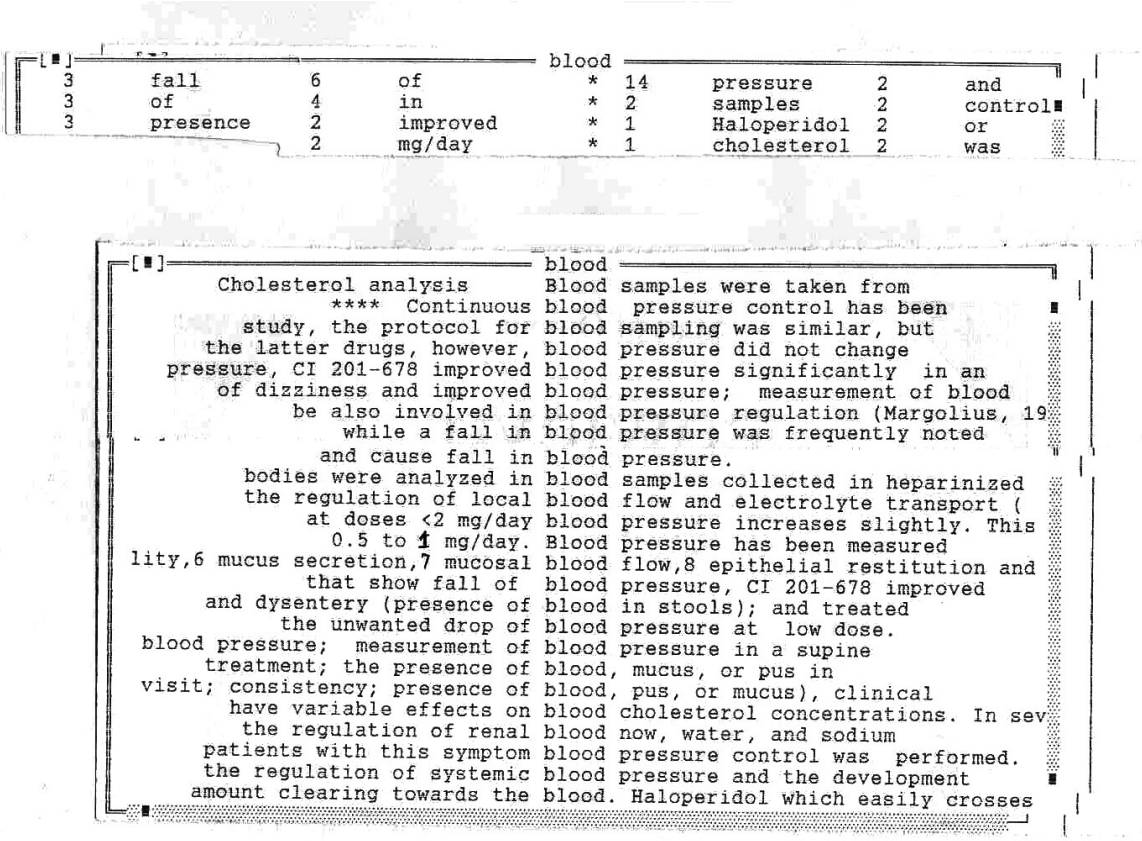
- *The task*

The teacher distributes a copy of the concordance of the item (that s/he believes it to be relevant to a point s/he is teaching), and explains to students how to interpret it.

The teacher assigns each student one or more of the items to prepare. Students (depending on the class size) can have one item each to prepare or they may share (if there are more students than items).

As an example, the teacher distributes concordance lines of the node 'BLOOD' (given below) and asks the students to list all the items that collocate with it.

SYNOPTIC PROFILE



blood

3	fall	6	of	*	14	pressure	2	and
3	of	4	in	*	2	samples	2	control
3	presence	2	improved	*	1	Haloperidol	2	or
		2	mg/day	*	1	cholesterol	2	was

blood

Cholesterol analysis Blood samples were taken from
**** Continuous blood pressure control has been
study, the protocol for blood sampling was similar, but
the latter drugs, however, blood pressure did not change
pressure, CI 201-678 improved blood pressure significantly in an
of dizziness and improved blood pressure; measurement of blood
be also involved in blood pressure regulation (Margolius, 19
while a fall in blood pressure was frequently noted
and cause fall in blood pressure.
bodies were analyzed in blood samples collected in heparinized
the regulation of local blood flow and electrolyte transport (
at doses <2 mg/day blood pressure increases slightly. This
0.5 to 1 mg/day. Blood pressure has been measured
lity,6 mucus secretion,7 mucosal blood flow,8 epithelial restitution and
that show fall of blood pressure, CI 201-678 improved
and dysentery (presence of blood in stools); and treated
the unwanted drop of blood pressure at low dose.
blood pressure; measurement of blood pressure in a supine
treatment; the presence of blood, mucus, or pus in
visit; consistency; presence of blood, pus, or mucus), clinical
have variable effects on blood cholesterol concentrations. In sev
the regulation of renal blood now, water, and sodium
patients with this symptom blood pressure control was performed.
the regulation of systemic blood pressure and the development
amount clearing towards the blood. Haloperidol which easily crosses

Figure 6.2: Selected concordance lines of the node *blood*

From the concordance lines, we notice that the node 'blood' collocates with 'samples', 'pressure' and 'flow', thus forming 2 noun NCs ('blood samples', 'blood pressure' and 'blood flow').
It also forms 3 noun NCs ('blood pressure control', 'blood pressure regulation.')

The student's individual task is to decide what meaning the node 'blood' has in their example(s), and to present the explanation orally to their mates.

This task is considered as a grammatical / stylistic exercise as well as an interpretative one. As such we prepare the students with this specific lexis.

- **Participles (or participial adjectives)**

The teacher and students first discuss the formation of participles, and give their own examples of how participles work and what functions they perform. The teacher writes these examples on the board to elicit the formation and placement of the participles in a sentence. This acts as a structural review and prepares the students for the next task.

Another possibility could be a task example of 'Find the rule', that rather than tell students a particular rule, teachers can give students a set of concordances and ask them to discover what rules can be deduced from the evidence.

- *The task*

The teacher can use some of the traditional exercises such as multiple choice questions (to choose between –ed or –ing endings) or cloze exercises.

- S/He may select a passage from a scientific article dealing with the students' speciality and put blanks in this passage and then ask the students to complete the passage with the appropriate form of the participles (-ed or –ing).

- Re-ordering the left or right context

To increase the students' understanding of the position of these participial adjectives in sentences, concordance output can be used as gap-fill and / or matching-up exercises.

Using a word processor, the teacher can delete for example, right or left collocates of a node, and ask the students to supply these.

The node 'cells' has been selected from the corpus for such a task.

*cells		
human embryonic kidney 293	**cells**	..... (transfected) with APP 695...
parameters of cl 19A	**cells**	..... (cultured) on HAHY Millipore ...
derivatives secreted by sf 9	**cells**	..... (infected) with baculovirus expressing.
in ..... (cultured)	**cells**,	the APP secretase is normally ....;
Microsomes prepared from COS-1	**cells**	..... (tranfected) with ....;
.. then examined the	**cells**	to identify the cellular (transfected)

(The correct answer of the word deleted i.e. the appropriate participial is given between brackets in this research, but it should not be given to students before doing the task in class)

• The teacher can also give the participial in its infinitive form and ask the students to give the –ed or –ing forms.

*protein		
kallikrein with the *(bind)*	**protein**	in both humans and
sequence of the *(purify)*	**protein**	and by sequence comparisons
segment of the **(***translate)*	**protein**	sequence from the cDNA
as that of the	**protein**	*(purify)* from rat sera
for comparison, a	**protein**	(*immunopurify)* from rabbit kidney
a 26 428-Da acidic	**protein**	whose *(derive)* amino acid

• The exercise can also be lexical, i.e. the students can be asked to supply all the adjectives that might modify the node 'protein' and that they already know, and then compare their information with all the adjectives that they can select from the concordance lines.

The most frequent adjectives in the concordance lines that are used to describe the node 'protein' are: mature / microsomal /translated / binding/ derived/ purified/ encoded/ etc.

• Another exercise can be grammatical. In the concordance lines of ‘protein’ the teacher can highlight for example, the following words ‘showed / called /encoded’ and ask the students to determine whether these are ‘verbs’ or ‘participials.’

*protein		
Furhermore, the rK10	protein	**showed** the same molecular
Part of a larger	protein	**called** the amyloid protein
Are conserved in the	protein	**encoded** by rklk10. The ...

- **Prepositions**

What preposition(s) is / are used before / after the node ?

Task 1: Give the node and ask students to fill in the blanks with appropriate prepositions. The nodes ***dna** and ***cells** are taken as examples.

***dna**		
06-methylguanine (in) .... the template	**DNA**	strand have been carried ...
reaction (with) .... all the	**DNA**	present as an enzyme-DNA
DNA is catalyzed (by)	**DNA**	polymerases (with) ..... speed and
Molecules are bound (to)	**DNA.**	
. When the	**DNA**	was digested (with) ..... BamHi
Synthesis (of) ..... the	**DNA**	Primer-Templates

*cells		
the mucosal solution and	**cells**	were assayed ..... (for) cAMP ...
cultivated the	**cells**	..... (on) porous-bottomed dishes ...
activity of the	**cells**	was not affected ..... (by) ...
.The	**cells**	were deproteinised ..... (by) ice ...

(The correct prepositions are given between brackets)

Task 2: Can you identify the missing keyword in each of the following printouts ? The reverse of *Task 1* procedure is to ask the learner to supply the keyword itself. A concordance output can be turned into a gap fill exercise by simply deleting the keyword. This can be done without a word processor (a bottle of correcting fluid will suffice).

- **Modals**

Modals which have been explained in class and focused on, because of their frequency in biology discourse, could also be dealt with in different activities.

• Students could be presented with selected concordance lines (or short paragraphs from the whole article) in which the modal verbs are blanked out, and their task would be to decide which modal(s) suit(s) the blank and what function(s) they perform because of the context.

• Teachers could also provide students with selected concordance lines (or short paragraphs from the article) containing the appropriate modals, and ask students to express the function of these modals.

• Another activity (suggested by Thompson, 2001:202), that could relate the use of the modal auxiliaries to specific rhetorical sections of the text-genre used by biologists, would

be the matching of selected concordance lines into rhetorical section categories, such as 'Introduction, Methods, Results or Discussion' referred to in the previous chapters and explained in 'Section 6.8.3.2 / Modals.' These activities are designed for advanced students.

Specialists are also exploring other ways in which web technology can be effectively used in grammar teaching. Much of this technology is still being developed. A main advantage is that Web Projects such as the IGE (Internet Grammar of English) are available to anyone who has access to the Internet.

According to the authors, IGE is an online resource which aims to describe English grammar mainly from a traditional viewpoint, but making use, where appropriate, of the insights of modern theoretical work. IGE is intended for those who want to learn about English grammar (or improve their existing knowledge of it) through self-study, but is not aimed specifically at learners of English as a Foreign language. Another possible audience are teachers and researchers.

6.9 Final tasks

At the final stage (at a much advanced level, in semesters 6 and 7 of the D.E.S programme, i.e. after students have dealt with the different traditional and more technological tasks e.g. concordancing), we can suggest tasks on problem-solving.

These tasks should be prepared jointly by the language teacher and the subject-speciality teacher, and could be considered as real 'sample community tasks' (Leaver and Willis, 2004:40).

Students in groups work through real world problems with the goal of learning how to apply and synthesise biological concepts and principles, find and evaluate biological information, and communicate ideas and information about biology to others.

The problem-solving tasks, integrated with other class activities, such as the ones suggested previously, will provide the opportunity to develop basic skills and understanding of content.

The content objectives include an understanding of basic biology concepts related to: -

- structure / function relationships in plants and animals,
- adaptation and homeostatis,
- biodiversity,
- relationships between organisms at the population community level,
- relationships between organisms and their environments at the ecosystem level, and
- to evolutionary biology. (as suggested in BISC-208 syllabus, 2000:2).

The aim from learning these concepts and their related principles is to guide the students' thinking in a way that prepares them for more advanced study in biology.

6.9.1 Suggested areas of problem-solving for Algerian biology students

Based on the biology course syllabus at university level, problem-solving tasks may be given to students in the following areas:

- plant reproduction,
- plant physiology,
- plant and animal diversity,
- population and community ecology,
- HIV and the Health Care (Immune system responses to viral invasion, etc.).

For further practice, when / if the Internet is available, both students and teachers are advised to use web-sites dealing with Problem-Based Learning (PBL) such as: http://udel.edu/ deallen/208syll.htm, where they can get a list of problem-solving areas from which they can choose.

6.9.2 Biology problem assignments

Assignments will be given to the class that will assess understanding of major concepts. These assignments may be individual or group, depending on their complexity and/or difficulty. They will consist of either a written summary (1-2 page essay) or of a quiz on the most important concepts introduced by the problems.

Group and individual assignments related to each problem, oral presentation, and group evaluations will provide the opportunity for the students to demonstrate the progress they have made in reaching both the content and the skills objectives.

6.9.3 Evaluation

The evaluation of such tasks should be:

a) joint-evaluation by the language teacher and the subject-specialist teacher

The language teacher should not focus too much on language errors, etc., but on how successful the students' understanding was (reading) and communicating the message (writing). The subject-specialist teacher should focus on how much content has been grasped and the efficiency of the solutions suggested.

b) Evaluation of Individual Performance in Groups

Individual performance in the group will be assessed in writing by both the teacher and by each group member. Students might also be asked to evaluate their own performance as part of this process. These evaluations will be done once during the semester, and can serve as a continuous assessment.

6. 10 Testing

The current practice of testing students in the English component at the 'Biology Department' is done through exams. Testing at the 'Department of Biology' is done twice a year (2 main exams) together with 2 make-up exams (one in June and another in September), following the administration exam timetable.

The exam procedure of the English component is that the teacher gives the students a short text in English (an ESP concocted text, not an authentic one), and ask them to answer some comprehension questions.

What we would suggest in this research is that students be assessed through formal exams (E.M.D 'Epreuve de Moyenne Durée', following the administration requirements 'Section 6.10.1'), and through continuous assessment for the tasks they do in class or as homework ('Section 6.10.2')

6.10.1 Exams

There will be 2 exams of 2 hours each, given throughout the year, on the dates specified in the course schedule. The exams of the English component should be prepared jointly by the language teacher and the subject-specialist. They will cover some aspects of their subject-speciality that they have dealt with in biology lectures and laboratory sessions.

It is generally agreed that similar activities should be used for testing as for teaching, and that topics for achievement tests should be familiar to learners. In other words "*test methods and inputs should be closely linked to teaching and learning experiences.*" (Leaver and Willis, 2004:36).

The main objective of these exams would be to assess the students' ability to understand texts in English and express themselves in writing with some reasonable accuracy, and to solve problems related to their field of study.

Both exams will be a mix of :

- essay questions (that will examine the students' knowledge and understanding of content, their ability to synthesise concepts, and the ability to build convincing evidence to support an argument); and / or
- one or two short problems: the exam questions about the short problems will focus on the process of working through the problem (such as identification of learning issues); and / or
- information transfer tasks (analyses of graphs, tables, pictures of plants/animals, etc. that are common themes to both lectures and laboratory work).

The purpose from information transfer tasks is "*to measure an ability to recognise information relevant to an established purpose or need.*" (Douglas, 2000:203)

6.10.2 Continuous assessment

Teachers can also consider the problem-solving tasks done in class (or as homework) by giving marks or grades, and use them as a continuous assessment, especially when students do well.

6.11 Conclusion

It is hoped, in this chapter, that we have managed to reach the research objective i.e. developing a teaching methodology for Algerian undergraduate students to read scientific texts quickly and efficiently.

We suggested Willis's Framework with its three main cycles:

- The pre-task,
- The task, and
- Language Focus.

This framework (which tried to apply the task underlying principles listed in 'Section 6.7.3.') was then supplemented with the different proposed tasks and activities.

These tasks and activities were ranked from easy (i.e. predicting text content, information transfer, concordancing, ...) to difficult (essay writing dealing with the subject-speciality topics and themes.)

Some pedagogical recommendations were also given to teachers i.e. to develop a teaching programme and activities in collaboration the subject-specialist, to refer to some Internet websites and adapt the possible propositions on how to develop students' reading. The chapter ends with some suggestions for the evaluation of the different tasks and activities, as well as testing students on the English course.

GENERAL CONCLUSION

GENERAL CONCLUSION

As stated in the 'General Introduction', the primary motivation for this study is pedagogic. The aim is give a full description of the situation of teaching English to biology students and to achieve a richer understanding of the nature of texts read by these students, in order to bring improvement in the English course tuition.
The previous chapters have given us an opportunity to deal with the teaching / learning of English in a particular university discourse community; and to analyse the scientific discourse used by this community.

The research questions raised in 'Chapter One' have been answered throughout the present research. With reference to the first question which asks: "How is the discourse community under study organised in terms of teaching / learning practices ?", the investigation carried out throughout this research showed that the biology students need to develop their reading skill in the English language, the latter being socially significant in the effective and efficient conversion of non-members into active members of the target discourse community. However, teaching English to science students is not an easy task for both language teachers and subject-specialists as was explained in 'Chapter Three.'

The study of the teaching / learning situation, through the teachers' and students' questionnaires, revealed that it is confined to the study of 'some' texts written in English followed by some comprehension questions. The English course is seen as 'Question / Response / Evaluation.' Thus, the teacher's role seems to be that of asking questions and evaluating students' answers; and the students' role seems to be that of answering the teacher's questions in writing. No interaction and no real communication seem to take place in the teaching classroom.

The suggested methodology in this research tries to put forward a more dynamic and interactional teaching / learning situation in which learners develop their linguistic competence by finding out the language items (lexis, grammar, etc.) and concepts related to their subject-speciality in order to become more independent (from the teacher) in using English language source material.

The teaching methodology was suggested with the aim to bring some improvement in the English language classroom, and more interaction between the teachers and learners, and between learners and their peers, as all these members form a discourse community.
This language teaching improvement is mainly represented in the introduction of Willis's teaching framework and the development of specific tasks for biology students. The tasks follow some progression from easy to more difficult ones.

With reference to the second question which asks, "How does the English language function in extremely specialised texts ? The element that was introduced in the suggested teaching methodology is the use of authentic materials (scientific texts) and their analyses (manual and computational). Both analyses showed the most frequent features used in biology discourse and explained the complexity of such features to non-native students of English. As far as lexis is concerned, the study focused on 'lexical density' and 'grammatical metaphor' i.e. biology lexis represented in the form of nominal compounds (NCs). For the grammatical features, the study focused mainly on the use of tenses, voice and modals.

These analyses illustrate how quantitative corpus research, can have direct implications for pedagogical practices. Additionally, these corpus analyses have been used to describe more complex patterns, and patterns that would remain hidden while using other analyses.
The investigation of lexico-grammatical structures of biology discourse showed how related grammatical constructions tend to be used with sets of words (e.g. the most common verbs, adjectives, nouns, etc.), as done in the detailed description of the 100 nodes in 'Chapter Five.'

However, the main difficulties encountered in this research are :

- The collection of data for computational analysis (a limited corpus);
- The author's / editor's permission for using such texts for computational analysis.

If a researcher is denied permission for using scientific texts, this will limit his/her corpus;

- Scanning and editing the texts

Scanning and editing texts on a computer is a time-consuming operation, as the scanner sometimes does not recognise some characters of some texts. Depending on the kind of texts, such an operation may be less time-consuming for texts dealing with Economics, for example, than with texts dealing with biology, which contain Greek letters, specific symbols, etc. In this case, the researcher will have to go through each word and add the missing letters or numbers, or any other element.

Recommendations of the present research

This research presents recommendations at three levels: the first and second recommendations are organisational and will be offered to the Ministry of Higher Education in Algeria and the university, and the third recommendations are pedagogical and they are addressed to teachers and classroom situations.

1. **Organisational recommendations of language teaching to the Ministry of Higher Education**

Language planning policy is concerned with the choice of languages to be taught in a country, and the organisation of the teaching of these languages as explained in 'Chapter One'. We also mentioned that English is necessary for Algerian science students who need to consult written publications (available mainly in English) in their field of specialism in order to be informed of and better equipped with the latest developments.

However, the planning of the teaching of English for Algerian science students shows some shortcomings.

a) The time allocated to the English course is not sufficient. The English course in Science Departments / Faculties have always been taught during one session (one and a half hours or two hours) a week.
b) The planning of the English course in the third and fourth years (biology D.E.S. programme) and the second and third years (biology L.M.D programme) does not seem to be pedagogically appropriate because of the two year void (in the D.E.S programme) and one year void (in the L.M.D programme).

As suggested in 'Chapter Six' the number of sessions of the English course for science students should be increased to, at least, two sessions a week (one session for the lecture and the second session for more practice and interaction). Some educationists (Roe, etc.) think that it is better not to have any language teaching than to have one teaching session a week.

The two years and one year void in the English instruction, between secondary school and its introduction in the Biology Department, is not pedagogically appropriate because the students might forget most of what they have learnt previously.

In view of this fact, the proposal made by the current research is to keep Arabic and French as the medium of instruction at the university level in the Arabophone and Francophone Departments, but offer to introduce the English course throughout the four year / three year university education in the D.E.S and L.M.D programmes respectively, as a separate subject, that is an ancillary course alongside the subject-specific courses.

2. Recommendations for science Departments at the university

It is suggested that curriculum design of the language course needs to based on more empirical foundations by carrying out systematic needs assessment This needs assessment allows language teachers to check what had worked out and what had not, and adjust their curriculum accordingly. The shortcoming of the English course is that it remains a separate part from the biology curriculum, not a complementary one.

It is the responsibility of both teachers (language teachers and subject-specialists) and the administration of the department (Head of Department, etc.) to form a team and work seriously on developing teaching programmes (specific courses in different disciplines) so that the future scientists are not deprived of the recent scientific and technical developments published through English. Collaborative effort is needed to make the teaching of English in science departments a successful experience.

We stated in 'Chapter One' that the findings of this research may be implemented in the frame of a research programme. Therefore, the suggestion we could make is to organise:

- an entry test in science departments (to check the students' level at entry) and an exit test (to check the level of students after undergoing the teaching course suggested in 'Chapter Six' of this research), to see whether any improvement has taken place, and to bring any other change if necessary, as advocated by Handy in the 'General Introduction / page 1' to use "*testing until the solution is a new part of our life.*" (1990:50)

- With the advent of the L.M.D system, teachers can also create a self-access centre where students can work independently of the teacher, or with the teacher if s/he is available.

As some of the tasks suggested in 'Chapter Six', e.g. concordancing, etc., referred to the use of computers in the classroom, the authorities should start thinking about creating a computer room in each department so that both teachers and students can use, either for teaching purposes or as a self-access room.

Teachers will have to develop materials dealing with language learning / teaching items (grammar, lexis, etc.) and activities for communicative purposes. Each exercise / task should be supplied with an answer key so that students can correct the activity they have done and evaluate themselves. Self-access activities provide opportunities for self-evaluation and progress in learning.

The self-access centre needs the collaboration of all teachers of the department (language and subject-specialist teachers) for the development of tasks, as mentioned in 'Chapter Six'.

3. Pedagogical recommendations for the language teachers

The last recommendations relate to the role of language teachers in the classroom.

The research findings in 'Chapters Four and Five' would hopefully direct the teaching course to deal with those elements that were found out by using the computational analysis.

These findings showed that many scientific terms are an integral part of the everyday language, but have different linguistic functions and meanings with collocates in a particular discipline e.g. biology.

The third recommendations relate to the third question of this research and which asks, "How can the research findings be dealt with in organising a teaching / learning course for undergraduate biology students ?"

- Therefore the role of the teacher is to raise the students' consciousness by making them aware of this shared understanding by explaining and through task-based teaching, to be able to communicate later with experts from the target discourse community.

- Teachers should also make effort and use 'authentic texts' in the classroom to help the students read what they have to read in order to succeed academically and socially.

- In view of the benefits of computer facilities in language teaching, the collaboration of the language teachers, subject-specialists and computer scientists is also becoming crucial nowadays.

The use of corpora in language teaching, to identify (using specific computer programmes) the elements related to any specific field of research in order to include them in teaching courses, is now developing in many countries (Great Britain, the United States of America, Australia, New Zealand, etc.)

In the classroom, corpora are used to create concordances for data-driven language learning , i.e. identify important elements related to the field of study and develop activities / tasks based on such findings.

For concordancing tasks, the language teacher needs to have some computer science background, especially in programming, that would help him / her develop specific programmes to analyse scientific discourse, and to produce such concordance tasks. If not, we need the collaboration of the computer scientist to develop such a programme.

- Therefore, it is time for the Algerian authorities to provide the necessary conditions to create computer programmes (like A.T.A) to be used locally, or even worlwide.

Such programmes can also be commercially-exploited (to sell such programmes in Algeria and abroad.)

These computer programmes can also have a much wider impact. They are not only implemented in the classroom, but they can also be used to compile dictionaries or glossaries.

Suggestions for further research

- Computer programmes can also be used to compile a scientific dictionary, or at least a monolingual glossary to be used locally by the Algerian students, by selecting the documents and identifying the terms either by using frequency lists or keyword lists (as done in 'Chapter Five').

- A more ambitious implication would be to create '*compara parallel corpus'* (Frankenberg-Garcia, 2005) of English and other languages (e.g. French, Arabic, etc.) to create bilingual scientific dictionaries.

- The above pedagogical implication would also apply to the L.M.D programme at the Department of Foreign Languages which includes a teaching unit called 'Initiation aux langues de spécialité' with its two components 'Typologie des Discours Disciplinaires' and 'Langues de Spécialité'.

The same computer programmes may be used by teachers at the Department of foreign languages to analyse literary texts (or any other text used for pedagogical reasons). The use of such a technique is to ease the task of the teacher by providing him / her with the most frequent elements of the language , topics and themes of the different texts, etc. so that s/he can develop his teaching tasks accordingly.

The investigation carried out in this research does not pretend to have covered all the pedagogical aspects of language teaching to science students , but it is hoped that most biology students will, by the time they graduate, "*have been trained to bring to their work an awareness of science as a key force in social change and economic advance."*

(Frankenberg-Garcia, 2005:195). This means that the students have succeeded in moving from non-members to members of the target discourse community.

Through the accumulated experience and availability of a whole team of researchers, this work can be implemented in an 'action research' perspective, and maybe, make suggestions and remedial proposals, that will ease the transfer of knowledge in biology from English language world users to a national English competence community.

BIBLIOGRAPHY

Abed Rabbo, K., (1996) A Corpus-based learner's Glossary (English-Arabic) for Electrical engineering trainees at Kalandia Training Centre, Palestine, MSc.Thesis, University of Aston in Birmingham.

Ager, D.E. (1984) *French outside France, North Africa, (Algeria),* in "La Francophonie" (A.F.L.S), University of Nottingham, 1984.

Aijmer, K and Altenberg, B. (1991) English Corpus Linguistics, Longman.

Allsop, J. (1988) Cassell's Students English Grammar, Cassell Publishers Limited.

Antaki, C. (1988**)** Analysing Everyday Explanation, SAGE Publications.

Arab, S.A. (1990) *Pour une rénovation de l'enseignement des langues étrangères*, in El Moudjahid Newspaper, Algeria.

Ashiurakis, M. A. (1987) The Influence of the Socio-educational Reading Environment in an Arab University upon English Reading Performance, Unpublished PhD Thesis, University of Aston in Birmingham, U.K.

Aslanian, Y. (1985) *Investigating the reading problems of ESL students: An alternative*, in ELT journal, volume 39/1, January 1985.

Baka, F. (1989) The discourse of biology lectures : Aspects of its Mode and Text Structure, Unpublished PhD Thesis, The University of Aston.

Barakat, H. (1986) A Sociological Investigation, Beirut,

℮ℸ ر ≺♌ ◘حليم ☹ ، íÚ̧ãȨ̑ íÚ?ØȨ̑Ó ËÍÈ(1986) ⑥ راسات الوحدة

Barber, C.L. (1962) Some Measurable Characteristics of Modern Scientific Prose, in Swales, J. (1985) Episodes in E.S.P., Pergamon Press.

Barron, C. (2002) *Problem-solving and EAP: themes and issues in a collaborative teaching Venture*, in English for Specific Purposes, Volume 22, pp.297-314.

Bartolic, L. (1978) *Nominal compounds in Technical English*, in English for Specific Purposes, English Language Institute, Oregon State University.

Bazerman, C. (1980) Shaping written knowledge, the genre and activity of the experimental article in science, The University of Wisconsin Press Ltd.

Beglar, D. and Hunt, A. (2002) Implementing Task-Based Language Teaching.

Belkhenchir, K. (1986) A Study of the Relationship between Verbal and Non-verbal Information in Technical Reports (Gas Treatment), MSc.Thesis, University of Aston.

Benblidia, A. (1985) E.S.P. Syllabus Design for Science Streams in Algerian Secondary Schools: Criteria and Implications, M.Sc. dissertation, Language Studies Unit, University of Aston.

Bennoune, M. (1988) Education and Development, chapter 9, in 'The Making of Contemporary Algeria 1830-1987, Colonial Upheavals and Post-independence Development', C.U.P.

Bernhardt, E.B. and kamil, M.L. (1995) *Interpreting Relationships between L1 and L2 Reading: Consolidating the linguistic Threshold and the linguistic interdependence hypotheses*, in Applied Linguistics, volume 16/1, O.U. P, 1995.

Bhatia, V.K. (1993) Analysing Genre: Language Use in Professional Settings, Longman.

Biber, D. and Conrad, S. (2001) *Corpus-Based Research in TESOL, Quantitative Corpus-Based Research: Much More Than Bean Counting,* in TESOL Quarterly, Volume 35/2.

Biber, D., Conrad, S. and Reppen, R. (1994) *Corpus-based Approaches to Issues in Applied Linguistics*, in Applied Linguistics, volume 15/2, O.U.P, 1994.

Biber, D., Conrad, S. and Reppen, R. (2000) Corpus linguistics, Investigating Language Structure and Use, C.U.P.

Bloor, M. (1998) *English for Specific Purposes: The preservation of the species (some notes on a recently evolved species and on the contribution of John Swales to its preservation and protection) in* English for Specific Purposes, Volume 17, Number 1, pp.47-66.

Boswood, T. and Marriott, A. (1994) *Ethnography for Specific Purposes: Teaching and Training in Parallel,* in English for Specific Purposes, Volume 13/1: 3-21.

Bosworth-Gerome, S. and Marret, R. (1994**)** Ecrire l'Anglais Scientifique et Technique, Ellipses.

Bousquet, G.H., (1961) Les Berbères (Histoire et Institutions), Presses Universitaires de France.

Bowen, E.C. and Schneller, B.R. (1991) Writing about Science, O.U. P.

Bowker, L. and Pearson, J. (2002) Working with Specialized Language, A Practical Guide to Using Corpora, Routledge.

Brierley, W. And Kemble, I.R. (1991) Computers as a tool in language teaching, Ellis Horwood Limited.

Brown, G. and Yule, G. (1983) Discourse analysis, Cambridge University Press.

Bulter, C.S. (1992) Computers and Written Text, Blackwell.

Carrell, P. (1985) *Facilitating ESL Reading by Teaching Text Structure*, in TESOL Quarterly, Volume 19 / 4, December 1985.

Carrell, P. L. Devine, J. and Eskey, D.E. (1988) Interactive Approaches to Second Language Reading, C.U.P.

Carter, R. (1993) Introducing Applied Linguistics, An A-Z GUIDE, Penguin English.

Carter, R. And McCarthy, C. (1988) Vocabulary and Language Teaching, Longman

Chapelle, C.A. (2001) Computer Applications in Second Language Acquisition, Foundations for Teaching, Testing and Research, C.U.P.

Chapelle, C.A. (2003) English Language Learning and Technology, Lectures on Applied Linguistics in the age of Information and Communication Technology, John Benjamins Publishing Company.

Checkland, P. and scholes, J. (1990) Soft Systems Methodology in Action, J. Wiley and sons.

Chimbganda, A.B. (2000) *Communication strategies used in the writing of answers in Biology by ESL first year science students of the University of Botswana*, in English for Specific Purposes, Volume 19/4, pp.305-329.

Church, K.W and Hanks, P. (1990) *Word Association Norms, Mutual Information, and Lexicography*, in Computational Linguistics, Volume 16/1, March 1990.

Clarke, D.F. and Nation, I.S.P, *Guessing the meanings of words from context: Strategy and Techniques,* in System, Volume 8, pp.211-220.

Clarke, M. A., (1994) *The dysfunctions of the theory /practice discourse,* in T.E.S.O.L. Quarterly, volume 28 /1/ Spring 1994.

Cook, G. (2000) Language Play, Language Learning, O.U.P.

Coulthard, M. (1977) An Introduction to Discourse Analysis, Longman.

Crane, D. (1972) Invisible Colleges: Diffusion of Knowledge in Scientific Communities, Chicago London : University of Chicago Press.

Crombie (1985**)** Discourse and Language Learning: A Relational Approach to Syllabus Design, O.U.P.

Cruse, D.A. (1986) Lexical Studies, Cambridge : C.U.P.

Dhieb-Henia, N. (2003) *Evaluating the effectiveness of metacognitive strategy training for reading research articles in an ESP context*, in English for Specific Purposes, Volume 22/4, pp.387-417.

Djité, P. (1992) *The arabization of Algeria: Linguistic and Sociopolitical Motivations,* in International Journal of Sociology of Language, Number 98, pp.15-28.

Douglas, D. (2000) Assessing Languages for Specific Purposes, C.U.P.

Dudley –Evans, T. (1994) Genre Analysis : An approach to Text Analysis for E.S.P, Chapter 14, in Coulthard , M. 1994, Advances in Written Text Analysis, Routledge.

Dudley-Evans, A. And St John, M.J. (1998) Developments in E.S.P: a multi-disciplinary approach, Cambridge University Press.

Ebel, Bliefert, and Russey (1987) The Art of Scientific Writing, VCH.

Ellis, R. (2003) Task-based Language Learning and Teaching, O.U.P.

Ellis, R. (1985) Understanding Second Language Acquisition, 13th Impression, O.U.P.

Engwall, G. (1994) Not chance but Choice: Criteria in Corpus Creation, pp. 49-82, in Atkins, B.T. S and Zampolli, A., Computational Approaches to the Lexicon, Clarendon Press Oxford.

Fairclough, N. (1992) Discourse and Social Change, Polity Press.

Fetni, A. (1987) Literacy and post-literacy in the Framework of Continuing Education: The Algerian Experience, in **'**Learning Strategies for Post-literacy and Continuing Education in Algeria, Egypt and Kuwait, Editors : Dave R.H., Ouane A. and Ranaweera A.M., UNESCO Institute for Education, 1987.

Flowerdew, J. (1993) *An Educational, or Process, Approach to the Teaching of Professional Genres*, in 'ELT Journal, Volume 47/4, October 1993, O.U.P.

Flowerdew, J. (1993) *Content-based language instruction in tertiary setting*, in English for Specific Purposes, Volume 12,pp.121-138, Pergamon Press Limited.

Frankenberg-Garcia, A. (2005) *Pedagogical uses of monolingual and parallel concordances*, in ELT Journal, Volume 59/3, O.U.P.

Freedman, A. (1993) *Show and Tell ? The Role of Explicit Teaching in the Learning of New Genres, in* Research in the Teaching of English, Volume 27/3, October 1993.

Freitas, A.C. *Considerations about the lexical aspect of a 'genre'*, in The Especialist, Volume 12 / numbers 1&2, pp.145-160.

Garside, R., Leech, G. and Sampson, G. (1987) The computational Analysis of English, a corpus-based approach, Longman.

Gee, J.P. (2005) An Introduction to Discourse Analysis, Theory and method, Second Edition, Routledge.

Geertz, C. (2000) Available Light, Anthropological Reflections on Philosophical Topics, Princeton University Press.

Gledhill, C. (1995) Scientific Innovation and the Phraseology of Rhetoric, Posture, Reformulation and Collocation in Cancer Research Articles, PhD Thesis, Aston University.

Gledhill, C. (1996) *Collocation and Genre Analysis*, in zeitschrift für Anglistik und Amerikanistik, Volume XLII 1/1, pp.11-36, stauffenburg verlag, Tùbingen.

Goodman, K.S. (1967) *Reading: a psycholinguistic guessing game*, in Journal of the Reading Specialist, 6:126-135.

Gordon, D.C. (1966) The Revolution and Culture (Chapter 8), in The Passing of French Algeria, O.U.P.

Graham, W.S. (1981) The Readability of Materials in Biology, M.Phil Thesis, Aston University.

Greenall, S. and Swan, M. (1986) Effective Reading, Reading skills for advanced students, Cambridge University Press.

Grellet, F. (1981) Developing Reading Skills, A practical guide to reading comprehension exercises, Cambridge University Press.

Gunawardena, C.N. (1989) *The Present Perfect in the Rhetorical Divisions of Biology and Biochemistry Journal Articles*, in English for Specific Purposes, Volume 8, pp.265-273.

Hadj Moussa, A. (1993) A discourse-Oriented Approach for the teaching of English to biology students at the University of Annaba, Magister Thesis, University of Annaba.

Halliday, M.A.K. and Martin, J.R. (1993) Writing Science: Literacy and Discursive Power, The Falmer Press.

Hanania, E. and Akhtar, K. (1985) *Verb Form and Rhetorical Function in Science Writing: A Study of Master's Theses in Biology, Chemistry and Physics*, in English for Specific Purposes 4, pp. 49-58.

Handy, C. (1990) The Age of Unreason, Arrow Books Ltd.

Hayane, O. (1989) L'Enseignement de la Langue Anglaise en Algérie depuis 1962, Office des Publications Universitaires (O.P.U.), Algérie.

Herbert, A.J. (1965) The Structure of Technical English, Longman, in Swales, J. (1985) Episodes in E.S.P., Pergamon Press.

Herzberg, (1986) in Swales, J. (1990) Genre Analysis, English in academic and research settings, C.U.P.

Hewings, A. and Hewings M. (2005) Grammar and Context, an Advanced Resource Book, Routledge.

Hoey, M. (2005) Lexical Priming, A new theory of words and language (Chapters 1&8) Routledge.

Hudson, A. (1993) Journal of Pragmatics, Volume 20, Number 5, November 1993.

Hunston, S. (2002) Methods in Corpus Linguistics: Interpreting concordance lines (Chapter 3) and Corpora and Language Teaching: General applications (Chapter 7), in Corpora in Applied Linguistics, C.U.P.

Hunston, S. Gill, F. And Manning, E. (1997*) Grammar and Vocabulary: Showing the connections*, in E.L.T. Journal, Volume 51/3, July 1997.

Hutchinson, T. and Waters, A. (1987) English for Specific Purposes, A learning-centred approach, C.U.P.

Hyland, K. (1992) *Genre Analysis: Just Another Fad*, in English Teaching Forum, April 1992.

Johns, A.M. and Dudley-Evans, T. (1991) *English for Specific Purposes: International in Scope, Specific in Purpose*, in TESOL Quarterly, Volume 25/2, 1991.

Johns, T. (1991) *From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-Driven Learning*, in ELR Journal, Volume 4, The University of Birmingham.

Johns, T. (1994) The text and its Message, Chapter 7, in Coulthard, M. (1994) Advances in written text analysis, Routledge, London.

Kennedy, C. (1983) Language Planning and Language Education, Allen, G. and Urwin.

Kennedy, C. (1998) An Introduction to Corpus Linguistics, Longman

Kjellmer, G. (1990) Patterns of Collocability, in Aarts, J. and Meijs, W., Theory and Practice in Corpus Linguistics, Rodopi.

Kneller, G.F. (1978) The Scientific Community (chapter 8) Columbia University Press, N.Y.

Knowles, F. (1995) in Knowles, F and Roe, P. (1995) Lexical Studies File, Language Studies Unit, University of Aston.

Knowles, F. (1995) *"Lexical cartography" in LSP texts*, Aston university, Birmingham.

Knowles, F. (1995) "*Collocability in Languages for Special purposes (LSPs): some preliminaries."*

Kocourek, (1982) La langue Française de la Technique et de la Science, Brandsletter Verlag.

Kowitz, J. And Carroll, D. (1991*) Using Computer Concordances for Literary Analysis in the Classroom*, in ELR Journal, Volume 4, Classoom Concordancing.

Kumaravadivelu, B. (1991*) Language-learning Tasks: Teacher intention and learner interpretation,* pp. 98-107, in E.L.T. Journal, Volume 45/2, April (1991) O.U.P.

Latorre, G. and Pons, H. (1986) *Pedagogic Models of the English Complex Noun Phrase*: A *Comparison between Registers*, pp.99-108 in E.S.P. in Practice, Models and Challenges for Teachers, General Editor, Wilcox Peterson, P.

Lave, J. and Wenger, E. (1991) Situated learning: Legitimate peripheral participation. Cambridge: Cambridge University Press.

Leaver, B.L. and Willis, J.R. (2004) Task-based instruction in foreign language education, Practices and Programs, Georgetown University Press.

Le Corbusier, in Collins Concise Dictionary, 1998, Harper Collins Publishers, fourth Edition.

Leech, G., Rayson, P. and Wilson, A. (1998) Word Frequencies in Written and Spoken English (based on the British National Corpus), Longman.

Le Quotidien d'Oran, A daily newspaper, Monday October 30th 2006.

Lerat, P. (1995) Les Langues Spécialisées, Paris : Presses Universitaires de France.

Lewis, M. (1993) The Lexical Approach, The State of ELT and a Way Forward, Language Teaching publications, England.

Locastro, V. (2001) *Large classes and student learning*, in TESOL Quarterly, Volume 35/3, pp.493-496.

Malavieille, M., Quivy, M. and Rotgé, W. (1999) Collection Bescherelle Exercices, Anglais, Hatier.

Marchand, H. (1991) Lecture Raisonnée de l'anglais scientifique, Ellipses.

Marco, M.J.L. (2000) *Collocational frameworks in medical research papers: a genre-based study*, in English for Specific Purposes, Volume 19:63-86.

Mavor, S. and Trayner, B. (2001) *Aligning genre and practice with learning in Higher Education: an interdisciplinary perspective for course design and teaching*, in E.S.P. Journal, Volume 20/4, pp.345-366.

McCarthy, M. and Carter, R. (1994) Language as Discourse, Perspectives for Language Teaching, Longman.

McEnergy, T. and Wilson, **A**. (1996) Corpus Linguistics, Edinburgh University Press.

Meyer, C. F. (2002) English Corpus Linguistics, An Introduction, C.U.P.

Miliani, M. (1996*) The Circulation of European Educational Theories and Practices: The Algerian Experience*, pp. 1-12 in Mediterranean Journal of Educational Studies, volume 1 (1).

Miliani, M. (2003) *Arabisation of Higher Education in Algeria : Linguistic Centralism Vs Democratisation*, in International Journal of Contemporary Sociology, Volume 40, Number 1, April 2003.

Miller, C. (1984) *Genre as social action, in* Quarterly Journal of Speech 70: pp. 151-167.

Milroy, L. (1980) Language and Social Networks.

Mindt, D. (1996) English corpus linguistics and the foreign language teaching syllabus, in Jenny, T. and Short, M. (1996) Using corpora for language research, Longman.

Moirand, S., Des Choix Méthodologiques pour une Linguistique de Discours Comparative.

Mparutsa, C. Love, A. and Morrison, A. (1991) *Bringing concord to the E.S.P. classroom*, in ELR Journal, Volume 4, Classroom Concordancing, pp.115-133, The University of Birmingham.

Muller, C. (1963) Le MOT, unité de texte et unité de lexique en statistique lexicologique, in Travaux de Linguistique et Littérature, pp. 155-173.

Myers, G. (1988) *The Social Construction of Science and the Teaching of English: An Example of Research*, in ELT Documents: 129, Modern English Publications.

Myers, G. (1990) Writing Biology: Texts in the construction of social knowledge, Wisconsin: University of Wisconsin Press.

Myers, G. (1992) *Textbooks and the Sociology of Scientific Knowledge*, in English for

Specific Purposes, Volume 11, pp.3-17.

Myers, G. (1994) Narrative of Science and Nature in Popularizing Molecular Genetics, in Coulthard, M., 1994, Advances in Written Text Analysis, Routledge.

Nagel's Encyclopedia-Guide (1973) Algeria, pp. 76-77, Nagel's Publishers.

Nattinger, J.R. and Decarrico, J.S. (1992) Lexical Phrases and Language Teaching, O.U.P.

Needham (1967:47) in Phillips, M. (1985:121), Aspects of Text Structure: An investigation of the lexical organisation of text, Elsevier Science Publishers.

Nessehauf, N. (2003) *The Use of collocations by advanced learners of English and some implications for teaching*, in Applied Linguistics 24/2, pp.223-42.

Nitta, R. and Gardner, S. (2005) *Consciousness-raising and practice in ELT coursebook*s, in ELT Journal, Volume 59/1, pp.3-11.

Nuttall, C. (1982) Teaching Reading Skills in a Foreign Language, Heinemann Educational Books, London.

Ooi, D. and Kim-Seoh, J.L (1996) *Vocabulary teaching: looking behind the word*, in ELT Journal, Volume 50/1 January 1996, O.U.P.

Oukil, M.S. and Damou, M. (2002) *Towards a Better Quality Education in Engineering and Technology in Developing Countries*, in Sciences Humaines n°18, Décembre 2002, pp.19-26, Université Mentouri, Constantine, Algérie.

Ourghi, R. (2002) The Implications of Critical Reflection and Discourse Awareness for Curriculum, Learner, and Teacher Development: The Case of University-Level Writing, Doctorate Thesis, University of Aboubakr Belkaid, Tlemcen.

Owen, C. (1996) *Do concordances require to be consulted* ? in ELT Journal, Volume 50/3, July 1996, O.U.P.

Pal, A. (1989) *Problem Posing in the EFL/ESL Classroom*, in The Especialist, volume 10/2: 119-148.

Paltridge, B. (1994) *Genre Analysis and the Identification of Textual Boundaries*, in Applied Linguistics, Volume 15/3, O.U.P.

Paltridge, B. (1996) *Genre, text type, and the language learning classroom*, in ELT Journal Volume 50/3 , July 1996, O.U.P.

Peacock, M. (1997) *The Effect of authentic materials on the motivation of EFL learners*, in ELT Journal, Volume 51/2, April 1997.

Phillips, M. (1985) Aspects of Text Structure: An Investigation of the Lexical Organisation of Text, Elsevier Science Publishers.

Pollack, R. (1994) Signs of Life, the language and meanings of DNA, penguin books.

Potts, P. (1988) Algerian Universities E.S.P. project : the Role of Research in this project.

Prakasam, V. (1996) *"NGp of NGp" constructions: A Functional-Structural Study, pp. 567-583*, in Halliday M.A.K. (1996) Meaning and Form: Systemic Functional Interpretations, in Advances in Discourse Processes.

Purvis and Hunt (1993) *Discourse, ideology, discourse, ideology, ideology,…*, in The British Journal of Sociology, Volume 44, Issue 3, pp. 473-499, September 1998.

Renouf, A and sinclair, J. (1991) Collocational Frameworks in English, in Aijmer, K and Altenberg, B. (1991) pp. 128-143, English Corpus Linguistics, Longman.

Richards, J., Platt, J. And Weber, H. (1985) Longman Dictionary of Applied Linguistics, Longman.

Richards, J.C. (2001) Curriculum Development in Language Teaching, C.U.P.

Richards, J.C. and Renandya, W.A. (2002) Methodology in Language Teaching, an Anthology of Current Practice, C.U.P.

Richards, J.C. and Rodgers, T.S. (1986) Approaches and Methods in Language Teaching, C.U.P.

Riley, P. (1985) Discourse and Learning, Longman.

Robinson, P. (1980) E.S.P: English for Specific Purposes, Pergamon Press.

Rodrigues Faria coracini, M.J., L'Hétérogénéité dans le Discours Scientifique Français et Brésilien : Un effet persuasif.

Roe, P. (1977) The Notion of Difficulty in Scientific Text, PhD Thesis, University of Birmingham; U.K.

Roe, P. (1993) A Doctoral Summer School lecture, Language Studies Unit, University of Aston.

Romer, U. (2004) A corpus-driven approach to modal auxiliaries and their didactics, in Sinclair, J.McH (2004) How to use Corpora in Language Teaching, John Benjamins Publishing Company.

Rossner, R. and Bolitho, R. (1990) Currents of Change in English Language Teaching, O.U.P.

Rosenthal, J.W. (1996) Teaching Science to Language Minority Students, Multilingual Matters Ltd.

Saad, M. (1992) *Année Scolaire 1992-93, Nouvelles Méthodes de Travail*, in El Watan Daily Newspaper, Algeria, Sunday 13th September 1993.

Sager, Dungworth and McDonald (1980) English Special Languages: Principles and Practice in Science and Technology , Brandsletter Verlag.

Sager, J.C. (1990) A Practical Course in terminology Processing, John Benjamins Publishing Company.

Salager, F. (1984) Compound Nominal Phrases in Scientific-Technical Literature: Proportion and Rationale, in Pugh and Ulijn (eds.) (1984), Reading for Professional Purposes, London : Heinemann.

Salager-Meyer, F. (1990) *Discoursal flaws in Medical English Abstracts: A genre analysis per research – and text-type, p*p.365-384, in TEXT (1990) Volume 10, Number 4.

Saville-Troike, M. (1982) The Ethnography of Communication, Oxford: Basil Blackwell.

Schiffrin, D. (1994) Approaches to Discourse, Blackwell.

Scott, M. (1991) Social Network Analysis: a handbook, London, Sage Publications.

Siddiqui, N.A. (1992) Effects of teaching and validity of testing reading comprehension skills in ESL context, M.A. in T.E.F.L., Centre for Applied Language Studies, University of Reading.

Sinclair, J. M (1987)) Looking up, An account of the COBUILD Project in Lexical Computing, Collins ELT.

Sinclair, J.M. (1991) Corpus concordance collocation, O.U.P.

Sinclair, J.M. (1995) corpus, concordance, collocation, (3rd impression), O.U.P.

Sinclair, J.M. (2004) How to use corpora in language teaching, John Benjamins Publishing Company.

Slougui, D. (1994) Qualifying PhD Report, Language Studies Unit, University of Aston.

Spencer, J. (1963) Language in Africa, C.U.P.

Spillner, B., Textes Médicaux Français et Allemands, Contribution à une comparaison inter-linguale et interculturelle.

St-John, M.J. (1989) *Thoughts on genre-based writing and reading courses,* in The Especialist, Volume 10/2: 167-179.

Stubbs, M. (1986) *Lexical Density: A Technique and Some findings,* pp.27-42 in Coulthard, M. (1986) Talking about Text, ELR, University of Birmingham.

Stubbs, M. (1994) *Grammar, Text and Ideology: Computer-assisted Methods in the Linguistics of Representation*, in Applied Linguistics, volume 15/2, O.U.P.

Stubbs, M. (1996) Text and Corpus Analysis, Computer-assisted Studies of Language and Culture, Blackwell.

Swales, J. (1981) Aspects of Article Introductions, Aston E.S.P. Research Reports N°1, Language Studies Unit, The University of Aston in Birmingham.

Swales, J. (1985) Episodes in E.S.P., Pergamon Press.

Swales, J. (1990) Genre analysis, English in academic and research settings, C.U.P.

Swales, J. (1993) *"Genre and Engagement",* Revue Belge de Philologie et d'Histoire 71(3):687-698.

Swales, J. and Fredrickson (1993) *Competition and Discourse Community: Introductions from Nysvenska studier.*

Taleb Ibrahimi, A. (1981) De la décolonisation à la révolution culturelle (1962-72), SNED, Alger, 3ème Edition.

Tarone, E. Dwyer, S. Gillette, S. and Icke, V. (1998) *On the Use of Passive and Active Voice in Astrophysics Journal Papers: With Extensions to other languages and other Fields*, in English for Specific Purposes, Volume 17/1, pp.113-132.

Taylor, S.J and Bogdan, R. (1984) Introduction to Qualitative Research Methods.

The British Council (1984) English Language and Literature Division, English Teaching Profile, Algeria.

Thom, R. (1972) Stabilité Structurelle et morphogenèse, Massachussetts: W.A. Benjamin, Inc.

Thom, R. (1975) Structural Stability and Morphogenesis, translated by Fowler and Waddington, W.A. Benjamin, Inc.

Thompson, P. (2001) A pedagogically-motivated corpus-based examination of PhD theses: macrostructure, citation practices and uses of modal verbs, PhD thesis, September 2001, The University of Reading.

Thompson, S. (1994*) Frameworks and Contexts: A Genre-based Approach to Analyzing Lecture Introductions*, pp. 171-186, in English for Specific Purposes Journal, volume 13, number 2, 1994.

Thouvenin, S. (1996) The Identification and Exemplification of Multi-word Units within a Technical Corpus of English, Including an Investigation of Nominal Groups, MSc; Dissertation, Language Studies Unit, University of Aston.

Thurstun, J. And Candlin, C.N. (1998) *Concordancing and the Teaching of the Vocabulary of Academic English*, In English for Specific Purposes, Volume 17/3, pp.267-280.

Tovar, L. (1993) *Grammatical Difficulties of Advanced ESP Reading Students*, The Especialist, volume 14/2: 127-146.

Treffgarne, C. (1994) *Bonjour l'Afrique*, in E.F.L. Gazette, July 1994.

Tribble, C. and Glyn, J. (1990) Concordances in the Classroom, Longman.

Trimble, L. (1979) Rhetorical, Grammatical and Lexical Elements of Written ESP, Mimeo.

Trimble, L. (1979) *A Rhetorical Approach to Reading Scientific and Technical English*, in English Teaching Forum, Volume 17/4, October 1979.

Trimble, L. and Drobnic, K. (eds), English for Specific Purposes: Science and Technology, Corvallis, Oregon.

Warren, K.S. (1981) Coping with Biomedical Literature, A Primer for the Scientist and the Clinician, Praeger Publishers.

Wenger, E. (1998) Communities of Practice: Learning, meaning and Identity. Cambridge: Cambridge University Press.

Widdowson, H.G. (1979) Explorations in Applied Linguistics, O.U.P.

Widdowson, H.G. (1990) Aspects of Language Teaching, O.U.P.

Widdowson, H.G. (2003) Defining issues in English Language Teaching, Chapters 6&9, O.U.P.

W**illiams, E. And Moran, C. (**1989) *Reading in a foreign language et intermediate and advanced levels with particular reference to English,* in Language Teaching 22/4:217-228.

Williams, I. A. (1996) *A contextual study of lexical verbs in two types of medical research report: clinical and experimental,* in English for Specific Purposes, Volume /3:175-197.

Williams, R. (1984) A Cognitive Approach to English Nominal Compounds, in Pugh and Ulijn (eds.). (1984) Reading for Professional Purposes, London: Heinemann.

Willis, D. (1990) The Lexical Syllabus: A New Approach to Language Teaching, Collins Cobuild.

Willis, J. (1996) A Framework for Task-Based Learning, Addison Wesley Longman Limited.

Willis, J. and Willis, D. (1996) Challenge and Change in Language Teaching, London: Heinemann.

Wilson, B. (1990) Systems: Concepts, Methodologies, and Applications (Second Edition)

Wingard, P. (1981) Some verb forms and functions in six medical texts, pp.53-64, in Selinker, L. Tarone, E. And Hanzeli, V., English for Academic Purposes: Studies in honor of Louis Trimble. Rowley M.A.: Newbury House.

Zappen, J.P. (1989) *The Discourse Community in Scientific and Technical Communication: Institutional and Social Views,* in Journal of Technical Writing and Communication, Volume 19/1, pp.1-11, Baywood Publishing Company.

Official Documents

« Conférence Nationale de la Formation Supérieure », Document de Travail de la Commission 3, Arabisation, 7 /8 / 9 july 1987.

Ministère de l'Enseignement Supérieur (1986) Direction des Enseignements, Développement de l'Enseignement des Langues Etrangères, Algérie.

Ministère de l'Enseignement Supérieur (1992) Direction des Enseignements, Enseignement de l'Anglais Scientifique et Technique.

Ministry of National Education (1984) Pedagogical Instructions, Inspectorate of English, Algiers.

Ministry of National Education (2006) New schedule for teaching English in secondary school, March 2006.

Conference Papers

Biber, D. (2001) Using Corpus-Based Methods to Investigate Grammar and Use: Some Case Studies on the Use of Verbs in English, pp.101-115, in Simpson, R.C. and Swales, J. (2001) Corpus Linguistics in North America, Selection from the 1999 Symposium, The University of Michigan Press.

Cherrad-Benchefra, Y. (1987) La réalité Algérienne: Comment les problèmes linguistiques sont vécus par les Algériens, in 'Langage et Société', Bulletin 41, 1987, Colloque: Contacts de Langues : Quels modèles ?

Gavioli, L. (2002) Some thoughts on the problem of representing ESP through small corpora,in Kettemann, B. and Marko, G. (2002) Teaching and Learning by Doing Corpus Analysis, Proceedings of the 4th International Conference, on Teaching and Language Corpora, Graz 19-24 July 2000, Editions Rodopi.

Gee, 1990, **in Swales (1992)** Genre and Engagement, A paper presented at Nottingham University conference, August 1992.

Granath, S. (1998) Using Corpora in Teaching English Syntax to EFL Students at the University Level, Teaching And Language Corpora (TALC, 24th –27th July 1998), Keble College, Oxford, Longman.

Hunston, S. (1995) Ideology, Genre and Text in Systemic Linguistics, BAAL/CUP Genre Analysis Workshop, Sheffield, July 1995.

Scott, M. (1998) Focusing on the Text and its Key Words, Teaching And Language Corpora (TALC) 24-27 July 1998, Keble College, Oxford, Longman.

Burnard, L. And McEnery, T. (2000) Rethinking Language Pedagogy from a Corpus Perspective (Papers from the Third International Conference on Teaching and Language Corpora, Lang P., in £ODZ Studies in Language, Volume 2.

Simpson, J. (2004) Language Learning in a virtual world: Lessons from an online language learning community, A paper presented at the International Conference on "Languages and Modernity", University of Oran (12-15 march 2004) in Passerelle / Numéro Spécial, LOAPL.

Swales, J. (2003) Is the University a community of Practice ? in British Studies in Applied Linguistics 18, Applied Linguistics and Communities of Practice, Selected papers from the Annual Meeting of the British Association for Applied Linguistics, Cardiff University, September 2002, Editors: Sarangi, S. and Van Leeuwen, T.

Willis, D. (1993) Syllabus, Corpus and Data-Driven Learning, IATEFL Annual Conference Report, pp.25-32.

Websites

Aarts, B., Nelson, G. and Buckley, J. , The Internet Grammar of English: New horizons in grammar pedagogy, University College, London
IGE-site: **http://www.ucl.ac.uk** /internet-grammar

BISC-208 Syllabus, Department of Biological Sciences (2000) at:
http://udel.edu/ deallen/208syll.htm

http://www.edict.com.hk/vlc/grammar course/lesson2_preposition/fixed.htm

Gunert, J. (1997) The Course Syllabus: A Learning-Centered Approach, Bolton, Massachusetts: Anker Publishing Company, Inc. at:
http://cte.udel.edu/syllabus.htm

http://www.edict.comhk/vlc/Study Guide/concordance-examples.htm

de Szendeffy, J. (2004) Vocabulary and Usage Activities with Concordances, Adapted from Computer-assisted Language learning : A practical Guide for Teachers, John de Szendeffy, University of Michigan Press, at:
http://www.bu.edu/celop/mll/tutorials/pdf_public/concordance.pdf

Sinor, J. And Kaplan, M. (2004) Creating Your Syllabus at:
http://www.crlt.umich.edu/crlttext/P_21text.html

http://www.nsknet.or.jp/ peterr-s/concordancing/usingconcs.html

http://www.nsknet.or.jp/ peterr-s/concordancing/glossary.html

Swales, J. (1998) Other Floors, Other Voices: A Textography of a Small University Building, Mahwah NJ, and London: Lawrence Erlbaum Associates at:
http://bowland-files.lancs.ac.uk/staff/greg/genre/SWALES.htm

APPENDICES

Appendix A	The teacher questionnaire
Appendix B	The student questionnaire
Appendix C	Titles of the articles in the biology corpus
Appendix D	List ‘Frequent non-grammar’ of the biology corpus
Appendix E	A selected article as a sample for use
Appendix F	Synoptic profiles and Concordance lines of nodes ‘Protein’ and ‘Acid’

APPENDICES A and B

THE TEACHER and THE STUDENT QUESTIONNAIRES

THE QUESTIONNAIRES

Dear Teachers / Students

I am currently engaged in preparing a Doctoral Thesis, and I would be most grateful if you would complete and return this questionnaire in order to assist the department of biology with its English course review. Your comments are very valuable to make any changes when we come to revise this course and to elicit other information which might help solving some of the problems you face in teaching / learning English.

Your responses will be confidential. Thank you for your co-operation.

Miss K. BELKHENCHIR .
Section of English
University of Oran.

APPENDIX A

QUESTIONNAIRE FOR THE E.S.P. TEACHER

PART A: The E.S.P. Teacher Background:

1) Sex : □
2) Age : □
3) Department : □
4) How many years have you been teaching English? □
5) Are you a language teacher or a subject –specialist? □

General instructional objectives of the E.S.P. Course :

6) What is the course main objective? (Give a number to show order of importance).

   a- Reading □
   b- Writing □
   c- Speaking □
   d- Listening □

7) What is the teaching load supplied to reach the course objective?

   Hours per week □

8) Is the number of teaching hours to reach the course objective?

   a- Sufficient □
   b- Insufficient □

Give reasons :

_____

Suggestions :

_____

Part B: The Teachers' Methodology

A. Instructional objectives of the reading component :

9) Do you teach reading to :

- a- Improve speed? ☐
- b- Improve comprehension? ☐
- c- Improve both, speed and comprehension? ☐
- d- Other (specify) ☐

10) Is reading taught to explain :

- a- How the English language works ☐
- b- How scientific experiments were done ☐
- c- How processes work ☐
- d- Specific terminology ☐
- e- Other (specify) ☐

Methodology :

11) Which approach do you use to teach reading?

(Give as much detail as you can).

_____

_____

Teaching materials :

12) How are the syllabus selected texts?

- a- Too specialised ☐
- b- Suitably mixed ☐
- c- Too general ☐
- d- Other (specify) ☐

13) What types of exercises your students have to do in the reading class?

- _____
- _____
- _____

14) Are the materials you use to teach reading :

a- In house. □
b- Published □
c- Other (specify) □

Evaluation

15) How do you assess reading?

_____

__

_____

_____

____

16) What are the students'main areas of difficulty in reading?

a- Grammar (tenses, concord, connectors, etc.) □
b- Specific terminology (= scientific terms) □
c- General vocabulary □
d- Non-verbal information (symbols, formulae, diagrams…) □
e- Other (specify) □

B. Instructional objectives of the writing component :

- Time allotted to the Academic writing component :

17) How often do your students have written exercises during class time?

a- Regularly □
b- Occasionally □
c- Never □

Methodology :

18) Which of the following sub-skills do you teach in writing?

(Give a number to show order of importance)

a- Writing summaries and paragraphs □
b- Report writing □
c- Information transfer (from diagrams, tables… to text) □
d- Writing essays □
e- Other (specify)

19) How do you teach the three most important sub-skills?

1- _____

_____

____

2- _____

_____

____

3- _____

__

Teaching materials :

20) Which materials do you use to teach Academic writing?

a- In-house. □
b- Published. □
c- Other (specify) □

Evaluation :

21) How do you assess writing?

_____

_____

____

_____

_____

____

_____

__

C. Instructional objectives of the listening component :

22) How often do you teach listening?

a- Regularly □
b- Sometimes □
c- Never □

23) Which of the following sub-skills do you teach in listening?

d- Note-taking. □
e- Summarising □
f- Outlining □
g- Commenting on…. □
h- Other (specify) □

Methodology :

24) How do you teach listening?

25) What are the materials you use when you teach listening?

D. Instructional objectives of the speaking component

26) How often do you teach speaking?

a) Regularly □
b) Sometimes □
c) Never □

Methodology :

27) Why do you teach speaking? (aims and objectives)

28) How do you teach speaking?

_____

_____

Teaching materials :

29) What materials do you use to teach speaking?

_____

_____

_____

30) What are your teaching difficulties in :

Reading	Writing	Listening	Speaking
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

Suggestions to solve problems in :

Reading	Writing	Listening	Speaking
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

Thank you for your co-operation.

APPENDIX B

QUESTIONNAIRE FOR THE 4TH YEAR BIOLOGY STUDENTS

Please read all the instructions and questions carefully. Try to answer all the questions and be free to comment.

PART A / RUBRIC 1 : THE STUDENTS' BACKGROUND

Please circle M or F to indicate your gender. M F

Age : Between 20 and 25 Over 25

Medium of instruction in your department

Arabic French

What foreign language did study at middle and secondary schools?

Spanish English German Other

How long have you been studying English?

Number of years □

Is your English level proficiency?

Advanced Intermediate Beginner

PART B / RUBRIC 2 : GENERAL INSTRUCTIONAL OBJECTIVES OF THE ENGLISH COURSE AT THE DEPARTMENT OF BIOLOGY.

1. When is English introduced in the programme in your deparment?

1st □

2nd □

3rd □

4th □

2. Type of English classes :

Literary English ☐
Scientific English ☐

3. Is English necessary in your field of study?

Yes, Why? Because.

No, Why? Because.

4. According to you, what is the main objective of the course? (Give a number to show order of importance : 1 the most important → 4 the least important)

Reading ☐
Writing ☐
Speaking ☐
Listening ☐

5. How many hours of English per week do you have?

1st year hours/week
2nd year hours/week
3rd year hours/week
4th year hours/week

Length of each class ☐

6. Was there a suitable number of English lectures on the course?

Far too many				Far too few
5	4	3	2	1

PART C / RUBRIC 3 : THE ENGLISH COURSE

1. Was the course well organised?

Contentwise ☐
Duration ☐
Activities ☐
Tasks ☐
Other (specify) ☐

2. What did the English course consist of?

- Reading comprehension ☐
- Vocabulary exercises ☐
- Grammar exercises ☐
- Writing exercises ☐
- Teaching your subject speciality ☐

3. Did you understand your English teacher's lectures?

- Completely ☐
- Partly ☐
- Not at all ☐

RUBRIC 4 : ENGLISH ASSIGNMENTS

1. What English assignments did you have to do?

- Report an experiment ☐
- Write an essay ☐
- Write a project ☐
- Read text plus answer comprehension questions ☐
- Do grammar exercises ☐
- Explain scientific terminology ☐

2. Were English assignments

- Very difficult ☐
- Very easy ☐

3. How often were you assigned work out of class?

Very often Not at all

5 4 3 2 1

RUBRIC 5 : THE DIFFERENT SKILLS (READING WRITING, LISTENING, SPEAKING) TAUGHT AT YOUR DEPARTMENT

A- The reading skill

1. How often do you read in English?

- Frequently ☐
- Sometimes ☐
- Never ☐

2. **Is reading English recommended in your field of study?**

Yes, why?

No, why?

3. **Do you have to read?**

Textbooks	☐
Articles in journals	☐
Handouts	☐
Other (specify)	☐

4. **Indicate the difficulties you meet when reading?**

Scientific terminology	☐
General English	☐
Grammatical structures (specify)	☐
Other (specify)	☐

5. **Which strategies do you use when reading English?**

B- <u>The writing skill</u>

1. **How often do you write in English?**

Frequently	☐
Sometimes	☐
Never	☐

2. **What do you have to write in English?**

An essay	☐
A project	☐
A summary of your thesis	☐
Other (please specify)	☐

3. **Indicate the difficulties you meet when writing?**

Word level	☐
Sentence level	☐
Paragraph level	☐

C- **The Speaking skill :**

1. **Is speaking English necessary in your field of study?**

   Yes ☐
   No ☐

2. **Do you need to speak English?**

   Yes, why?

   No, why ?

3. **How often do you speak English?**

   Frequently ☐
   Sometimes ☐
   Never ☐

4. **Indicate the difficulties you meet when speaking English.**

D- **The Listening skill :**

1. **Is listening to English necessary in your field of study?**

   Yes ☐
   No ☐

2. **How often do you listen to English?**

   Frequently ☐
   Sometimes ☐
   Never ☐

3. **When/where do you listen to English?**

RUBRIC 6 : COURSE CONTENT

1. **Please assess the standard of teaching in your English lectures:**

Very good Very poor

5 4 3 2 1

2. **In your opinion was English class content generally relevant to your subject speciality (biology)?**

Very relevant Not relevant at all

5 4 3 2 1

3. **Did your teacher help you?**

To understand scientific facts ☐
To learn the grammar rules of the English language ☐
To learn scientific terminology ☐
Other (specify) ☐

4. **Did your lecturers make use of the following teaching aids?**

	Very often				Not at all
Handouts	5	4	3	2	1
Overhead projector	5	4	3	2	1
Tape-recorder	5	4	3	2	1
The blackboard	5	4	3	2	1

5. **Do you feel teaching aids were constructively used?**

Yes, because_____

No, because_____

6. Learning activities : which of the following learning activities did you do in class, and how often ?

	Never	Sometimes	Often
Individual work	□	□	□
Pair work	□	□	□
Group work	□	□	□
Free composition	□	□	□
Written/oral exercises	□	□	□
Individual project	□	□	□
Group project	□	□	□

7. How, would you rate the availability of the resources you may have needed for your course?

	Very good				Very poor
Library books	5	4	3	2	1
Lab facilities	5	4	3	2	1

Other (please specify)

8. Did the English course offered to you help you acquire the necessary skills?

Yes □
No □

9. If not, the problems in learning English are due to the.

Lack of appropriate syllabus □
Lack of adequate material □
Lack of teacher's experience □
Other (specify) □

RUBRIC 7 : STUDENTS' ASSESSMENT :

1.What assessment methods did your teacher use?

Exams □
Assignments □
Continuous assessment □
Project □
Other □

2. **Please place these assessment methods in your order of preference :**

Most preferred				Least preferred
5	4	3	2	1

Exams	☐
Assignments	☐
Continuous assessment	☐
Project	☐
Other	☐

3. **Do you think that the assessment is relevant to the course as it was conducted?**

Yes, definitely			No, not at all
1	2	3	4

4. **If not, what was/were the main problem (s)?**

RUBRIC 8 : OVERALL COURSE EVALUATION

1.Please indicate your overall view of the teaching quality of your English course in each year of study :

	Very good				Very poor
3rd Year	5	4	3	2	1
4th Year	5	4	3	2	1

2. **Which aspect (s) of the English course do you regard as having been the best taught in each year of study?**

3rd Year	4th Year

3. **Which aspects (s) of the English course do you regard as having the least well taught in each year of study?**

3rd Year	4th Year

4. **Please list any aspect (s) of the English course you regard as having been irrelevant to your course?**

3rd Year	4th Year

5. **Please state below the main sources of difficulty you encountered in pursuing your studies.**

6. **What suggestions would you make for improving the English course?**

RUBRIC 9 : SOME MORE GENERAL QUESTIONS :

1. **Please list 2 strengths of the English course.**

2. **Please list 2 weaknesses of the English course.**

3. **What do you feel you have gained from the English course?**

Thank you very much for your assistance, it is much appreciated.

APPENDIX C

TITLES OF THE ARTICLES IN THE BIOLOGY CORPUS

APPENDIX C

THE BIOLOGY CORPUS

(The fifteen Journal Articles are presented from the shortest to the longest, in terms of words, not in pages, as referred to in 'Chapter Four / Section 4.8.4.1' / Tables 4.4, 4.5 and 4.6)

Text 1: Cloning and Expression of a rat D2 Dopamine receptor cDNA (1104 words).

Text 2: Treatment of Parkinson's Disease with the Abeorphine CI 201-678, A Rigid Analogue of Dopamine (1988), in Biogenic Amines, Volume 5, No.3/4, pp.269-274, (1961 words)

Text 3: Smectite in Acute Diarrhea in Children: A Double-blind Placebo-controlled Clinical Trial (1964 words)

Text 4: Inhibition of Apomorphine-Induced Yawning and Penile Erection by Neurotensin (2540 words)

Text 5: Alzheimer's Disease and Control Brain Contain Soluble Derivatives of the Amyloid Protein Precursor that End within the βamyloid Protein Region, in Biochemistry 1992, 31/ 44, 10936-10940 (2805 words)

Text 6: The Discriminant Dopamine Antagonist Property of Benzamides is Observed at Various Times after their Systemic or Intracerebroventricular Administration (2867 words)

Text 7: Hypercholesterolemic Responses in Rabbits to Selected Groups of Dietary Essential Amino Acids, 1994, (2909 words)

Text 8: Relation between Chloride Secretion and Intracellular Cyclic Adenosine Monophosphate in a Cloned Human Intestinal Cell Line HT-29 CI 19A (3712 words)

Text 9: Cloning and Expression of Three Rabbit Kidney cDNAs Encoding Lauric Acid w-hydroxylases, in Biochemistry, 1990 / 29, pp.873-879 (4017 words)

Text 10: A Packaged Rice-Based Oral Rehydration Solution for Acute Diarrhea (4050 words).

Text 11: Molecular Cloning and Primary Structure of Rat α1-Antitrypsin (4296 words)

Text 12: Involvement of the Myenteric Plexus in the Cholera Toxin-Induced Net Fluid Secretion in the Rat Small Intestine, in Gastroenterology, 1993, 105/5, pp.1286-1293 (4400 words)

Text 13: Unexpected Potentiation by Discriminant Benzamide Derivatives of Stereotyped Behaviours Elicited by Dopamine Agonists in Mice (4667 words)

Text 14: Molecular Cloning and Characterization of rklk10, a cDNA Encoding T-kininogenase from Rat Submandibular Gland and Kidney, in Biochemistry, 1992 /31, 10922-10928 (4893 words)

Text 15: Kinetic Analysis of the Coding Properties of O6-Methylguanine in DNA: The Crucial Role of the Conformation of the Phosphodiester Bond, in Biochemistry, 1994 /33, pp.5335-5346, (9815 words).

APPENDIX D

LIST " FREQUENT NON -GRAMMAR" OF

THE BIOLOGY CORPUS

APPENDIX E

A SELECTED ARTICLE AS A SAMPLE FOR USE

APPENDIX F

SYNOPTIC PROFILES AND CONCORDANCE LINES OF NODES 'PROTEIN' AND 'ACID'

1- Operation				
Operating principle	***Working substance in operation***	***Means of operation***	***Characteristic Working Part***	***The person who formulated the operating principle***
- amino acid analyser. - Immobilon –N filter. - Elution buffer . - Rapid-quench apparatus.	- Rice flour. - Glucose – electrolyte. - Milk formula. - Mycoplasma contamination - Drug regime. - Spiroperidol binding. - Blood samples. - Amino acids. - High-salt and low-salt buffer. - Agarose gels. - Magnesium ions. - Gas- light syringes. - Compressed-air supply. - Snake venum / Water bath. - Saline – injected mice. - Elution buffer.	- Ear vein; etc.	- Weighing scale. - Syringe plungers. - Gas-tight syringes. - Rapid-quench apparatus	- Wilcoxon rank. - Ussing chambers. - Transwell dishes. - Transwell filters. - Madopar treatment. - Madopar doses. - Ringer solution. - Denhardt's solution. - Klenow fragment. - Watson – Crick pairs. - Michaelis – Menten Kinetics. - Sprague – Dawley rats. - Alzheimer's disease. - Parkinson's disease.

Table 4.11a : Functional Analysis of Ncs (following Bartolic's Categorisation)

2 - Materials	*3- Purpose of application*	*4- Location*	*5- Professional Engagement*	*6- Shape / form*
- Wire bottons. - Metal bars. - Wire netting. (about cages where mice under treatment are kept)	- Treatment group. - Control group. - Dehydration solution. - Feasibility study. - Voltage clamp. - Filter paper. - Control mice.	- Binding sites. - Cleavage sites. - Cell- lines. - N-Terminal region. - Coding region. - Enzyme subsites. - CAMP- binding sites. - Colony hybridization	Manufacturer's Instructions (for instruments etc.)	- Inhibition curves. - Loop structure. - Bell-shaped curve.

Table 4.11b : Functional Analysis of Ncs (following Bartolic's Categorisation)

6.7 The General Teaching Framework (based on Willis's, 1996)

The general teaching framework of the reading course for biology students should be based on different phases. These phases are well-represented by Willis (1996:60) in the following diagram.

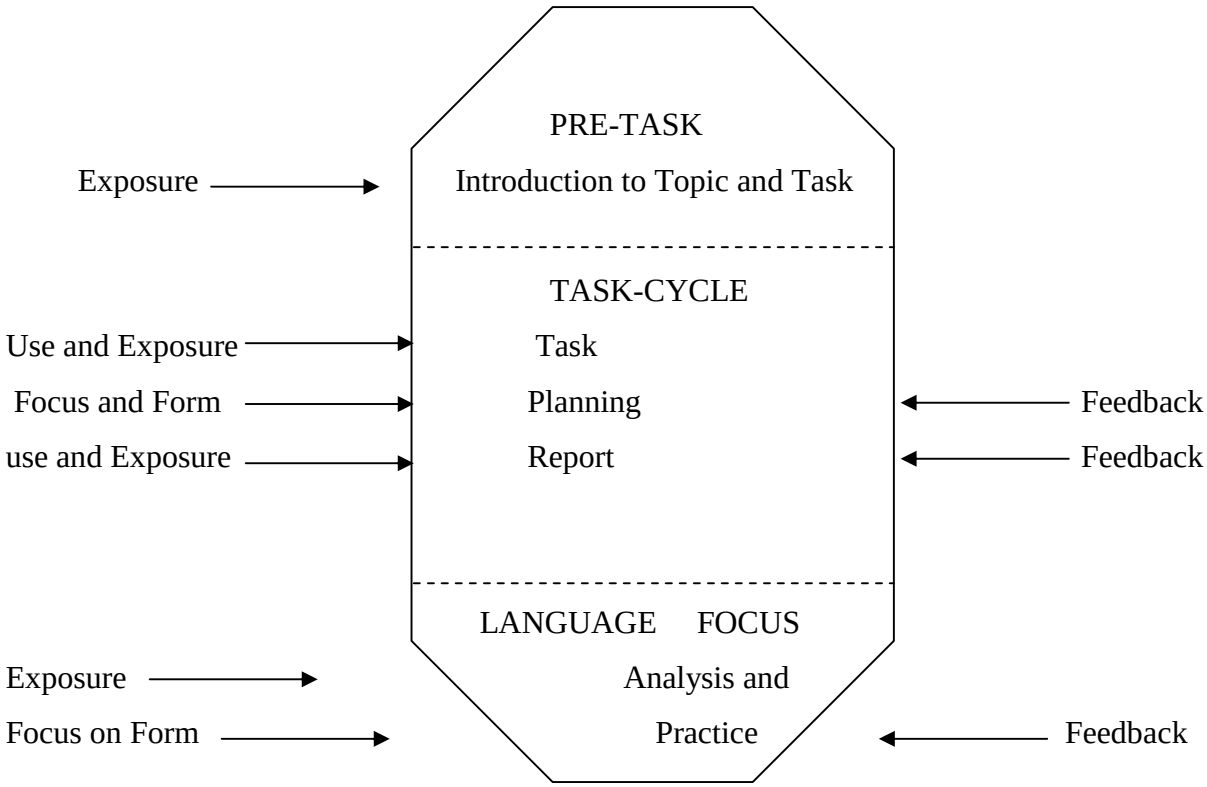


Diagram 6.1 : Components of the General Teaching Framework

(Source: Willis:1996:60)

Within this framework, we can notice that the four conditions for language learning are fully taken into consideration i.e 'Exposure to the Target Language', 'Use of the Target Language', 'Instruction' including 'Analysis and Practice', and finally 'Motivation' expressed through 'Feedback' by guiding, praising the students, etc.

The framework also follows a natural progression (from the holistic part to the specific) and could be explained as follows.