

**USING COMPUTER CONCORDANCING FOR THE
LEARNING OF FRENCH AS A SECOND
LANGUAGE**

JEREMY ARTHUR WILLIAM WHISTLE

Master of Philosophy

ASTON UNIVERSITY

October 2004

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Aston University

USING COMPUTER CONCORDANCING FOR THE LEARNING OF FRENCH AS A SECOND LANGUAGE

Jeremy Arthur William Whistle

Master of Philosophy

2004

Summary

The present thesis describes an experiment in the use of computer concordancers with undergraduates learning French as a second language in the United Kingdom. The main hypothesis to be tested by the project was that working with a concordancer would make students more autonomous as well as enrich their learning experience. Evaluation of this hypothesis would focus on student reaction rather than on actual results and would require qualitative data based upon the students' own assessment of the value of the concordancer and the level of autonomy achieved.

The original intention was that students should use the software themselves to create their own concordances which they would then analyse, individually or in groups, as part of a process of personal rule derivation, in order to make them more autonomous as language learners.

Technical problems and the reluctance of the students to adopt the learning methods proposed meant that the experiment could not be completed as originally conceived. The original student-centred approach was then replaced by a more tutor-centred approach leading to a much greater integration of computer generated concordances into course materials used in the teaching of French to first and second year undergraduates.

The description of the experiment is preceded by a survey of Computer Assisted Language Learning (CALL), language teaching methodology and the various uses of concordancers in language teaching in order to provide a context for an evaluation of the experiment.

In the conclusion, the experiment is evaluated both in terms of its own outcomes and in the framework of other research into the use of concordancers and the encouragement of autonomous learning.

Key words

CALL
Corpus
Second Language Acquisition
Autonomous Learning

Acknowledgements

I should like to thank Dr Pierre Larrivée for his support and above all the understanding that he demonstrated during the frequent periods of silence that occurred during the writing of this thesis when my professional responsibilities required my full attention. I should also like to thank my employer, University College Northampton and especially the Northampton Business School, for generously funding the early years of the project.

LIST OF CONTENTS

Summary	2
Acknowledgements	3
List of Contents	4
List of Tables and Figures	6
1 INTRODUCTION	7
1.1 The National Context	7
1.2 The Research Problem	7
1.3 The Institutional Context	10
1.4 The Research Hypothesis	11
1.5 Conclusion	13
2 LANGUAGE TEACHING: THE EDUCATIONAL CONTEXT	14
2.1 Introduction	14
2.2 The Four Methodologies	14
2.2.1 The Grammar Translation Methodology	14
2.2.2 The Behaviourist Methodology	16
2.2.3 The Cognitive Methodology	19
2.2.4 The Communicative Methodology	21
2.3 The Position of Grammar	23
2.4 Conclusion	27
3 THE HISTORY OF CALL	29
3.1 Introduction	29
3.2 Why Use CALL?	31
3.3 The Development of CALL	34
3.3.1 CALL and Computer Technology	34
3.3.2 CALL and Language Teaching	37
3.3.2.1 The First or Behaviourist Phase	39
3.3.2.2 The Second or Communicative Phase	42
3.3.2.3 The Third or Humanistic Phase	46
3.3.3 CALL and Higher Education in the United Kingdom	47
3.4 Conclusion	49
4 CONCORDANCERS IN LANGUAGE LEARNING AND TEACHING	52
4.1 Introduction	52
4.2 The Development of Computer Concordancing	53
4.2.1 Concordancing and the Computer	53
4.2.2 Concordancing and Modern Corpus Linguistics	56
4.2.3 Concordancing and Corpus Design and Construction	58
4.3 Concordancers and the Teaching Context	66
4.3.1 Concordancing and Lexicography	66
4.3.2 Course and Syllabus Design	67
4.3.3 Grammar / Course Book Evaluation	68
4.3.4 Testing	69

4.4	Teaching with Concordancers	71
4.4.1	Pedagogic Principles	72
4.4.2	Literature	75
4.4.3	Lexis	79
4.4.4	Grammar	82
4.4.5	Translation	86
4.5	Conclusion	91
5	THE RESEARCH PROJECT	93
5.1	The Research Question	93
5.1.1	Pedagogical Issues	93
5.2	The Experiment	94
5.2.1	The Corpus	94
5.2.2	The Concordancer	96
5.2.3	The Experimental Groups	102
5.2.4	Implementation of the Concordancer: Stage 1	102
5.2.5	The Students' Response	108
5.2.6	Practical problems that Arose	108
5.2.7	Preliminary Conclusions	110
5.2.8	Implementation of the Concordancer: Stage 2	112
5.3	Conclusions	119
6	CONCLUSION	121
6.1	Introduction	121
6.2	Operational Issues	122
6.3	Pedagogical Issues	123
6.4	Conclusion	129
6.5	Postscript	131
	References and Bibliography	133
Appendix 1	Questionnaire	154
Appendix 2	Worksheets	156
Appendix 3	Course materials	164
Appendix 4	Guide Sheets	178
Appendix 5	Notes on KWIC-CONCORD	185

LIST OF TABLES

Table 1	Adjectives collocating with <i>absolutely</i>	74
Table 2	student approaches	127

LIST OF FIGURES

Figure 1	“Below is a list of core vocabulary items...”	114
Figure 2	“Some very common words can only be understood...”	115
Figure 3	“What are the dominant grammatical roles...”	116
Figure 4	Concordance of <i>avant</i>	118

CHAPTER 1

1 INTRODUCTION

1.1 THE NATIONAL CONTEXT

The background to the experiment described in the present study was the widespread feeling among foreign language teachers in Higher Education in the UK that students arrive at university with a grammar deficit (Metcalf 1992; Corness *et al* 1997; Goodfellow and Metcalf 1997). By the mid 1980s many universities were already organising remedial teaching “specifically directed at those areas of the target language system that cause the learner trouble”(Ellis 1987:56). Many if not most foreign language degree courses now contain a substantial grammar revision/development component. The grammar sessions are however often perceived, even by students who are successful, as being boring and receive relatively low popularity ratings in the half-yearly feedback returns. This is in spite of the fact that students generally want grammar teaching and recognise its usefulness¹. The low popularity ratings, combined with the perception that teaching grammar seems to have little effect with certain students, highlighted the need for a re-examination of existing practice and a search for new methods. Ellis (1987) suggests a way forward:

Simply practising correct language forms cannot guarantee that learning will occur. Therefore, if remedial teaching is to continue it would do well to eschew mechanical practice and concentrate instead on *consciousness raising*, i.e. adopt an explicit approach (op.cit.:70).

One activity he explicitly excludes is language laboratory drilling (ibid). One can reasonably assume that this exclusion extends to any kind of mechanical drilling.

1.2 THE RESEARCH PROBLEM

The precise problem which inspired this research project was one which was causing increasing concern across languages. This was the apparent inability of too many students to develop beyond a certain point. Their language seems to remain fossilised at a stage which can only be described as “pre-grammatical”. It

¹ In a survey carried out by the author in October 1998, 74% of Y1 students and 95% of Y2 students thought grammar was important. 73% of Y1 and 88.5% of Y2 thought grammar teaching was helpful although only 33% of Y1 and 46% of Y2 thought it interesting. Slightly fewer, 20% of Y1 and 39% of Y2, thought it enjoyable.

is only necessary to observe the occasional mature student, 15 or 20 years after taking 'O' or 'A' Level, in a class of 18/19 year olds who sat their 'A' Level examinations a few months before entry to see that there has been a decline in language skills. However slow and hesitant the mature students may be at the beginning, by the end of the first year they have begun to outstrip the others in all skills, including the spoken. A similar situation is described by Klapper (1997:25-26). Many, if not most, of the rest of the students are handicapped by what can only be described as "grammar-blindness". Because in the first crucial years of language learning, essentially up to GCSE, they are not taught to see features of language such as gender, number, tense, subject, object, word order, they become blind to them. The responsibility for this lies largely with the National Curriculum for Modern Languages which has led to a "marginalisation of explicit grammar teaching (considered counterproductive) as well as little attention to knowledge about language" (Block 2002:20). The level of grammatical confusion in the minds of secondary learners of French in their GCSE year is analysed in considerable detail by Metcalfe *et al* (1998). As a result learners fail to acquire a core of grammatical structures enabling them to progress beyond the stage of mechanical production. This observation is supported by Klapper (1998): "My experience of examining GCSE candidates and teaching the products two years on, is that far too many of them do not perceive language structurally, and therefore have no chance whatsoever of ever forming language independently"(op.cit.:22). This inability to perceive language structurally also leads to poor and even non-existent reading skills.

The changes that have taken place in teaching methodology in schools (discussed in more detail in Chapter 2) mean that it is necessary to teach undergraduate students more than in the past in order to bring them to a level where they can read, let alone write, with reasonable accuracy. Yet, at the same time the teaching hours have over the years been progressively reduced. Whereas fifteen years ago, first year Combined Honours students at University College Northampton (then Nene College) had six hours of tuition in French, they now have four. It is therefore even more important that they can use their non-timetabled time to extend their knowledge of the language. One way to achieve this is to help them develop as autonomous learners.

Experience of working with such students had led to doubts about the efficacy of formal grammar teaching, the principal objections being:

- 1) it leads to simplifications when we have to teach our students to cope with complexity;
- 2) it tends to treat forms in isolation and not in opposition to other forms;
- 3) it involves decontextualisation which effectively excludes many of the discourse features which, alone, can often explain why a specific word or structure is used.

However, in spite of personal scepticism, it was decided to put in place a revision grammar syllabus going back to the absolute basics. This was in part because of pressure to come into line with colleagues in other languages and also because the institution requires detailed schemes of work to be given out at the beginning of the year. Unfortunately, there remained an irreducible number of students who made no visible progress and for whom, to take one example, use of the relative pronouns *qui* and *que* remained totally arbitrary. This could in part be imputed to the passive role that the grammar classes imposed on the students. The lesson format did not require them to remain passive. Opportunities were provided for questions but when these came they were always from the mature students. Not only did the non-mature students for the most part not ask questions, they often positively resented the answering of mature students' questions, complaining that class time was being wasted on unnecessary complications.

It was in the hope of resolving these problems that the decision was taken radically to change the way that grammar was taught to first and second year students. The priorities were:

- 1) to increase students' grammatical awareness;
- 2) to make them more autonomous as language learners so that more learning could take place outside formal contact hours.

Personal experience of using a concordancer as a research tool provided a model for introducing it as a language learning tool with first and second year students. The principal reason for the decision was the concordancer's ability to search large numbers of texts incredibly quickly for examples of real language usage:

Concordancing is a means of accessing a corpus of text to show how any given word or phrase in the text is used in the immediate contexts in which it appears. By grouping the uses of a particular word or phrase on the computer screen or in printed form, the concordancer shows the patterns in which the word or phrase is typically used (Flowerdew 1993b:87).

Taking these patterns as the starting point, the intention was to put students in a position where they had to ask questions about meaning or structures and then find answers. The aim was that, through having to search texts for forms and provide explanations, they would not only become more active participants in their own language learning but, crucially, they would begin to see the surface features of language and progressively lose their blindness. The general conceptual framework was that of discovery learning or, more specifically, “data-driven learning” (DDL) as described in (Johns 1991a:1):

What is distinctive about the DDL approach to inductive language teaching is the principle that the data is primary, and the teacher does not know in advance exactly what rules or patterns the learners will discover. Indeed [students] will often notice things that are unknown not only to the teacher, but also to the standard works of reference in the language. It is this element of challenge and discovery that gives DDL its special flavour and stimulus.

At the heart of data-driven learning is the concordancer which provides the raw data. Requiring students to use the concordancer themselves would involve them in both extracting and interpreting the linguistic data. The real test would be how far the rules formulated were an adequate explanation of the phenomena, even if they did not conform to the conventional explanations provided by the published grammars. Underlying and in part justifying the experiment was an awareness of a gap in the research more recently summed up by Kennedy and Miceli (2001:80): “Much has been written on *what* can be done with corpora in language learning [...] but relatively little in the literature on *how* students actually do this, and especially on how they fare on their own.”

1.3 THE INSTITUTIONAL CONTEXT

The specific context in which the research project described below took place is the teaching of French to undergraduate students at University College Northampton. Until the academic year 1999-2000, students were taught separately according to the degree for which they had enrolled. The two degree courses which were relevant when the project was originally conceived are the BA European Business and the BA Combined Honours. The former was a four year course, including a compulsory year abroad, in which students study a single foreign language in parallel with a broad-based core of business subjects. In 1998, the minimum entry requirement for French in the European Business degree, since

renamed BA International Business, was changed from a pass at 'A' Level to a pass at GCSE. The BA Combined Honours is a three year course, with the possibility of an optional year abroad, in which students can study French as a Major, Minor or Elective. In 1999, the minimum entry requirement was brought into line with that for European Business although students wishing to major are still expected to have a pass at 'A' Level. One reason for this change was the introduction in the same year of a new modular structure, the main consequence of this being that students are now taught together regardless of the degree on which they are enrolled. As a result the European Business students were no longer available as a potential separate control group and small overall numbers made it impossible to subdivide the language classes.

Until recently, there were no specialist language degrees. The highest level of specialisation that students could achieve was as a Major in French in Combined Honours where French accounted for 4/9 of the course. Progressively German and Spanish were added as Minor subjects amounting to 3/9. As the Combined Honours regulations required students to study three subjects, two languages could thus be combined to make 7/9 and three languages combined to make 9/9 of the total. The new modular structure allows students to study two subjects making it easier to specialise in languages. Until October 2002, French was the only language available as a Major requiring a minimum of 9 modules and a maximum of 11 modules out of 18. Since then German has been approved as both a Major and Joint subject.

1.4 THE RESEARCH HYPOTHESES

In order to address the situation described in Section 1.1 above and the specific problem described in Section 1.2, it was decided to introduce concordancing into the first and second year French courses as part of a general strategy of making students more aware of the processes involved in learning a foreign language and more willing to assume responsibility for their own progress as described by Little (1991:3-4, quoted in Littlemore 2001:41): "Essentially autonomy is a capacity – for detachment, critical reflection, decision-making, and independent action. It presupposes, but also entails that the learner will develop a particular kind of

psychological relation to the process and content of his learning.”. The main hypothesis to be tested by the project was that:

working with a concordancer would make students more autonomous.

Evaluation of this hypothesis would focus on student reaction rather than on actual results and would require qualitative data based upon the students’ own assessment of the value of the concordancer and the level of autonomy achieved.

A secondary hypothesis to be tested was that:

working with a concordancer would reduce the number of errors that students make and thus improve their end of year performance.

It was planned to subject this second hypothesis to quantitative evaluation, based upon comparing students’ performance with 1) that of previous cohorts and 2) that of a control group which had not used the concordancer. In the first year of the experiment there was a small parallel group. However, the size of this group made its use as a control problematic. From 1999, as a result of modularisation, there would no longer be a control group. Therefore, the only basis of comparison would be the scores of previous cohorts. The only anticipated difference between cohorts from the point of course content and delivery would be use of the concordancer.

The research paradigm within which the experiment was to be conducted can be described as “loosely ethnographic”(cf Johnson 1992:132-163) since its aim was to investigate (and if possible change) students’ language learning behaviour in the light of widespread assumptions about what is involved in the language learning process. In addition, there would be personal involvement in the experiment as both participant and observer. As with any experiment, obtaining and evaluating the data can prove to be problematic. This proved to be the case. In the event, as a consequence of the short test period and the reorganisation of teaching groups, it was decided not to proceed with the second hypothesis and, instead, to concentrate on the first which was the primary purpose of the experiment. Students were therefore asked to judge the perceived usefulness not only of the concordancer but of the other software² being used via a questionnaire (Appendix 1).

2 Essentially the TELL grammar and translation practice activities (GramEx and TransIt Tiger).

1.5 CONCLUSION

The experiment began at the beginning of the academic year 1998-1999. Two stages were planned: the first stage of the project was to take place with Combined Honours students all of whom had a pass at 'A' Level; the second stage was to be undertaken with a mixed level group, the majority but not all of whom would have a pass at 'A' Level. Because of the non-specialist nature of the language courses, there tend to be relatively few high 'A' Level grades in French, the majority being between C and E.

Before, however, beginning the description of the research project and its outcomes in detail, it is necessary to place the project in its wider pedagogic and technical context. Therefore the study will examine the wider context of language teaching in Chapter 2, the narrower context of Computer Aided Language Learning in Chapter 3 and the application of concordancers in Chapter 4. The experiment itself will be described in Chapter 5. The Conclusion will try to assess the extent to which the experiment succeeded in achieving its aims and, by extension, to evaluate the contribution that the concordancer can make to the general issue of improving learners' motivation and autonomy.

CHAPTER 2

2. LANGUAGE TEACHING : THE EDUCATIONAL CONTEXT

2.1 INTRODUCTION

Any discussion of the use of concordancers and, by extension, of any application of the computer to language learning must be set in the wider context of language learning and teaching within which tools such as computers are employed. The history of foreign language teaching can be traced back over 2000 years and it is therefore not surprising that there have been changes in methodology. However, it is not the purpose of this chapter to provide a comprehensive history of these methodological changes¹. The intention is to provide a framework within which to situate, in the next chapter, the development of computer aided language learning (CALL) by examining the three methodological orthodoxies which dominated language teaching in the English-speaking world in the second half of the 20th Century.

The development of state education systems in the last hundred years has led to a situation where a single methodology can dominate once it has received the *fiat* of the national decision makers. Bureaucracies are usually slow to react and do not of themselves welcome change or too much variation. Changes in methods can also require massive investment in new materials and equipment. In the last 60 years there have been four methodologies which have dominated language teaching in the West and these will be discussed below. Each of these has had to consider the value and status of grammar and its role in the language learning process. As one aim of the present project is to increase students' awareness of grammatical concepts and structures, it will also be necessary to address the question of the position of grammar in language teaching.

2.2 THE FOUR METHODOLOGIES

2.2.1 THE GRAMMAR-TRANSLATION METHODOLOGY

From the 1940s until the 1970s, in spite of occasional experiments with the so-called "direct" or "immersion" method (McArthur 1983:-99; Morgan and Neil

1 For a brief survey of the last 500 years see Hawkins 1994 and for a more detailed survey of the last 100 years see Stern 1983:97-116

2001:2-3; Stern 1983:456-460) in which only the foreign language was used in class, the customary way of teaching a foreign language was by the grammar-translation method (McArthur op.cit.:98-99; Morgan and Neil op.cit.:2; Stern op.cit:453-456).

The primary aim of the grammar-translation method was to equip students to read complex texts in the foreign language drawn almost exclusively from the literary canon. A secondary aim was to enable students to write the foreign language with a high level of accuracy, and once again the models were taken from literature. The two main activities were translation in both directions, more often than not of passages taken from novels. In French this required a good command of the past historic, the past anterior and the imperfect tenses, and even some familiarity with the imperfect subjunctive. The emphasis on literature and the translation and analysis of literary texts had a profound influence on teaching methods.

These aims are achieved in the classroom by long and elaborate grammatical explanations and demonstrations in the native language, followed by practice on the part of the student in the writing of paradigms, in the applying of the rules he has learned to the construction of sentences in the foreign language, and in the translation of consecutive passages of prose from the native to the foreign language (Rivers 1968:16).

Discussion, both oral and written, of the texts was largely carried out in the native tongue. Oral work in the target language was limited and largely mechanical and there was little or no attempt to relate the language being learned to real situations in the real world. A common platitude is that after five years of learning the foreign language, learners still could not buy a stamp or a cup of coffee in the country. If, however, one judges the method by what it set out to achieve, it must be recognised that, with certain students, it was highly successful. In the final years of secondary education and then at university they were expected to read and analyse the most demanding of texts, including those of the mediaeval period. However, as Rivers acknowledges, there were those for whom the method did not work.

The grammar-translation method is not successful [...] with the less intellectual, who muddle through, making many mistakes over and over again and thus building up cumulative habits of inaccuracy which are difficult to eradicate at a more advanced stage. Such students find foreign-

language study very tedious and they drop out of the class as soon as this is permitted (op.cit.:17).

In spite of such criticism and the challenge from alternative methodologies (see below), the grammar-translation method proved to be remarkably resilient throughout the world (Rivers 1968:28; Stern 1983:454; Brown 1987:16). In England and Wales throughout the 1960s translation dominated the General Certificate of Education (GCE) modern language examinations at both 'O' and 'A' Level. Even though, towards the end of the decade, some boards had begun to abandon translation into the foreign language at 'O' Level, they still retained translation into English (Incorporated Association of Assistant Masters 1967:149-169). The situation in England and Wales changed radically with the passing of the 1976 Education Act which accelerated the move away from selective towards comprehensive secondary education and which, for the first time, offered a modern language to all but a few pupils. The changes raised the question of the suitability of the grammar-translation method with the less intellectual. However, even when the method had largely disappeared from the schools, it survived in many university language departments (cf Bate and Hare 1986:14).

2.2.2 *THE BEHAVIOURIST METHODOLOGY*

The grammar-translation method was increasingly challenged from two directions. The earliest, historically, was by the supporters of the Audio-lingual Method. This had two intellectual foundations. The first was derived from psychology and was based on the behaviourist theories of B.F. Skinner and the second from linguistics and based on the structuralist theories of Bloomfield. Skinner's contribution was to take the earlier work carried out by Pavlov on animal behaviour and to develop it into a general theory of human behaviour, under the label of "behaviourism". The conditioned reflexes of Pavlov's dogs provided the basis for the rather more sophisticated concept of operant conditioning (cf Nunan 1991:229) in which the responses to stimuli acquired greater importance insofar as they were seen as the main reinforcers of human behaviour. Skinner extended this theory to include learning in general and language learning in particular (Skinner 1957).

One might expect the influence of theoretical linguistics to have been merely complementary but there was in fact a substantial degree of overlap between the two disciplines which helps explain the undoubted power and success of the audio-lingual method in practice. Just as Skinner projected his behaviourist paradigm on to language, Bloomfield (1933) had a great deal to say about the psychological aspects of language learning. Bloomfield and his colleagues had formulated their theories as a result of field work with American Indians which had forced them to develop new tools for linguistic analysis since many of the languages they studied did not share the familiar concepts and categories of European languages. Because, almost uniquely up until then, these linguists had to rely on non-educated informants for their insights into the grammar rather than on their own or other educated speakers' intuitions, they were struck by the fact that the native-Americans could use their language but could not describe it. Clearly, it was not necessary to memorise rules and grammatical paradigms in order to learn a language. From here it was only a small step to suggesting that teachers, instead of teaching about the foreign language, should simply teach the language. The combining of Skinner's operant conditioning with the experiences in the field of these linguists led to the elaboration of a new and, for the first time, 'scientific' methodology² which analysed each language in its own terms rather than through the use of traditional classical categories:

They attempted to isolate significant linguistic structures and patterns which were segmented into smaller and smaller constituents, which were considered to be the basic building blocks of the language under investigation. This technique for linguistic analysis had a persuasive influence on the selection of input for language teaching, and also on the development of classroom activities. The development of patterns and substitution drills represented a marriage of structural linguistic analysis and behaviourist habit formation (Nunan 1991: loc.cit.).

Although the theory was developed from the observation of real languages in action, all too often the methodology failed to bring the learner to the point of using the language in a real (or even simulated) situation. This was generally postponed, largely on the grounds that if the learner were to attempt to produce spontaneous language too soon, he or she might well make errors which would then become internalised, especially if a transaction were to be successful since

2 The method owed much of its success to its adoption by the American armed services during the Second World War (cf Stern op.cit.:102 and 463)

this would, in behaviourist terms, reinforce the error. As a result, audio-lingual language lessons rarely went beyond the stages of presentation and practice. Through use of the target language linked to gestures or props, in the form of pictures or actual *realia*, the teacher had to present new structures in as clear and unambiguous a way as possible. Because of the practical difficulties of bringing the world in the form of *realia* into the classroom, teachers became increasingly dependent on projectors and filmstrips, usually used in parallel with tape recorders. Gradually the term “audio-lingual” became replaced by the term “audio-visual”. In Britain the most popular published courses were *En avant* (primary) and *Longman’s Audio-Visual French* (secondary)³. In many classes, especially where the teacher was inexperienced, the lesson was essentially a mechanical operation mediated through the equipment, involving extensive repetition, mostly by pupils in unison⁴.

The audio-lingual or audio-visual method finally succumbed to an attack from two very different forces. One was the total discrediting of behaviourist linguistic theory and the partial discrediting of contemporary structural linguistic analysis by Noam Chomsky (1957 and 1959). The other was simply the failure of the method to achieve its goals. There occurred no great breakthrough in language performance as had been predicted. Instead, there was often widespread boredom, linked to the technology on which much, and sometimes most, of the teaching was based. The almost total exclusion of the written word even produced results which have been described as disastrous (Krashen and Terrell 1983:15). In the UK, the negative conclusions of the primary French experiment (Burstall *et al* 1974) relating to both levels of performance and learner attitudes marked the beginning of the end of audio-lingualism, at least in its “pure” form, in schools. The materials often lived on in the classrooms, if only because the funding was not available to replace them, but the routine of repetition and drilling was largely abandoned.

3 In France *De vive voix* and *Voix et images de France* were widely used for the teaching of French as a foreign language - for a description and critique of the latter see Blanchet (n.d).

4 This is based on the author’s own experience of observing primary and secondary school French classes between 1972 and 1983 (cf also Power 1977).

2.2.3 THE COGNITIVE METHODOLOGY

The title is slightly misleading insofar as there is no single cognitive methodology. What distinguishes Cognitivism from Behaviourism is that it sees mental processes as being distinct from physical processes and thus situates language learning within the wider field of psychology. It therefore provides a framework for a number of theories of second language acquisition which can be invoked to justify a variety of teaching methods which have in common that they are predicated upon the belief that language is a form of knowledge⁵. Within this framework it allows for differences of opinion as to whether language is a specialist form of mental activity requiring a specialist language acquisition mechanism or whether it is subsumed within a general psychological mechanism.

The best known exponent of the former position is Noam Chomsky. In fact it is not so much his general linguistic theory⁶ that has influenced language learning methodology (cf Roulet 1975:40-64) as his postulation that every child is equipped with a genetically inbuilt 'Language Acquisition Device' or 'LAD' (cf Ellis 1986:12 *passim*; Van Els *et al* 1984:134), which, in response to exposure to meaningful input and largely without the need for external intervention, constructs progressively and spontaneously the language system of the adult. In his earliest formulation of the concept Chomsky speaks of a "data-handling or 'hypothesis-formulating' ability of unknown character and complexity"(1959:57). The influence of Chomsky's concept of the LAD probably owes much of its success to the fact that it was taken up by Stephen Krashen and incorporated into his own theories of language acquisition (Krashen and Terrell *op cit*:39). While there is disagreement about the existence of a specific language acquisition device (Mitchell and Miles 1998:73), there is more general agreement that language acquisition is a largely sub-conscious psychological process of incredible complexity.

5 For a discussion of the different approaches see Mitchell and Miles (1998), chapter 4

6 Especially his concept of a "universal grammar" which has given rise to an approach often described as "linguistic" as opposed to "cognitivist" since it focuses on linguistic rather than psychological mechanisms

Cognitivism also allows for different views as to whether L2 learning follows the same pattern as L1 acquisition. On the one side there is the so-called nativist position typified by Chomsky and Krashen, which holds that the learner unconsciously derives a set of linguistic rules from the language he or she is exposed to. On the other side of the debate are a number of approaches which have in common the belief that language knowledge is of two kinds, 'declarative' or conscious knowledge of the rules, and 'performative' knowledge, which is the automatic ability to apply the rules without conscious reflection (Færch and Kasper 1987:12-13). While cognitive linguists agree that language learning requires the acquisition of automatic responses, they disagree about the role of conscious mental processing (cf Lightbown and Spada 1993:25-26) as opposed to learned reflexes. However, all would agree with O'Malley and Chamot (1990) that "language acquisition cannot be understood adequately without reference to the interaction between language and cognition" (op.cit.:42-54).

As far as language teaching is concerned, it is perhaps safer to speak of cognitive methods rather than methodology insofar as teachers and course writers did not publicly proclaim their allegiance to it as the new answer to the problems of language learning. In addition it cannot be considered as a new departure. It has in fact been described as "a modified, up-to-date grammar translation theory" (Carroll 1996, quoted in Stern 1983:469). What is important is that all but the most nativist of cognitivists recognise a clear role for grammar in the classroom. "The behaviouristic view of learning in terms of conditioning, shaping, reinforcement, habit-formation, and overlearning, has been replaced by an emphasis on rule learning, meaningful practice, and creativity" (Stern op.cit.:470). It is thus very much a continuation of previous assumptions and practice and provides a justification for many of the criticisms of audio-linguistic methods. Often, however, the term cognitive has been applied to learner styles (Ellis 1986:114-116) and strategies (Wenden 1991:19-24) rather than to a precise teaching methodology. Nevertheless, there is no doubt that cognitive theory and its application to language learning was instrumental in preparing the ground for the next major paradigm shift in methodology.

2.2.4 THE COMMUNICATIVE METHODOLOGY

The perceived failure of both the grammar-translation and the audio-lingual methods prepared the way for the next major development in language teaching methodology, the communicative method⁷, which emerged within a cognitivist framework. Once again, the inspiration came from outside the classroom. The first impulse came, as has been mentioned above, from Chomsky whose theories⁸ about both the structure of language and the way that language is learnt had a profound influence on language learning theory (cf Rivers 1983:6-14), although Chomsky himself has questioned the relevance of his theory to language teaching (cf Rivers 1983:30). Various terms have been applied to the classroom methods that derived from Chomsky's contribution but what is common to most of them is the attempt to recreate in the classroom the natural process without overt emphasis on grammatical structures. Especially revealing is the use of the term 'acquisition', popularised by Krashen (1981), in place of 'learning' in order to reinforce the view that first and second languages share the same learning processes.

The second challenge to traditional grammar teaching, and the one that gave its name to the methodology, came from the field of "communicative" linguistics (Celce-Murcia 1991:476). Dell Hymes' *On Communicative Competence*, published in 1971, questioned the quasi-absolute assumption that many linguists, following Chomsky, made about the nature of grammar (often referred to simply as syntax). This posited language as being a set of rules which were essentially context-free and shared by all native speakers, the definition of which should be based on an adult native speaker's intuitions and not on observation of language in use. Competence, in Chomsky's view, is the ability to discriminate between the grammatical and the ungrammatical. For Hymes, however, "[t]here are several sectors of communicative competence, of which the grammatical is one. Put otherwise, there is behavior, and, underlying it, there are several systems of rules reflected in the judgements and abilities of those whose messages the behavior manifests" (Hymes 1972:281). Communicative competence involves

7 A good overview of the theory and practice of communicative language teaching can be found in Richards and Rogers (1986) Chapter 5.

8 For a succinct account of the development of Chomsky's linguistic theories in the crucial years, from phrase structure via transformational to generative grammar, see Lyons (1977)

both 'knowledge' and 'ability for use'(op.cit.:282) as well as 'sensitivity to situations'(op.cit.:292). A child acquires not just a set of grammatical rules but also a set of behavioural rules which, Hymes implies, are from the social point of view often more important (op.cit.:277-278). Though Hymes' notion of communicative competence led many in the world of language teaching to challenge Chomsky's concept of grammar, it did not lead to the same questioning of the notion of a child's innate ability to acquire language largely unconsciously.

The third challenge, at least in Europe, was the project to develop functional-notional syllabuses that developed in the 1970s supported by the Council of Europe. Originally intended to define a minimum level of proficiency for adults within an overtly communicative framework (the so-called 'Threshold Level' : cf. Van Ek 1975; Coste *et al* 1976; Baldegger *et al* 1980), the project extended its work to cover language learning in schools (Van Ek 1977; Porcher *et al* 1980). Its debt to Hymes is clear: "The organisation of the content of a language learning system for adults must result from the analysis of the acts of language communication in which they are to participate. Any description of language contents must take into consideration language systems, language skills and language situations"(Trim *et al* 1973:5). The approach can be summed up as "what the learner can *do with* the language rather than what he *knows about* it" (op.cit.:7). Traditional syllabuses are explicitly criticised:

A logico-grammatical progression underlies most courses. Simple propositions are almost always the basis [...] with a successive introduction of noun phrase elements [...], categories [...] and functions [...] with the consequent cases and associated morphology. Syntactic extension covers the traditional government and concord, order of elements, derivation of questions and imperatives, and then the various forms of sentence embedding (op.cit.:12).

Because the grammatical order is primary, the situational element is relegated to secondary status:

Situations are selected at random or because they are suitable carriers for a new point of grammar, notions and functions are ordered by the extent to which the grammatical point in focus lends itself to their expression, vocabulary is selected haphazardly or because of its morphology or syntax (ibid).

While the authors recognise the necessity of grammar in the new order of things (op.cit.:19), this is clearly subordinate:

The classification of language learning objectives is based upon an analysis of the language-using operations required by a member of a speech community. Each operation can be described in terms of (a) the behavioural input-output chain involved, (b) the communication function performed, (c) the notional/semantic content expressed, (d) the formal linguistic resources employed, (e) the situation in which it occurs (op.cit.:18).

The theoretical base is clearly reflected in the English version of the definition of the Threshold Level (Van Ek 1975) and the order of description: specification of situations, language activities, language functions, topics: behavioural specifications, general notions, specific notions, language forms, degree of skill. The grammar descriptions contained in the English and German specifications are little more than inventories (this is in fact the term explicitly used) and are lexical lists of grammatical items rather than a structured syllabus. In the English Version (Van Ek 1975) grammar even seems to be an afterthought, included as an appendix. This is less true of the French version (Coste et al 1976) in which the grammatical syllabus is included as part of the core and has 79 pages allocated to it (as opposed to 48 in the English version).

One of the most influential members of the Council of Europe team was Wilkins (1973) whose contribution to Trim *et al* (1973) is one of the earliest expositions of a functional-notional syllabus. The new style syllabus, linked to the notion of needs analysis, also a by-product of the Council of Europe project (cf Richterich and Chancerel 1980), had a positive effect on language teaching to adults in the UK, in FE colleges, polytechnics and in-house courses for industry. The effect in schools was less positive, often leading to a constant repetition of learnt “situational” dialogues, involving little or no acquisition of lexis or grammar or even of useful communicative skills (cf Metcalfe *et al* 1998).

2.3 THE POSITION OF GRAMMAR

Traditionally, grammar was taught through explanation and practice. Rules were first presented, usually through use of a specialist grammatical metalanguage, by the teacher who then required his/her pupils to first demonstrate an understanding of the forms and concepts through successful completion of a

variety of exercises whose sole focus was the practice of the new rule or form. Exceptionally, there might follow a more open-ended and creative activity which required demonstration that the rule had been internalised. However effective this methodology was with some, possibly a minority, of learners, it was generally accepted that, for many, it was markedly unsuccessful. Usually, rules were presented and practised without reference to any semantic or pragmatic context. There was only one right answer and any discussion of alternatives was impossible. It is precisely this kind of explicit form-focused teaching devoid of any communicative context or purpose that the research surveyed by Johnstone (1997) and Spada (1997) often found to be ineffectual and demotivating.

It was recognition of this failure that, in part, led to the widespread adoption of so-called “communicative” methods which all but banished explicit grammatical teaching (cf Wright 1999), so much so that no less a person than the head of the Centre for Information on Language Teaching and Research (CILT), at a conference called to discuss the question of grammar teaching, could declare: “Had we not learned as young teachers in the 70s and early 80s that all such meaningless formality was a failed orthodoxy to be consigned to the dustbin of educational theory?” (King 1994:1). This was in spite of a general recognition of the central role of grammar by some of the most authoritative figures in applied linguistics (Strevens 1965; Wilkins 1972 and 1974; Corder 1973; Hornsey 1975; Roulet 1975; Widdowson 1978, 1979 and 1984; Stern 1983; Nunan 1991). These writers were indeed critical of a style of language teaching that was solely form-focused, just as they were critical of the unthinking application of linguistic theory to language learning, but, while they called for “communicative” or “discourse-based” methods, they re-affirmed the centrality of grammar as the means by which communication and discourse is realised. What they proposed was not the rejection of grammar but the elaboration of a new “pedagogic” grammar.

Nevertheless, the call for communicative courses (Munby 1978; Brumfit and Johnson 1979; Johnson and Morrow 1981; Johnson 1982), the introduction of notional-functional syllabuses (Wilkins 1976; Finocchiaro and Brumfit 1983) encouraged by the Council of Europe (Van Ek 1975 and 1977; Coste *et al* 1976; Porcher *et al* 1980; Baldeger *et al* 1980) and the introduction in the UK of

graded objective tests (Harding *et al* 1980) seem collectively to have evicted the teaching (and even any mention) of grammar from the curriculum. Reviewing the situation, Page and Hewett (1987) see a clear cause and effect relationship:

In the style and content of foreign language teaching against which the graded objectives system was a reaction, grammatical accuracy, and grammatical accuracy in the written word particularly, was the major objective. It is therefore not surprising if teaching language for communication was seen, at least in the first instance, as a liberation from the necessity of teaching grammar. At last teachers felt they could with a good conscience teach their pupils to buy a cup of coffee without having to convey at the same time that one of the nouns was masculine and the other feminine and the articles had to be corrected. Consequently, in the early schemes no mention was made of grammar at all (op cit:76).

The lack of grammar is obvious in the syllabuses, most of which were drawn up by local working parties for the most part made up of teachers, language advisers and teacher trainers (cf Gordon and Downes 1978) without any central direction at all. As a consequence, the York syllabus prescribed only the tasks to be completed, leaving the linguistic content entirely to the teacher, while the Oxford scheme prescribed everything the learner was expected to know except grammatical content which the teacher was expected to derive from the syllabus and its pedagogical implications (Page and Hewett:op.cit:loc.cit.). Teachers may have been able, if they thought it necessary, to extract grammatical information from the “syllabus implications” but to expect the learners, especially in the early years, to do so requires an extraordinary belief in the “input hypothesis”, especially given the restricted transactional types of activities that made up the content of many syllabuses. The authors note the appearance of grammatical content in later syllabuses and see in this the influence of the Council of Europe’s threshold descriptions (cf *supra*). What is missing from the threshold specifications, and the graded objectives syllabuses, is any notion of grammar as system. Far from “being put in its place - that is serving meaning and not as having an independent existence of its own”(ibid), it was quite simply banished. It is not surprising that teachers were quite often unsure and confused (Dunning 1983; Mitchell 1988). The anti-grammar “phrasebook approach to the early stages of language learning”(Mitchell 2003:18) was enshrined in the National

Curriculum for Modern Foreign Languages (DES 1991) introduced in the early 1990s and still in force with some modifications today⁹.

Another important factor was the generalisation of comprehensive schools in the UK, which required languages to be taught for the first time to the whole ability range, and the lack of any linguistics component in most foreign language and teacher-training courses. In the 1970s and 1980s, it was common to hear ML teachers in English schools claim that the only way of teaching languages to low-ability pupils was through rote-learning without any reference to rules or structure at all¹⁰. Error-correction was also often expressly forbidden. As most classes were mixed-ability, this effectively meant a total lack of grammatical instruction for all, with the result that “there is a significant proportion of 14-year-olds who have failed to internalise enough target language material to develop and control any stable language system, and who consequently cannot themselves engage in meaningful target use”(Mitchell 2003:19). Older pupils were found to be critical of a curriculum content that is “too narrow and too centred on mundane everyday life”(op.cit.:20).

While the rejection of explicit grammar teaching in the early years may be justifiable, in part at least, on the grounds that little knowledge of the rule system is necessary for simple communicative task fulfilment¹¹, this is to ignore the fact that the early years do not constitute for all learners the sum total of their foreign language experience. The underlying assumption often seems to be that all individuals will ever need to do in the foreign language is write a letter to a pen friend or engage in face-to face communication (cf Page 1994). There often seems little awareness of the fact that some (too few no doubt) pupils will want to continue beyond GCSE to ‘A’ Level and beyond. As a result, reading is often a neglected skill and yet it is the skill which provides the foundation for later study. There has been in fact little attention paid to the link between grammar and reading both from the point of view of developing grammatical competence and extracting meaning from written texts (cf Heafford 1993).

9 For a detailed description of the effects of the National Curriculum on modern language teaching in schools see Morgan and Neil 2001: 22-42

10 Based on the author’s experience in Initial Teacher Training between 1972 and 1984.

11 It is possible to argue that little in the way of lexis is necessary - pointing with a finger may be all that is needed!

One consequence of this shift away from the written word in schools has been the perception, certainly in Higher Education, that levels of grammatical knowledge and accuracy were declining, not just in French but also German (Perrin 1991; Metcalfe 1992; Sheppard 1993; King and Boaks 1994; Whitton 1994; McCulloch 1995; Wright and Craig 1995; Cornell 1996; Engel and Myles 1996b; Nott 1996; Rogers 1996; Goodfellow and Metcalfe 1997; Klapper 1997; McEnery *et al* 1997). A detailed analysis of both the research evidence and of school examination board reports (Metcalfe *et al* 1995) confirmed the general perception that there had in fact been a decline in grammatical knowledge over the previous twenty years with the effects being most noticeable in written work. For most commentators it is the effect on writing that causes the greatest concern. There is little if any attention, perhaps because of the difficulty of quantifying the effects, paid to the consequences of this decline for reading. And yet reading is perhaps more important since, certainly at advanced levels, the written word is the main medium through which learners will have contact with the foreign language. Certainly in the UK students will have little exposure to French language music, films or television. Developments in computing and especially use of the Internet have only reinforced the importance of the written word. Among the few to recognise this development is Johnstone (1994):

This puts increased responsibility on learners themselves to find their own ways of processing the input to which they are exposed. To this end it is essential that they should acquire an understanding of the system of words and rules that underlies the language they are learning. Otherwise, they will at best be able to extract the gist at a fairly superficial level. Significantly, this points to the importance of learning grammar for comprehension, whereas in the past the heavy emphasis has been on grammar for production (op cit:10).

It is precisely because they lack a sound grasp of syntax and grammar that many of our students are incapable of extracting anything but the most minimal information from the texts they are asked to read. Questions about the author's purpose or position will, in the circumstances, be largely pointless.

2.4 CONCLUSION

Changing fashions in methodology have had a profound effect on both the style and the content of language teaching. This is on the whole more true of schools than higher education. Although the concepts of communicative competence and

a needs-based functional-notional syllabus, mediated via the CNAA modern languages panel, were extremely influential in British polytechnics in the 1970s and 80s, the teaching of grammar was never abandoned but simply reformulated. As a consequence of the growing concern about the situation in schools and performance at GCSE and 'A' Level, there has been a general if not universal recognition (Saunders et al 1994, Saunders 1996) that there needs to be movement towards more explicit teaching of the rules of grammar. In the words of Celce-Murcia (1991): "During the last 25 years we have seen grammar move from a position of central importance in language teaching, to pariah status, and back to a position of renewed importance"(op.cit.:476). How it should be taught or learned is still a matter of debate. To use one metaphor, should grammar be taught as "statute law: pre-determined rules to be learned and applied"(McEnery *et al* 1997:10) or as "case law: looking for majority usage and using it to define the standard language [and engaging] students in interaction with real texts in a wider context"(ibid). The stance that teachers take in this debate will be crucial to the way that they approach using computers in the teaching process. The computer itself is totally compliant. The next chapter will look at the ways that the conception and application of CALL have been determined as much by methodological positions as by the technological potential.

CHAPTER 3

3. THE HISTORY OF CALL

3.1. INTRODUCTION

Is CALL anything more than an opportunistic response to a new technology? Can it enhance the process of language learning in a way that other technologies cannot? Above all, can its use be underpinned by a coherent theory of learning? Before considering these issues, it is important to situate CALL within a continuing history of new technologies being applied to foreign language learning. In the 1950s and early 1960s it was the radio, via the BBC language programmes, that brought into the classroom the nearest that one could get at that time to “real” French spoken by multiple voices. Radio was followed by television with both the BBC and independent companies transmitting language programmes to schools. However, radio and television never had the same effect as the portable tape recorder which, from the mid 1960s onwards, enabled the teacher to bring the sound of the foreign language into the classroom at any time in a much more authentic and varied way as well as a much more controlled way, at least in principle. Without the tape (and later cassette) recorder, the audio-lingual method (see section 2.2.2 of the previous chapter) might never have come to dominate language teaching methodology in the way that it did during the 1960s and 1970s¹.

The first applications of the computer to language teaching can be traced back to the 1960s (cf Ahmad *et al* 1985; Davies and Higgins 1985). However these were almost exclusively mainframe-based which considerably restricted access. The so-called micro-computer did not appear in ordinary classrooms² in the UK until the final years of the audio-lingual era in the early 1980s, triggering a new teaching revolution (cf Ahmad *et al* 1985:35-36; Kenning and Kenning 1983:142-144).

As usual with any new technology, the first applications tend to emulate familiar practices until the full potential becomes apparent. Gradually a new area of

1 The effect of the tape recorder on primary school French is documented in Burstall *et al* 1974.
2 It was only with the appearance of the BBC Micro microcomputer in 1981 that computers became a familiar feature in many British school classrooms.

language teaching / learning methodology began to evolve under a variety of names until the label Computer Assisted Language Learning (CALL)³ gained almost universal acceptance. In the “narrow” sense of the term, its meaning is restricted to applications “designed specifically to promote language learning” (Davies 1999:2) and this is how the term and its variants were generally used in the early days (Kenning and Kenning 1983; Higgins and Johns 1984; Last 1984), understandably as there were few if any other software applications in general use outside large offices. However most recent authors (Warschauer and Healey 1998; Bax 2003; Beatty 2003) follow Levy (1997:1) and use the term in its “broad” sense to include the use of any computer application in the process of language learning, whatever its original purpose, including word-processors, web browsers, communications programs, as well as those programs which have been specifically designed for the language classroom: “Computer Assisted Language Learning (CALL) may be defined as ‘the search for and study of applications of the computer in language teaching and learning’.” It is the broad definition that will be followed here.

In addition to disagreements regarding the scope of the term CALL, there exists a variety of approaches to classification. In the 1980s, when the field was still too young for observers to identify trends or patterns, there was a tendency to categorise CALL according to the type of activity (Davies and Higgins 1982; Higgins and Johns 1984; Last 1984). An alternative slightly later approach was to describe the evolving technologies and the potential they offer (Ahmad *et al* 1985; Davies 2001). Delcloque (2000) favours a different, purely objective and historical approach complete with dates and sources while Beatty (2003) focuses not on the programs but on their application to teaching and research. There have been relatively few attempts to relate CALL to pedagogical theory. Kemmis *et al* (1977) were probably the first in the UK to attempt to do so, classifying activities not by program type (e.g. gap-filling exercise) but by fitting them into a scheme of “paradigms” of learning (see Section 3.2 below). Kenning and

3. Ager (1986) identifies, in addition to CALL, the following acronyms:
 CALI - Computer-assisted language instruction
 CALT - Computer-assisted language teaching
 CML - Computer-managed learning
 CBL - Computer-based learning
 but does not mention CAI (Computer-assisted instruction).

Kenning (1990) discuss in some detail the use of CALL within a framework of both teaching methodology and language acquisition theory but avoid a more general classification. Cook (1985) is the first to link the development of CALL to the dominant teaching methodologies of the period, proposing four distinct categories, followed by Warschauer and Healey (1998) who sought to conceptualise two decades of CALL under three broad categories derived like Cook's from the prevailing teaching methodologies (see section 3.3.2 below for more detailed discussion).

The goal of this chapter is both to examine the reasons for using CALL and trace the history of its application to the process of language learning. The evolution of CALL will be considered first in relation to developments in computer technology and secondly in relation to changes in language teaching methodology. Section 3.3 will examine how, in the early days from the 1950s through to the 1970s, the main constraining factor was the computer itself and how, as it freed itself from its physical straitjacket, CALL found itself to a large extent constrained by the prevailing methodological orthodoxy. Before however addressing these historical issues, it will be necessary in Section 3.2 to consider the pedagogic advantages of CALL.

3.2 WHY USE CALL?

Although the first use of the computer to analyse foreign language texts was for the decoding of German war-time encoded messages, this was achieved through purely mathematical techniques (Knowles F 1990:39-40). Because the computer is essentially a "number cruncher" its earliest applications were in the physical sciences and business. This was true of the projects that were the basis for one of the earliest evaluations of Computer Assisted Learning (CAL) produced in 1977 by Stephen Kemmis and colleagues at the University of East Anglia (Kemmis *et al* 1977). Of the ten projects listed, two were in the field of mathematics, one was in the field of statistics, one in the field of engineering and three were in the field of the natural sciences. Of the non-quantitative applications, two involve decision making in medicine and business. What makes the computer so useful as a learning tool, in these cases, is its ability to store large amounts of information and, through branching procedures and algorithms, to establish

rapidly and exhaustively links between the categories of information. The only application that could be categorised as belonging to the Humanities is a local history project based on census returns supplemented by other documentation. In this case what matters is the ability of the computer to store and interrogate a large database. None of the applications listed and studied relate to language learning. The earliest applications in the languages field were essentially for research and exploited the computer's ability to store and search large amounts of information, in this case in textual form, for specified patterns (cf Burton 1981; Ahmad *et al* 1985:39-40).

What made the Kemmis study so influential, including in the field of CALL (cf Higgins and Johns 1984:18-19; Phillips 1985:30; Ager 1986:194; Backer 1995:1), was less the projects it evaluated than the model of learning that it derived from observation and analysis of the interaction between programs and users. If the model could be demonstrated to have a general validity, computer-assisted learning could be extended to disciplines other than those involved in the study, including language learning. The study identifies three major paradigms of education: the *instructional*, the *revelatory* and the *conjectural* (op.cit:24), each of which can be broken down for evaluation into a limited number of components. Because of the influence that the Kemmis study has had, it will be useful to examine each paradigm in some detail.

The *instructional* is equated with the classic drill-and-practice programs of computer assisted instruction (CAI) and with adaptive-tutorial projects. These involve the "belief that the knowledge students need to acquire can be specified in language and learned by the transmission of verbal messages"(ibid). In this paradigm the student is required to master a defined body of content through structured presentation of material and feedback. The learning theory underlying it is essentially behaviouristic.

The *revelatory* is equated with simulations and certain types of data-handling programs. It identifies gaps in the student's knowledge and sets out to help the student fill those gaps by a process of gradual revelation. The focus is not on transmitting chunks of information but on helping the student acquire key

concepts and ideas, often via simulations. The underlying theory is of learning by discovery.

The *conjectural* equates more closely with Artificial Intelligence and other intelligence modelling applications. For this reason, it lends itself to business style decision making processes which are based more on reflection on experience and proposing courses of action than on establishing statistical or other quantifiable links between items of information. This paradigm can be situated within a problem-oriented theory of knowledge as well as general cognitive theory.

A possible fourth paradigm is identified though not elaborated in as much detail. This *emancipatory* paradigm sees the computer as a labour-saving device which 'emancipates' the student from 'inauthentic labour'. Here the value of the computer lies in its ability to free the student from many of the chores involved in study such as hunting down the information or plotting graphs (op.cit.:27-30). It could be argued that this itself could be sufficient reason for using CALL in the absence of any other pedagogic justification.

The four paradigms, thus, provide a framework within which applications can be situated and evaluated. They also provide a learning rationale for their design and use. No one paradigm is necessarily better than the others although all are subject to the value judgements of changing educational fashion. Clearly a behaviourist approach would privilege the *instructional* whereas a discovery approach would tend towards the *revelatory*, as would concordancing which would, in addition, give an important place to the *emancipatory*. Whichever paradigm is adopted, the authors warn that there is no necessary link between the educational value of the paradigm and the quality of the software. They insist that what they have identified is only the potential of CAL and go on to counsel caution. Instructional software can impose the developer's questions and logic and control the process to such an extent that it demotivates the student. Revelatory software may reveal answers but not the problematic, and possibly contested, nature of those answers. Conjectural software may not help students formulate complex ideas and may end up imposing categories of thought. Emancipatory software may take the chore out of the activity but also the

learning value of performing the chore. The ideas contained in the programs or their functioning may be too complicated for students to master and the role and position of the software in the overall course structure may be ill thought out or simply inadequately explained. These caveats are highly relevant to the immediate research context.

The emphasis in the Kemmis study is very much on the pedagogical model and on the difficulties of implementing this model caused by the inability of developers to match the software to the learner's needs. This explains the influence of the study in the evolution of CALL and why, nearly twenty years after its publication, it should be adopted as a framework for a retrospective survey of the theoretical literature of the teaching of grammar using CALL (Backer 1995). The study both provides a powerful pedagogical case for the use of CALL and warns of some of the dangers. The paradigms are universal and hence timeless but, while they constitute a valuable grid for understanding and classifying CALL activities, they do not necessarily provide the best framework for understanding and evaluating the development of CALL in its historical dimension. Section 3.3 will therefore examine alternative paradigms in an attempt to assess whether they provide a better explanatory framework.

3.3 THE DEVELOPMENT OF CALL

3.3.1 CALL AND COMPUTER TECHNOLOGY

The earliest applications of the computer to language learning, as opposed to machine translation or literary exegesis, predate the microcomputer. In the 1960s the Universities of Stanford and New York at Stony Brook in the USA developed programs for Russian and German. In Britain the University of Essex was involved in a large and well-financed project to develop reading skills in foreign languages and especially Russian. The emphasis on Russian reflects the political priorities during the Cold War. One of the best-known and ambitious of these early programs was the PLATO⁴ project at the University of Illinois. By the end of the 1970s, CALL projects were in progress in the UK at the University of Hull, the University of Aberdeen, the University of East Anglia,

4. For an idea of the scale and scope of PLATO as well as other early projects, see Ahmad et al (1985:30-32); Chapelle and Jamieson (1986:30-32); Delcloque (2000); Beatty (2003:18-19)

the University of Surrey and Ealing College of Higher Education (Davies and Higgins 1985:8-9).

The reliance on mainframe computers seriously limited access. The cost of the hardware and the process of software development restricted use to a few well-endowed institutions or departments. There were also practical difficulties. Keyboards often had a limited range of characters (Burton 1981:139; Last 1984:12) based on the standard American typewriter keyboard, thus excluding all diacritics (Kenning and Kenning 1983:147). The introduction of the microcomputer in the late 1970s, and in the UK of the BBC Micro in 1981, radically altered the situation and for the first time made CALL a viable curriculum activity for most language teachers (Davies and Higgins 1985:8-9), helped in the UK by official support through initiatives such as the Microelectronics Education Programme (MEP).

This potential empowerment of teachers by the arrival of the microcomputer was accompanied by a wave of publications devoted to CALL including books (Davies and Higgins 1982,1985; Ahmad et al 1985; Higgins and Johns 1984; Kenning and Kenning 1983; Last 1984), journals (*CALICO* was probably the first) and newsletters (one of the earliest in the UK was *Callboard* which first appeared in 1982). Even though it represented an advance, the microcomputer imposed, certainly by modern standards, severe hardware constraints. The normal amount of memory one could expect was 16k, recommended as a minimum by both Davies and Higgins (1985:87) and Last (1984:29), although the BBC micro and the Commodore, two of the most popular machines available in the early 1980s, could later be expanded to 64k. Floppy disks gradually replaced tapes and cassettes and were available in 40 or 80 track formats giving up to 400k of storage. A screen width of 80 (as opposed to 40) characters or columns was becoming the norm. As an illustration of just how tight the physical constraints were, the first experimental version of the *Micro-Concord* concordancer ran within a shell of 1k of memory (Johns 1986:151).

The potential of the micro-computer was initially limited by the lack of available software. It was widely assumed that, as a result, teachers would need to become engaged in programming and many of the early authors in the field of CALL

gave examples of programs or routines written, for the most part, in BASIC (Davies and Higgins 1982; Kenning and Kenning 1983; Higgins and Johns 1984; Davies 1985; Higgins 1985). However, most teachers were soon spared the need to learn to program themselves thanks to the vitality of the market. In spite of the hardware limitations, by 1985 Davies and Higgins (1985:109-159) were able to list a number of commercial language programs for the microcomputer, most of them for EFL but some for foreign languages.

Because of the variety of machines in use (institutions in the UK were variously equipped with IBM, Apple II, Commodore, BBC Micro and RML 380Z/480Z PCs) and the lack of a standard operating system, the choice of software was in practice more limited. Higgins and Johns (1984) give examples of a wide variety of activity types, including drills, games (BINGO, JACKPOT) and quizzes (NOUGHTS AND CROSSES, SNAKES AND LADDERS), text manipulation activities (JUMBLER, TEXTBAG, STORYBOARD - some of which are still available in commercial form), simulations (ADVENTURE, PHOTOKIT, TOWN PLAN) and even simple artificial intelligence programs (SHRDLU, JOHN AND MARY). As an indication of the progress (or the relative lack thereof) over a period of 16 years, the C&IT Centre software database⁵ contained in January 2000 details of over 1100 language-related software applications.

Although the number of programs is considerable, it must be noted how primitive they appear in retrospect. The screen displays are, by modern standards, often crude (cf Davies 2001:5), limited by the low resolutions then available, the restricted range of colours, and the absence of graphical user interfaces, the only display option available being - partly due to the hardware but also to the limitations of early versions of BASIC - text mode which divided the screen into a number of rectangular boxes⁶ (for examples see Kenning and Kenning 1990:69-90). Even though, by the end of the 1980s, screen displays were becoming more sophisticated (op.cit.:42-44), it was still impossible to import images which meant that the visual display was only as good as the programmer's skills and could not compete with video. It may well be a general

5 accessible at <<http://www.hull.ac.uk/cti/resources/searchdb.htm>> (last accessed 12 March 2003)
6 Even in the 1990s, the maximum number of columns available in text mode under MS-DOS was 80 per row. The normal number of rows was 25 although 43 was sometimes possible.

lack of programming skills among many of the program creators that explains the decline in standards during the early 1980s noticed by Ahmad *et al* (1985:36).

It is the mechanical nature of much of the software that contributes to its apparent lack of sophistication. The presentation of information and entering of responses appear clumsy when compared to modern systems with graphical interfaces and mouse input. Much of this can be explained by the hardware limitations discussed above. Since then, however, the technology has made enormous advances. All standard computers now have GUI interfaces which means that not only does image quality match that of traditional film but also, thanks to parallel developments in storage media, that images can be viewed as video. All machines can now be equipped with sound reproduction devices with the result that images can be accompanied by realistic soundtracks. There is now software available capable of a credible level of voice recognition and an acceptable if not authentic level of voice synthesis. As a result, there has been a manifold improvement in the technical potential. How the realisation of the inherent potential of the modern computer has been influenced by the shifting fashions in teaching methodology will be examined in the next section.

3.3.2 CALL AND LANGUAGE TEACHING

As the variety of applications increased, so did the difficulty of classifying these in a paradigm that found general acceptance. The simplest classification of CALL activities is by reference to purpose, for example testing⁷, classroom aid or direct teaching (Cook 1985:13). However, this does not take us much beyond a classification based on activity type, understandable in the early days but inadequate later as a conceptual framework for interpreting and evaluating three decades of CALL. Cook does in fact prefer an alternative classification – *audiolingual*, *cognitive code*, *communicative* and *humanistic* (op.cit.:14-18) – based on the teaching methodology the activities serve, recognising that, at any specific period, this will be a powerful influence. Stevens (1992) uses a simple binary classification, *behaviourist* and *humanistic*, based on the amount of control and independence the software allows the student. Backer (1995), like

7 The use of computers for testing can be traced back to 1935 in the United States when an IBM model 805 was used to set and score multiple-choice questions on a massive scale.

Cook, relates developments in CALL to teaching methodologies – audiolingual, cognitive-code and communicative – although he prefers the headings *instructional* and *revelatory* borrowed from Kemmis *et al* (1977). Bax (2003) reverts to a descriptive framework based on types: *restricted*, *open* and *integrated*.

Perhaps the most comprehensive analysis of CALL is by Warschauer and Healey (1998:57) who identify three distinct stages: *behaviouristic CALL*, *communicative CALL*, and *integrative CALL*, with each corresponding to a certain level of technology as well as a certain pedagogical approach, but which also represent a progression through time and thus introduce into the analysis a historical dimension. However, these stages “do not fall into neatly contained timelines. As each new stage has emerged, previous stages continue”(op.cit.:58). The first of these stages, behaviouristic CALL, is described as “conceived in the 1950s and implemented in the 1960s and 1970s”(op.cit.:57), the second, communicative CALL, as having “emerged in the late 1970s and early 1980s”(ibid) while by implication the third, integrative CALL, which is seen essentially as a new “perspective”(op.cit.:58), is perceived as emerging during the 1990s. Two years later, Warschauer (2000:4) uses the same headings but moves each stage on in time by a decade.

In spite of the apparent terminological inconsistencies of the above taxonomies, there is a high level of convergence under the surface. There seems to be an implicit agreement on two phases of development for which the preferred labels are *behaviourist* and *communicative*. These echo the teaching methodologies discussed in the previous chapter and it does not seem unreasonable to expect that at any one period the prevailing teaching methodology would exert a major influence on CALL just as it did on other teaching materials.

There is less agreement about what to call the third and present phase of development. The term *post-communicative* is obviously unsatisfactory since it could be taken to imply that communication is no longer a goal. One possible solution to this problem would have been to adopt Warschauer and Healey’s (op.cit.:58) term *integrative*. This sees CALL as operating at several levels, by integrating social and cognitive knowledge through setting activities in social

contexts: by integrating CALL into the learning process and thus making it central rather than marginal (echoing Kenning and Kenning 1990:90); by integrating learners into authentic environments through task, project and content-based approaches, and by integrating the various skills. This can be achieved through the use of multimedia programs, concordancing software and exploitation of the internet (op.cit.:67). Bax (2003:5) is critical of this analysis on several counts. There is for him nothing new in setting language learning activities in social contexts. He also finds little evidence of the integration of CALL into language learning at the level suggested. He puts forward his own preference, *integrated CALL*, which largely involves the use of communications applications⁸, but recognises that it is an aim one should aspire to rather than a reality (op.cit.:6)

The disadvantage of both terms is that they are intrinsically technical and do not really take into account the aims and purposes of CALL. In the circumstances, it is proposed to follow Cook (1985) and Stevens (1992) and use the label *humanistic*, in the specific sense coined by the latter who saw it as “explicitly avoiding the rigid behaviourist-based approaches of earlier years” and “moving towards granting learners greater control over their own learning [...] by providing exploratory environments for language learning, presenting problems in need of resolution, and providing tools for further work and learning”(Stevens op.cit.:34). Not only does the term express a general philosophy of teaching, it excludes the behaviourist model without making any claims to being specifically communicative. An additional and important justification for its use here is that it reflects the methodological basis of the present experiment.

3.3.2.1 THE FIRST OR BEHAVIOURIST PHASE

The theories that lay behind the *behaviourist* model of language learning and teaching have been discussed in Section 2.2.2 above. Early CALL, certainly from the 1960s to the late 1970s, was very much an electronic extension of the practice of “programmed learning” or “programmed instruction”(Backer 1995:3)⁹. Language teaching very much became the inculcation of grammatical structures through a process of drill and practice, first in the language laboratory

8 which he also sees as being genuinely communicative (p7)

9 For a more detailed exposition of the influence of Skinner see Ahmad *et al* (1985:36-39)

and then in the computer room (Backer *ibid*; Stevens 1992:12). The appeal of CALL was, as Backer (op.cit.:3) observes, in large part explained by the computer's ability not only to free the language teacher from the drudgery of classroom drilling but also to manage much of the learning process. The computer was able to monitor, guide, evaluate and report on the progress of each student. The professed advantage for students was that they could work at their own pace, ideally choose their own level, solicit help and even feedback, and above all not be compelled to perform before the whole class. However, the outcome, as Backer (op.cit.:4) also observes, can be very different. The whole learning process is pre-determined. The student has to progress step by step, in a rigid order, a point that was noted much earlier by Higgins and Johns (1984:17): "[A]s long as the programmed learning paradigm dominated, pace was the only element that was individualised. The learner was not free to make significant choices about the sorts of activities undertaken, or the order in which they would be tackled". Every decision, every error (and sometimes even every key-stroke) was recorded and open to examination by the teacher. Far from being free to work at their own speed and cater to their personal needs, students found themselves totally controlled by the program¹⁰. Conceptually such a level of programmed instruction underpinned much of the software but, in reality, few institutions had the resources necessary for such an extreme situation and there was little appetite in the UK for such levels of control¹¹ (Ager 1986:199).

Drills were however not the only activity. In addition to vocabulary practice and text manipulation exercises (for examples see Kenning and Kenning 1990:54-60), the ability of the computer to monitor and record learners' input, combined with the need to restrict the number and type of responses, made it also an ideal medium for multiple choice exercises and even automated testing. In spite of evidence of disenchantment¹², such activities survived in CALL. The main factor was undoubtedly the attitude of teachers, many of whom had not only been taught but also trained to teach within a behaviourist framework (cf Backer 1985:3). Other teaching materials such as textbooks and especially language

10 For an ironical picture of what many educators saw as the pedagogic ideal see Brumfit *et al* (1985:1).

11 It is perhaps significant that use of the term *instruction* is rare in the UK compared to the USA.

12 The term 'drill-and-practice' was often replaced by the pejorative 'drill-and-kill' (Warschauer and Healey 1998:57).

laboratory materials were also largely behaviourist in conception and helped reinforce the prevailing methodology. Theoretical considerations were further reinforced by the fact that such materials are relatively easy to write, in part by the fact that many learners liked them, and in part by the fact that teachers and institutions were often unwilling to write off the investment involved in their purchase.

Because such activities have been widely labelled as behaviourist, it is perhaps appropriate at this point to consider to what extent activities such as drills are really behaviourist in conception. They were in fact very much a legacy from the grammar-translation era which saw the learning of morphological forms as being at the heart of language learning. The dividing line between an exercise and a drill is fairly fine. At a procedural level, it has much to do with the sheer number of repetitions, the immediate provision of correct responses and the refusal to allow the learner to proceed until the correct response has been entered. At a conceptual level, it is the intention of the person who designs or uses the activity and his or her perception of the role that cognition plays in language acquisition, rather than the nature of the exercise itself, that determine whether the activity is seen as behaviourist or cognitive.

It is even more difficult to claim as behaviourist gap-fill and text reconstruction activities. These are essentially cognitive. Their purpose is not to inculcate a set of habits but to encourage learners to use their existing knowledge to solve new problems. Insofar as they help students to see patterns of structure and usage, they are *revelatory* rather than *instructional*. Insofar as they require learners to hypothesise, they are clearly *conjectural*. Such activities predate the behaviourist era but seem to have acquired the label by association. They can justifiably be called *restricted* (Bax 2003:6) because they do not normally allow alternative answers. CALL activities which simply require students to repeat what they have just heard or to participate in a role-play by simply supplying rote learnt responses could, conversely, quite rightly be described as behaviourist, especially if they included rewards and punishments. Such activities are rare in CALL even if they were frequent in the classroom.

3.3.2.2 THE SECOND OR COMMUNICATIVE PHASE

If the behaviourist paradigm added little to CALL, apart perhaps from an aspiration towards ever greater control of both the learning environment and the process, it might be expected that the move towards a **communicative** approach would bring new goals and new types. As the communicative tidal wave gathered momentum (Munby 1978; Johnson and Morrow 1981; Johnson 1982; Lunt 1982; Brumfit 1983; Dunning 1983; Bate and Hare 1986; see also Section 2.2.4 above), CALL was recruited in the service of the new methodology. Underlying this was a belief that grammar should be learned implicitly rather than explicitly and that learners should be free to create their own utterances rather than manipulate prefabricated structures (Warschauer and Healey 1985:57).

In 1984 Underwood (quoted in Backer 1995:5-6) produced a comprehensive framework for communicative CALL based on 13 premises. The most important requirements in his view were a rejection of drills and evaluation by the computer, the implicit rather than explicit teaching of grammar, the encouragement of original use of language and the permissibility of multiple responses. In no way should the computer try to replace the text book. Instead it should encourage exploration and discovery on the part of the student and, above all, be fun. Examples of communicative CALL activities recommended by Underwood are simulations, communicative games (actually simulations perceived as games), text manipulation programs such as hypertext, storyboard, cloze formats, etc. and text generation programs such as poetry generators. By this time, improvements in hardware, and especially in memory and graphics capabilities, had made possible more sophisticated software. However, many of these programs were surprisingly, and disappointingly, familiar even if the presentation was slicker (cf Kenning 1996).

Ager (1986), while recognising that the computer cannot itself communicate, surveys both the programs and the literature then available and identifies certain parameters that, in his view, have a role in developing materials which, while not in themselves communicative, “can help in a language-learning programme which intends to follow communicative approaches”(op.cit.:193). He then

proposes a number of activities: exercises on “notions such as quantity, space or time”; “referential exercises” relating to the “natural world outside the computer” or to “a smaller universe of discourse” based on games or data-bases; “sociolinguistic exercises” focusing on the “illocutionary force” of language in the context of borrowing money or apologising for lateness; “contextual exercises” ranging from “sentence building” to the “identification of discourse markers indicating sections or sequences in the text”; “translation support programs” ; “note-taking exercises based on summarizing or expanding text”, and a “range of exercises based on Cloze gapping techniques”(op.cit.:197-199).

To what extent can such activities described be regarded as communicative? Backer (1995) seems to concede that some applications have the potential to be communicative: “They present microworlds for the students to experience where meaning is primary. The simulations and games can generate meaningful interaction among the students who are using them”(op.cit.:6). Warschauer and Healey (1985:57) make a similar point: “the focus was not so much on what students did with the machine, but rather what they (sic) with each other while working at the computer”. This is echoed by Higgins and Johns (1984) and Seedhouse (1995). Backer (ibid) challenges this view by stating that this argument is only valid in the context of a mixed learner group where there is no single shared language other than the target language, a condition which is rarely fulfilled in the average L2 classroom. Even when conversation takes place, the language produced can be very limited (Kenning and Kenning 1990:75) and may even be less effective than that generated by a non-CALL activity (Evans 1992). Another objection is that the argument Warschauer and others use could be applied to any activity, computer-based or not. Surely, if software is to be labelled communicative, it should engender student-machine communication rather than student-student communication.

Backer (ibid) also challenges the claim that text manipulation programs and text generation programs are inherently communicative. The main weakness, in his eyes, is their inability to embody meaningful interactions in the real world. He cites as an example the cloze technique which, though it draws on people's ability to infer meaning, remains for him simply a test of syntactic and semantic

knowledge¹³. In fact he sees such activities as belonging to the world of instructional CALL, suggesting that little has changed since the heyday of the audio-lingual method. Bax (2003:3) goes further, stating that, if the two key aspects of communicative language teaching, namely the notion that learners learn in order to communicate and that they may learn to do this best through the process of communication itself, are absent from so-called communicative CALL, “this must raise the suspicion that the term ‘communicative’ may not be appropriate to the software or the uses of CALL during the 1980s at all.”

It is difficult not to agree with him. Part of the problem is the fact that the term ‘communicative’ became such a blanket term that it became void of meaning. At times it seems that all an activity needed in order to be described as communicative was not to be overtly behaviourist. Had the activities been more modestly described as ‘pre-communicative’(cf Kenning and Kenning 1990:53), they would have attracted less criticism. However, they may well not have sold as many copies.

Bax and Backer’s criticisms could be considered excessive. The programs described require attention to be paid to meaning as part of the process of language acquisition within a variety of contexts and without the need for constant repetition. Success requires awareness of vocabulary and a range of grammatical features instead of an exclusive focus on a single linguistic form such as the perfect tense of French verbs conjugated with *être*. It could be claimed that such activities fall within the *conjectural* category of CALL by encouraging students to adopt a more exploratory approach. As a result of questioning each other and their teacher, the learning process becomes *revelatory*.

What Backer and Bax are criticising is not so much the activities themselves, the usefulness of which they largely accept, as the attempt to claim them as being communicative. This is implicitly recognised by Ager (1986.) who sees them more as a step towards achieving communicative competence. The point is developed further by Kenning and Kenning (1990:38-41). They base their

13 Research by Windeatt (1986) quoted in Stevens (1992:20) suggests that cloze exercises implemented on the computer may be harmful because students proceed one blank or line at a time rather than using the whole text and other blanks in a global manner to help them.

argument on the notion of communicative competence which, following Canale and Swain (1980), they break down into:

1. grammatical competence
2. sociolinguistic competence
3. strategic competence.

Because the first is essentially made up of a set of finite rules, it is capable of being processed by a computer and hence be incorporated into CALL. The other two are more complex, depending largely on individual and situational factors, and are, as a consequence, incapable of being incorporated into CALL in any but the most minimal way. Without them, however, the computer is unable to engage in real communication. This fact does not however disqualify it from being a very useful tool to aid the acquisition of grammatical competence which is, in the authors' view, an essential element of communicative competence. In this, they would presumably have the support of Widdowson (1984: 219): "Drills which provide repetitive practice for sound discrimination or the absorption of sentence patterns into habitual behaviour do not of their very nature deal with authentic language *as* communication. But they could provide an indispensable service in developing language *for* communication."

Most advocates of CALL during the pioneering years did not try to claim that it could be communicative. They were more concerned with the task of persuading teachers to embrace the new technology and saw the computer first and foremost as a patient and even friendly tutor whose role was to help learners to acquire outside normal language classes the vocabulary and grammatical forms that communication in a foreign language requires (Karttunen 1986; Hägg 1986; Brown 1988). Implicit in their arguments is a cognitive view of language learning. Until such time as computers can engage in conversation, it will be difficult to claim that they can generate genuinely communicative CALL. This will require much greater advances in both Artificial Intelligence and speech recognition and production, the present limitations of which are described by Davies (2001). However, as will be seen below, communicative CALL does not necessarily imply that one of the partners has to be the computer.

3.3.2.3 THE THIRD OR HUMANISTIC PHASE

At a minimum, humanistic activities should not only give the learner the maximum of control but they should also be *exploratory*. In addition, in Kemmis' (1977) terms, they should be *revelatory* and *conjectural*. Ideally, but not necessarily, they should also be *communicative*, at least insofar as the main focus should be on content rather than form. Familiar activities such as adventure games and simulations as well as applications such as databases and spreadsheets which have not been conceived as teaching tools, can be used as the starting point for language activities (Motteram 2000). The fact of having to interrogate a database or spreadsheet and extract and process information, all in the target language, can represent for the learner an act of communication, albeit somewhat formalised. It is perhaps communications applications that offer, at least in the short term, the best chance of fusing the communicative and the integrative (Motteram op.cit.). Requiring learners to use websites in the target language to hire cars, book *gîtes* or rail / air tickets or fill an on-line shopping basket¹⁴ (but stopping short of actually paying), likewise involves a process of two way communication, not with the computer but with the firm or organisation that has created the site (cf Vogel 2001). What such applications do is allow escape from the educational "dead-end" of a spurious communication *with* the computer and replace it by real communication *via* the computer. In addition, such uses achieve what Bax (2003:7) calls 'normalisation' - "the stage when the technology becomes invisible, embedded in everyday practice." When that day arrives Bax assumes that we will no longer talk about CALL, just as we do not talk about Book Assisted Language Learning or Pen Assisted Language Learning (ibid).

What is important in the humanistic paradigm is that the software is a tool enabling the learner to perform a task which will have two major goals, one the seeking of information, the other the acquisition of new language. Many of the activities mentioned above date from the early days of computer aided language learning. Humanistic CALL has been present from the beginning but obscured by being incorrectly labelled. Therefore, it is perhaps inappropriate to see this humanistic phase as following on from the behaviourist and communicative

14 These latter example are taken from the author's personal teaching.

phases. Rather, there have co-existed from the beginning the three paradigms adopted in this chapter. However, each has become in turn the prevailing methodological ideology that has shaped the world of language teaching and to this extent it is justifiable to speak of each as a phase. If radically new activities have arisen, this is due to technological advances which have opened up a whole area of technological communication and hence new possibilities for CALL. Perhaps these activities will even lead, as Bax hopes, to the demise of CALL. For the humanistic model to succeed, it will be important, as the technologies create new opportunities, for teachers to seize them not in the service of a monolithic methodology but for the greater empowerment of their students. Concordancing is seen as one possible means of empowerment, allowing students to interrogate the raw linguistic data directly and encouraging them to observe patterns and seek explanations, without however denying them the support and guidance of the tutor. The various applications of the concordancer to language learning and teaching will be the subject of the next chapter.

3.3.3 *CALL IN HIGHER EDUCATION IN THE UNITED KINGDOM*

The previous section has examined CALL in the wider national and international community. It must however not be forgotten that the experiment which is the object of this study is set within Higher Education in the UK and has to be situated within the wider institutional context. There are two principal sources of information on the application of CALL in British universities. The first comprises specialist journals such as *ReCall* and *System* as well as conference reports such as Thompson and Chesters (1994) in which individual practitioners describe their personal or institutional initiatives in this field. What emerges from these sources is an amazing variety of applications which almost defy categorisation but which, for the most part, fall under one of the headings mentioned above (behaviourist or communicative, instructional or revelatory). For more systematic information, it is necessary to refer to a small number of surveys covering the 1980s and early 1990s.

In a report based on a conference held in April 1987, within the framework of the UK Higher Education Computers in Teaching Initiative (CTI), called to map the use of CALL across the country, Chesters and Gardner (1987) provide

examples of activities in individual universities. The activities described include multiple choice questions as part of a text-only adventure game in which a British pilot shot down in Occupied France has to ask questions in correct French; a beginners' course in Latin requiring syntactic and grammatical analysis of sentences; grammar and reading comprehension exercises in Italian, mainly but not exclusively for beginners; the modification of a wordprocessor to teach phonetic transcription (mainly of mediaeval French); and a speculative scheme for analysing French poetry. There is also a useful survey of computational text analysis (Butler 1987) which both identifies the problems and provides a picture of current activity. What distinguishes several of these projects is their "artisanal" quality involving programming in BASIC or SNOBOL and liaison with computer services. Some were still tied to the mainframe but others made use of the PC, predominantly the BBC micro. What is apparent from several contributions is the technical difficulty of programming and this no doubt explains the significant move away from the "home-made" to the "ready-made" that has taken place in the period between the publication of the 1987 CTI report and later reports.

A slightly later survey of UK universities carried out in 1992 (Thomson 1993) showed that CALL packages (essentially the same as those used in schools and including grammar-drills and games) were the most widely used type of software, closely followed by word processing and authoring packages. Behind these came databases, CD-ROM and interactive video, and, even further behind, teleconferencing and "other methods [including] desk top publishing, simulations, speech analysis, electronic mail, packages for the analysis of text corpora, and translation tools developed in-house"(op.cit.:8). The "ready-made" packages were largely used for grammar development with less advanced groups (ibid) while one suspects that the "other methods" reflect individual lecturers' teaching or research interests. In slightly more than half of institutions that responded CALL was integrated into the syllabus (ibid). Given the advances in both hardware and software, especially of GUI based operating systems, the limited range of activities might appear surprising. However, the report also identifies a number of negative factors likely to inhibit both the development and adoption of CALL, among which are lack of official recognition of the field for

research and publication, the feeling that institutions saw the introduction of CALL as a means of reducing staff numbers, and a perceived lack of suitable software for advanced students¹⁵.

Of relevance in this context is the Technology Enhanced Language Learning (TELL) project financed in the 1990s by the UK Universities' Funding Councils as part of the Teaching and Learning Technology Programme (TLLP). The project, which took the form of a consortium of 37 universities, was set up to develop and disseminate CALL materials for use in HE. In her report evaluating the project, Hewer (1998) found that the three most widely used items of software were: a CD-ROM multimedia package (*Encounters*) designed to improve listening and speaking skills via a number of loosely communicative scenarios; a package of grammar exercises (*GramEx*) with inbuilt correction and monitoring procedures linked to sister programs providing definitions (*GramDef*); a basic translation passage (*TransIt Tiger*) providing help, including with vocabulary, but lacking any correction procedures apart from the provision of one or two model versions for comparison purposes. None of these packages could be described as communicative. All of them could have been conceived in the 1980s. Nevertheless, they can be useful in developing grammatical skills and translation techniques, especially among weak students, for whom the main priority is plausibly not conversation with a foreign native speaker but a good grade in their assessments. Anecdotal evidence suggests that such activities are quite popular with students.

3.4 CONCLUSION

A survey (Fox *et al* 1990) carried out into the use of technology in language learning across the whole range of education in the UK at the end of the 1980s speaks of "disappointment and even disillusionment among users" which it ascribes in part to excessive expectations but also, more significantly, to the practical difficulties in developing "human communicative skills", especially in the area of oral language work (op.cit.:21). It is interesting to note the then contemporary preoccupation with the communicative and the naïve expectations about the role that the computer could play. The authors' assessment of the

15 The same complaints can be heard at almost any CALL conference or meeting today.

situation is generally negative, finding that development was largely at a standstill and that, apart from small group and self-access work, software was of limited usefulness (ibid). What optimism they express rests mainly “on extensions to CALL, on communications technologies and, for the future, on multimedia applications”(ibid).

The feeling of disenchantment with existing CALL software and the identification of multimedia as offering the greatest potential for the future find their echo in Davies (1997) and Warschauer and Healey (1998) and anticipate the developments discussed in the previous section. The future, in the view of the latter, lies as we have seen in their third ‘integrative’ stage of development in CALL which should seek to combine more closely technology and skills’ development. This would require students to use a variety of technological tools integrated into the process of language learning as opposed to a weekly session of isolated exercises (op.cit.:58). They provide details of a number of commercially available programs that exemplify for them the new style of software that is coming on to the market. While recognising that there are still many grammar and vocabulary drill programs available, they see many of these as being both more contextualised and more sophisticated in their use of graphics, audio recording and playback and video (op.cit:59).

The types of programs that the authors list are: drills (often integrated into games), simulations, strip stories and text manipulation, crosswords and word search puzzles, outliners and word-processors, and finally concordancers (op.cit.:59-61). The benefits of using the computer they list as multimodal practice with feedback, individualization in a large class, pair and small group work on projects, the fun factor, variety in the resources available and learning styles used, exploratory learning with large amounts of language data, and real-life skill-building in computer use. The authors perceive a shift towards communicative CALL and yet the activities they cite are open to Bax’ (2003) criticisms quoted above. Many of the programs mentioned are familiar and do not represent any great methodological breakthrough. None of them is new and many echo those proposed by Last in 1984 (op cit:100-102). The problem would seem no longer to be one of justifying the use of CALL in methodological terms

but more one of addressing the practical and organisational implications of its implementation (cf Hewer *et al* 1999).

This seems to indicate that the last twenty years have added little in the way of activities to CALL. However, programs have become more user-friendly, largely as a result of GUI operating systems on both IBM PCs and Macs, and less mechanical, helped in part by the flexibility allowed by applications such as hypertext. Apart from increasing the visual attractiveness of programs, these developments make it easier for users to move backwards or forwards within a program, to exit or to solicit information or help. One activity that seeks to avoid the trap of “drill-and-kill”, because it can be seen to put the student in control, both physically and mentally, is concordancing which requires concordances to be first constructed and then analysed. It thus fits into Stevens’ (1992:11) humanist model in that it is “an informant on language rather than a preceptor, task-master, or programmed instructor”. Concordancing is usually associated with an area of computer use that is not always perceived as belonging to CALL and that is corpus linguistics. Although many of the authors mentioned above refer to concordancers (Kenning and Kenning 1990:60-61; Stevens 1992:17; Warschauer and Healey 1998:67), they do so briefly, almost *en passant*. Set against this is Tribble’s (1990) resounding endorsement: “Because it is relatively simple to use, does not require authoring before it can be deployed, and has an exceptionally wide range of applications, the concordancer will perhaps be the pre-eminent software tool in this next stage in the development of computer-assisted language learning”(op.cit.:475). Another advantage perceived was that using a concordancer would avoid the types of student dependency that some CALL programs can induce (Littlemore 2001:43). The various uses to which the concordancer has been put are examined in the next chapter.

CHAPTER 4

4 CONCORDANCERS IN LANGUAGE LEARNING AND TEACHING

4.1 INTRODUCTION

Since the pedagogic potential of the concordancer is central to the present study, it is necessary to survey the field and evaluate previous practice in order to learn from the experience of others. Most of the references to the use of concordancing in language learning and teaching are to be found in articles. Only two books have been published focusing exclusively on the use of concordancers (Tribble and Jones 1990 – updated and republished in 1997 - and Johns and King 1991), the latter of these being in effect a separate imprint of an issue of the ELR Journal. Both describe potential classroom applications in the hope of encouraging wider use. The 1990s witnessed a growing interest in concordancing primarily in the context of the construction and exploitation of corpora and it is here that one finds many of the later contributions (Aijmer and Altenberg 1991; Sinclair 1991; Barnbrook 1996; McEnery and Wilson 1996; Stubbs 1996; Thomas and Short 1996; Habert *et al* 1997; Wichmann *et al* 1997; Biber *et al* 1998; Hockey 1998; Kennedy 1998; Oakes 1998; Ooi 1998; Burnard and McEnery 2000). In general, though, the concordancer is mentioned only insofar as it is a necessary tool in the analysis of the corpus material (cf Blanche-Benveniste 1996). The focus is primarily on *theoretical*, as opposed to *applied*, linguistic, research with only Wichmann *et al* (1997) being exclusively devoted to applications in teaching. As the present focus is on the use of concordances in language teaching, and specifically in the teaching of French as a foreign language, sources which deal exclusively with concordancers as a pure research tool or with the construction of corpora will only be included here if they are felt to throw useful light on the application of concordancing to language learning, albeit indirectly.

Most of the examples quoted of the use of concordancers in language learning/teaching are from EFL/ESP in the UK and elsewhere in the world, primarily Australia and the Far East. There seems to have been little interest in using concordances for L2 teaching in the UK or US, at least until relatively recently (cf Wichmann *et al* 1997). This raises the question of to what extent

applications cited for EFL are transferable to French. Practice in EFL provides a useful model which cannot be ignored but there are aspects of French (gender, complex verb inflections, subjunctive, etc) that do not exist in English and for which the latter cannot provide models. What the sources quoted do provide, in addition to precise examples, is a theoretical and pedagogic framework for the present study. Only by experimenting is it possible to ascertain how easily and how far transference can be achieved.

Before examining the application of concordancing to language teaching, this chapter will trace in Section 4.2 the manner in which the process has been transformed by developments in computing and the influence that these have had on modern linguistics. It will then be possible to explore in Sections 4.3 and 4.4 the role that the concordancer has played in language teaching both outside and inside the classroom.

4.2 THE DEVELOPMENT OF COMPUTER CONCORDANCING

4.2.1 CONCORDANCING AND THE COMPUTER

Concordancing as an activity can be traced back to the Middle Ages. "In its original sense a concordance is a reference book containing all the words used in a particular text or in the works of a particular author (except, usually, the very common grammatical words such as articles and prepositions), together with a list of the contexts in which each word occurs. Each context may be indicated by means of a precise line reference, or by way of a short citation, or by both"(Tribble and Jones 1990:7). The authors cite as the first example of concordancing the case of Hugo de San Charo (also known as Hugues de St Cher) who, in the 13th Century, employed 500 monks to produce a complete concordance of the Latin Bible. In this case the concordancers were the individual monks and the task was long and laborious. As the authors further remark:

Computers have changed all this. Provided a text is stored in electronic form, a suitably programmed computer can perform all the tasks involved in compiling a concordance - locating all the occurrences of a particular word and listing the contexts - very rapidly and absolutely reliably. Scholars were not slow to realise this, and so computer programs, called "concordance generators", "concordancers" or, generically,

“concordancing software” were developed specially for this purpose (ibid).

The first attested use was in 1962 at Cambridge University in the analysis of a German literary corpus (Ahmad *et al* 1985:39).

Another application of modern concordance software is to provide information on word frequencies. To appreciate how computers have transformed this task, one only has to consider the example of Kaeding who, in Germany in 1897, “is reported to have worked with over five thousand assistants on a corpus of 11 million words to gather statistical information on the German language relevant for the training of stenographers”(Kennedy 1998:335). Not only can the task now be performed by a single computer in minutes but the output can be almost instantly rearranged in ascending or descending order or filtered to highlight items in a desired range of frequencies, tasks which until recently required massive manual re-editing. It should not however be imagined that the transition from processing by hand to processing by computer was instant and painless. Kennedy (op.cit.:6) also quotes the case of scholars indexing Indian Dravidian scripts in the early 1970s when all input data had to be laboriously punched on to tape by hand. However, even this was preferable to making individual card indexes of every word in the source texts, a task which, in the case of one text of nearly 12,500 stanzas, each with four lines containing on average six words, took six years to complete and amounted to 3,500 typed pages¹.

The early concordancers capable of performing all these functions were mainframe based and required familiarity with a complex programming language (Reed and Schonfelder 1979). This was still true of the Oxford Concordance Program which appeared in the 1980s and only slightly less true of the PC version, Micro OCP (Urdang 1991). Hardware and software constraints meant that use of concordancers was restricted to literary and linguistic research. Among the first to suggest the relevance of concordancing to language teaching were Davies and Higgins (1982): “Although the microcomputer is at present incapable, say, of producing a complete concordance of the works of Shakespeare, it can certainly handle poems and short stories and demonstrate to

1 For a much less ambitious approach to “manual” concordancing see Willis (1998).

students the scope of electronic wizardry in stylistic analysis”(op.cit.:2). However, largely because of the limitations of the hardware, the authors still see the concordancer as primarily of use outside the classroom: “Teachers who are more concerned with the practical aspects of language teaching will appreciate the value of concordance packages in checking the frequency and intervals of repetition of lexical items in course books and in the production of dictionaries, where contexts are an essential element”(ibid).

Almost the only people in the 1980s to see the full potential of concordancing were Higgins and Johns (1984) although Last (1984) does devote several pages to the topic, remarking that the field seems to be dominated by mediaevalists and Germanists. Otherwise mention of concordancing is fairly rare and usually brief. In general, during the decade most commentators (discussed *passim* in Chapter 3) see the future lying in either highly structured CAI/CALL programs, falling within what Higgins and Johns (1984:18), after Kemmis (1997), call the “instructional paradigm”, or in artificial intelligence (Brumfit et al. 1985 and Cameron 1989, in addition to those mentioned above). The limitations of the available hardware and software were very real but may not be the only explanation for the lack of awareness of the possible uses of the concordancer. Part of the explanation may lie in the move towards a so-called “communicative” teaching methodology (see Section 2.2.4 above) and the consequent emphasis on non grammar-based interactive learning activities, largely oral and of an everyday transactional nature based on “real-life” scenarios. Text-based work was consequently downgraded and metalanguage banished from the classroom.

It required the commercial success of the microcomputer² and the development of programs for use with MS-DOS such as the Longman Mini Concordancer, Micro-Concord, and WordCruncher before concordancing as widely practised today became feasible. The fact that the first version of Micro-Concord, written for the Sinclair Spectrum, contained approximately 1k of code, could only work on texts up to 20k and store at maximum 180 citations (cf Johns 1986) shows how crucial the development of the modern PC was to computational

2 The disadvantage of this shift was that in future much of the development work would be carried out on Windows or Mac platforms rather than on Unix with its superior text handling functions (cf Lawler 1998).

linguistics³. Once the breakthrough had begun, the potential of concordancing software as a tool for language teachers, mainly in higher education, was quickly perceived. It was only when computers became widely available for student use, at least in vocational and higher education, that it became practical to give students personal experience of using concordancers as part of their learning experience. Among the essential criteria for student use, Stevens (1995) identified: speed of loading; speed and flexibility of sorting, and the ability to interrupt searches. With the arrival of the new generation of PC and Mac based concordancers, these criteria could be met. Concordancing could henceforth be integrated into the language learning experience of large numbers of foreign language students. What had only been possible in a few well-resourced institutions with well-equipped student resource centres and staff with the expertise to produce the necessary software (cf Stevens 1988) became, at least in theory, possible in the majority of institutions.

4.2.2 *CONCORDANCING AND MODERN CORPUS LINGUISTICS*

The development of concordancing software has run in parallel with the development of large computerised language corpora. The latter have only become possible as a result of the exponential increase in the capacity of the modern computer's memory. However, the construction of a computerised corpus is only a continuation of the well established tradition of basing deductions about language on a body of written texts or extracts of texts (cf Grévisse 1955: preface). The development of large scale computerised corpora intended for language research began in the English-speaking world (for a detailed account see Kennedy 1998:23-60). The first was the Brown University corpus constructed during the 1960s in the United States followed by a joint British and Norwegian corpus of English (LOB or Lancaster-Oslo/Bergen project) which was completed in 1978. Since then the number has increased rapidly⁴. The Birmingham University Collection of English Text was the basis for the construction in the 1980s of the COBUILD Corpus of written and spoken text which has been used to develop dictionaries (*Harper-Collins*) and grammars. Recently more attention has been paid to the spoken language,

3 Even as late as 1995 the maximum number of words that the Longman Mini-Concordance Program (MS-DOS version 1.01) could handle was 50,000 (Qiao 1995:1).

4 Dossena (1996:246) lists twelve English corpora.

leading to the completion in 1987 of the Lancaster-IBM Spoken English Corpus (SEC). Gradually corpora in other languages (ABU, FRANTEXT and SILFIDE⁵ are French examples) are being developed but until recently none had reached the size of the largest English ones - the British National Corpus compiled between 1991 and 1994 contains 100 million words, The Bank of English⁶ has a total holding of over 500 million words, with 200 million accessible on-line, whereas the Spanish corpus⁷ created by Mark Davies of Illinois State University contains 100 million words and the *Deutscher Wortschatz* project now claims 123 million words (Quasthoff and Wolff 1999). Yet the idea of a corpus ran counter to the prevailing essentially Chomskyan linguistic orthodoxy of the period (cf Tognini-Bonelli 2001:50-52) which held that the truth of language was to be found in the innate cognitive knowledge of the native speaker and not in some external body of texts or utterances.

Such was the mood of the times that the sober virtues of computer research had little power to counteract the abstract intellectual fascination of generative grammar, based on the 'evidence' of native speaker intuitions. For me and for many others, the excitement of linguistics lay in discovering universal constraints on the structural diversity of languages, which were held to reflect the structure of innate language-processing machinery built into our minds as part of the human genetic endowment (Sampson 1996:14).

As a result, "objective data were not really relevant: either they corroborated the speaker's intuitions, in which case they offered a slow and cumbersome way of getting at facts which intuition revealed more immediately, or they contradicted intuition in which case the objective data were suspect"(op.cit.:15).

One important factor in sustaining the development and exploitation of corpora, in spite of the lack of interest on the part of most linguists, was the increasing involvement of publishers of dictionaries who saw the advantage of basing their lexica on examples of real usage backed up by figures indicating relative frequencies (see Section 4.3.1 below). The incorporation of syntactic and semantic information into the corpus databases has in its turn given linguists ideas for more sophisticated exploitation of the data, mainly through the

5 for more information see Romary *et al* (1999)

6 see <http://www.cobuild.collins.co.uk/> for details

7 see <http://www.corpusdelespanol.org/> for details - for details of associated corpora in Spanish and Portuguese see Davies M (2000 pp.174-175)

increasingly automatic parsing and tagging of the textual material (cf Garside *et al* 1997). In parallel, a number of smaller specialist corpora focusing for example on aspects of spoken (Beeching 1997) or scientific language (Gledhill 1996) have been or are being built up. Many of these projects are located in one or two universities as part of a programme of research. More recently, the widespread availability of easily accessed material on the Internet and on CD-ROM has made it possible to construct *ad hoc* corpora for specific and usually limited purposes (Inkster 1997; Maia 1997; Minugh 1997; Pearson 1996; Rézeau 1997; Tribble 1997). There is no doubt that the 1990s saw corpus-based approaches gaining momentum (Leech 1997), whether their partisans supported the “strong case [which] suggests that without a corpus (or corpora) there is no meaningful work to be done” or the “weak case” according to which “there are additional descriptive and pedagogic perspectives facilitated by corpus-based work which improve our knowledge of the language and our ability to use it”(Murison-Bowie 1996:182). First however it is necessary to have access to a corpus and it is the question of corpus design and construction that will be the focus of the next section.

4.2.3 *CONCORDANCING AND CORPUS DESIGN AND CONSTRUCTION*

Concordancing is impossible without the raw linguistic material on which to work. A corpus differs from a library or archive in that it is a “body of texts assembled according to explicit design criteria for a specific purpose”(Atkins *et al* 1992:5). It can be conceived as a *reference* corpus to serve as a benchmark of a given language at a given period in time or as a *monitor* corpus which follows changes in the language and can be added to or even destroyed if it has outlived its usefulness. The mass of selected texts comprising the corpus is often referred to as the ‘population’ (op.cit.:4-5) and, just as with opinion polls and other surveys of human populations, the results are only considered valid if the sample is representative (Biber 1993a). In the case of corpora this means deciding at the outset whether the sample is to represent the whole range of a language or certain regional, social, educational or professional sublanguages, whether it is to include written or spoken language or both, which genres will be included, what time span the sample will cover, whether it will differentiate between age groups

and sexes. Will it be made up of whole texts or samples⁸ of texts? A reference corpus should ideally be balanced (Atkins *et al* 1992:6) and not skewed in favour of any one group or type. The authority of any conclusions that are drawn from the corpus depends upon their ‘external validity’, which requires avoiding both ‘random error’ and ‘bias error’: “Random error occurs when a sample is not large enough to accurately estimate the true population; bias error⁹ occurs when the selection of a sample is systematically different from the target population it is intended to represent” (Biber 1993b:219-22). It is therefore essential for theoretical linguists, lexicographers and writers of standard grammars that the sample should have the maximum representativeness and validity (for a discussion of issues such as sampling and representativeness see Tognini-Bonelli 2001:47-63).

Such issues raise the question of **size**. If a corpus is to be truly representative of a language in all its variety it must seek to be as comprehensive and, by implication, as large as possible. Only a few years ago the target figure for an electronic machine-readable corpus was a million words (Sinclair 2001:viii) but with modern computing power ten to twenty million words can be considered a “useful, small, general corpus” (Sinclair 1991:20) in a world where “a corpus should be as large as possible, and should keep on growing” (op.cit.:18). Exponents of **large corpora** are mainly theoretical linguists who “view corpora as a mass exhibit *in extensi* of the facts of a language, yielding data on the relative frequency of phenomena of all kinds, and in particular providing a check on their own, or their informants’ intuitions”(Atkins *et al* 1992:14). Among them can be counted professional lexicographers and grammarians.

This attitude of “the bigger the better” has been challenged and from the early 1990s on there has been increasing interest in **small corpora** (Biber 1993b; Aston 1997b; Tribble 1997; Ghadessy *et al* 2001). Small corpus supporters tend to be applied linguists¹⁰ who are not seeking universal truths about the language but precise examples of usage to serve as models for their students:

8 It is important here to distinguish between *sample* in the sense of an extract from a text (hence a *sample corpus*) and *sample* in the sense of a representative cross-section.

9 One source of bias error can be the lack of balance in the text population.

10 An exception is Bernardini (2000) who argues the case for using a large corpus (BNC) with students.

In teaching a second language, corpora provide a substantial resource for extracting and studying authentic language data with the authority of attested use. This data might be presented directly to students for classwork or may underpin the preparation of teaching materials of every kind.[...] Subcorpora of restricted topic areas may be particularly appropriate for use in teaching Languages for Specific Purposes (ibid).

An example of such a **subcorpus** is Gledhill's (1996) scientific corpus of cancer research articles referred to in 4.2.2 above containing over 500,000 words which was created in order to investigate typical patterns of discourse in a specific genre. Another is Someya's (1999) corpus of business letters. What distinguishes a subcorpus from a small corpus is the fact that the former is a reference corpus designed to identify the words and features that are typical of a certain type of text or genre or social or professional group whereas a small corpus is usually less systematic in conception and constructed for a specific purpose. However, there is some overlap between the two types and a small corpus, if its coverage is sufficiently comprehensive, may well constitute a subcorpus.

By 1993 a number of practitioners were using a variety of small corpora ranging from 6,853 words to just under a million words drawn from, among others, academic lectures and journals and covering subjects as varied as biology and philosophy (see Flowerdew 1993b:90). One example of a small corpus created to serve as a reference source for the production of teaching materials is Facques' (2002) collection of 146 articles drawn from the major French daily newspapers over a two week period and designed to investigate uses of the past tenses. Another is Polezzi's (1993 and 1994) corpus composed of a series of short lectures on the Italian Renaissance and containing 10,209 words. From it was derived a set of teaching materials for use in an intensive *ab initio* course for ERASMUS students about to spend a term in an Italian University.

The arrival of the modern desktop computer and especially the generalisation of the Internet have made it much easier for individuals and small departments to create *ad hoc* corpora suited to their individual needs¹¹. This is especially true in the field of translation studies (see Section 4.4.5 below). What is arguably the smallest corpus is described in Bowker (1999) and consisted of a single text of

11 Detailed guidelines for constructing a special purpose corpus are to be found in Bowker and Pearson (2002:43-133).

100 words taken from a specialist French computing magazine for use in a translation class. One strategy is to use a commercial CD-ROM such as MicroSoft's *Encarta* and create from it topic or theme-based subcorpora (Tribble 1997). Another strategy is to transfer the burden of constructing the corpus to the students (Gavioli and Zanettin 1997; Maia 1997; Schäffner 1998; Zanettin 2002). Such a procedure has the added pedagogical advantage that it forces the students to assess the quality and suitability of the texts they find and can help make them more sensitive to features such as register as well as content. Whatever the method of construction, small corpora will by their nature be unbalanced and unrepresentative of the language in general and, if they are to be used with students, the teacher must make students aware of this fact in order to avoid potential over-generalisations from a restricted sample (cf O'Keefe and Farr 2003).

A more recent but easily available source of corpus materials is the **Internet** (cf Maia 1997; Varantola 1997; Foucou and Kübler 2000; Ooi 2001; Zanettin 2002) and this was the solution adopted for the present experiment (see Section 5.2.1 below). The advantage is the sheer amount of material available. Fletcher (2003a:1) estimates that by the autumn of 2001 there were between 3 and 5 billion publicly-indexable pages and that this number is expected to grow to between 10 and 15 billion by the middle of the decade. According to Fletcher 85% of these pages are in English (ibid) although the number of pages in other languages, including French, is growing at a proportionally greater rate (op.cit:2 and nn). What makes the Internet particularly tempting is the availability of powerful search engines such as GOOGLE which facilitate the task of seeking texts. These can be added to a corpus at any time and disposed of with no visible financial cost if they prove to be unsuitable. The disadvantage of search engines, however, is that they can retrieve several thousand pages in one go, many of which may have moved or simply disappeared. In addition, many of these texts may be ephemeral and of "dubious authorship and authority" (Zanettin 2002:3; cf also Friedbichler and Friedbichler 1997:3; Pearson 2000; Fletcher 2002:2). What is true of the quality of the content may well be true of the quality of the language¹², with the result that disappointingly few will be appropriate as

teaching materials (Fletcher 2003a). Another disadvantage of the Internet is that students may not see it as a serious tool for language learning (Vogel 2001).

It must also be borne in mind that texts retrieved from the Internet may well have characteristics that are not typical of the language as a whole. A thorough evaluation of the potential of the Internet was carried out by Ide *et al* (2002) as part of the American National Corpus (ANC) project. Because of the claims of Fletcher (2003b) and others (Cavaglià and Kilgariff 2001; Kilgariff 2001b) that the Internet is making traditional methods of text-gathering redundant, they performed an experiment to test (1) whether the Web removed the need for a corpus such as the ANC and (2) whether selection of texts could be performed automatically (op.cit.:2). Instead of relying solely on mathematical models (cf Caviglià and Kilgariff 2001), they used a combination of quantitative tests (e.g. word to paragraph ratios) and qualitative tests (see Biber 1995) to classify the texts into four main groups: institutional documents, expository, science, prepared speeches (Ide *et al*, op.cit.:4). The range of texts selected was fairly narrow, partly because of the difficulty in guaranteeing that the texts retrieved were in fact by American native speakers which obliged the researchers to restrict their search largely to governmental and educational sites in the US. The texts, totalling 2,190,196 words, were analysed and compared with known scores for other types of text. The researchers found that “web texts are more informationally dense and elaborated than texts from others sources” (ibid) and that they “can be regarded as falling into a genre of their own”(op.cit.:6). They conclude: “Our study suggests that web-based texts are, in any case, representative of only a small slice of the range of genres encountered by human readers everyday, and therefore cannot be used to provide a comprehensive view of American English in the 1990s”(ibid). It seems reasonable to assume that their findings can be applied to other languages.

The arguments surrounding the issue of the Internet highlight some of the problems associated with large reference corpora, such as the minimum and optimum sizes, the representativeness and balance of the population, the sampling criteria to be achieved to meet these goals. There is another problem

and that is the question of **comparability**. For example, at what point will the ANC be comparable with the BNC or the Brown-LOB corpus? If the corpora are not comparable, does that make any comparisons, for example between American and British English, invalid? The question of comparability is highly complex and requires familiarity with a whole range of statistical procedures (cf Oakes 1998; Kilgariff 2001a) which the average linguist may well find intimidating. The question is particularly relevant to corpus based translation study (see below).

Recent years have seen the development of **learner corpora** which are designed specifically to focus upon the language produced by students either in their mother tongue (Biber *et al* 1998) or in a second language (Granger 1994; Seidlhofer 2000; Cobb 2001; Ragan 2001; Granger *et al* 2001). There is nothing new in the practice of having students analyse their or other students' work as part of the learning process and several authors have brought the practice up-to-date by collecting coursework in electronic form and converting it into a subcorpus, usually quite small, which they then integrate into their teaching (Ching & Wong 1994; Qiao 1995; Turnbull and Burston 1998). In the literature, there is surprisingly little emphasis upon error analysis. Instead the emphasis is mainly on style and fluency, focusing on features such as over and under-use of words and phrases in comparison to native speakers. One advantage for the researcher of learner corpora is that it should be easier to get copyright clearance from the text owners (for more detailed discussion of the question of copyright see below).

All the examples quoted above have been of **monolingual** corpora but in recent years there has been increasing interest in creating **bilingual** and even **multilingual corpora**¹³. These are designed to be employed in a contrastive linguistic framework, serving to highlight systematic differences between languages (prosodic in the case of Roussel 1991, lexico-semantic in the case of Salkie 1997 and 2001), often within the context of training translators (Bernardini 2003), and sometimes combining both (Johansson 1997). Such

13 Details on many of these can be found on-line at <<http://www.ruf.rice.edu/~barlow/para.html>>. It is possible to build one's own parallel corpus using on-line sources such as *Label France*, which is multi-lingual, and *Courrier international* which contains many translated texts.

corpora can be *comparable* where texts “are put together on the basis of similarity of text-type and topic” (Gavioli and Zanettin 1997:Zanettin 1998), *parallel* where texts in language B are translations of texts in Language A (Bernardini 2003) or *aligned* where “every sentence in the source text matches its translation in the translated text”(Salkie 2001:1). Because of the difficulty in finding comparable or parallel texts in two or more languages¹⁴, such corpora tend to be relatively small, one million words in the case of the CEXI corpus half of which is made up of comparable texts and half of parallel texts in English and Italian (Bernardini op.cit.:530). Such corpora may deal with highly specific fields and thus contain specialist parallel subcorpora numbering no more than a few thousand words (Gavioli and Zanettin op.cit.:2).

The use of small corpora highlights one of the potential weaknesses of large corpora. Because the latter are designed to be normative, their design often excludes or marginalises highly specialist fields or texts precisely because of their perceived abnormality, from fear that they may unbalance the corpus and skew it in a particular direction. As a result of the balance or timespan chosen, even common words may be missing. The 1999 BNC Sampler (two million words) does not contain a single instance of “internet” or of the related sense of “web” so a learner trying to see if the words were synonymous would draw a blank . The only instance of “greenhouse effect” is in a poem by Benjamin Zephaniah, a fact which would give little help to learners in gauging its current frequency or the range of genres or text types it is to be found in.

There remains one highly complex issue relating to the compilation and use of corpora, the question of **copyright**. As soon as an individual creates a work that has physical substance, that person enjoys copyright: “Under the Berne copyright convention, which almost all major nations have signed, every creative work is copyrighted the moment it is fixed in tangible form”(Templeton n.d.). The creator does not have to append a statement to this effect or use the conventional symbol and does not have to make a charge. The question of copyright has become more acute as the Internet has developed and made so many texts immediately accessible. Even if the work is made publicly available, as is the

14 A relatively unusual monolingual corpus comparing translated and non-translated texts and designed to investigate the phenomenon of simplification is described in Laviosa (1998).

case with the Web, the copyright remains with the author. Although the law is complex and has not been fully tested in the courts, the presumption seems to be that the Internet enables the researcher to consult but not store and use the material posted, unless invited to do so by the copyright owner or in cases where the material is in the public domain, as is the case with US government documents, or unless it falls under the heading of 'fair dealing' which means that "you are allowed to make one copy for your own research or self-study"(CIEL 2001:14) Making such material available to students would not be covered. Even where copyright clearance has been given, the educational institution will probably have its own regulations and will no doubt require written confirmation before authorising the storage and accessing of any corpus material by students. One ingenious suggestion for avoiding the copyright problem is to create what Kilgariff (2001b:1) calls a "virtual multicorpus" which brings together individual corpora each made up of a set of URLs. Upon accessing the corpus, the researcher would be pointed to a carefully monitored bank of references to texts classified according to strict criteria of type and quality. As a result, nothing would be stored centrally. The developing of web-concordancers such as KwiCFinder (see Fletcher 2003a) makes such a proposal much more feasible.

The size, scope and ultimate purpose of individual corpora can, as has been seen, vary enormously. About the only constant in the world of corpora is the role of the concordancer. Exploitation of a corpus may at times require highly specialist software tools (Barnbrook 1996:chapter 6; Biber *et al* 1998:chapter 3; McEnery and Wilson 1996:chapter 5) but the workhorse remains the concordancer. A corpus may well be "a rich resource of authentic data containing structures, patterns and predictable features that are waiting to be 'unlocked' by the human intelligence" (Leech 1997:3), but access requires a key. It is the concordancer which allows the researcher to search the corpus for the presence of an item, to establish its frequency and the "company that it keeps" and it is precisely these features of the concordancer that make it potentially so useful in teaching. The role of concordancing in both course design and classroom practice will be examined in the next section.

4.3 CONCORDANCERS AND THE TEACHING CONTEXT

The construction of corpora and subcorpora specifically for the purposes of language teaching has already been discussed in Section 4.2.3 above. The objective quantitative data about the frequency of words and constructions derived by the concordancer from these corpora provide a foundation for the preparation of course books and other materials. It is perhaps not surprising, given its development as a research technique, that concordancing should have been used first to answer questions about what should be taught rather than how.

4.3.1 CONCORDANCING AND LEXICOGRAPHY

Although it might be argued that lexicography falls outside the present focus, it merits mention in part because it would be difficult to conceive language teaching without dictionaries and in part because it was the lexicographers who provided much of the impetus (and funding) for the construction of the earliest non-literary corpora, especially the Longman's and the Collins' COBUILD project (cf Sinclair 1987), and it was the exploitation of these corpora that encouraged the development of concordancing software. From the perspective of teaching foreign languages, it was not until 1994 that the first bilingual dictionary based on the use of corpus material, *The Oxford Hachette French Dictionary* (Corréard and Grundy 1994), appeared. In the introduction, the authors provide extensive examples of concordance extracts from their corpus as well as a justification for their use:

[N]o human reader can capture all the examples of how even one word is used in a single book, let alone in a hundred or a thousand books. A computer does just that, for every word in all the texts. No human being, even with all the time in the world, can be sure of remembering all the ways of using even one small word, let alone discover how other people use it. A computer allows us to do just that, for every word in the language. No human being can take ten thousand sentences and put them in alphabetical order of one of their words. A computer does that in a few seconds (op.cit.: xxi).

Speed is not the only consideration. Concordancing allows a much wider range of sources to be searched and ensures a much higher level of generalisation and hence validity. The lexicographer's need for data drawn from real examples of native-speaker usage parallels that of anyone engaged in the learning or teaching of a language.

4.3.2 COURSE AND SYLLABUS DESIGN

The primary value of concordancers for course designers is that they allow them to do two essential things: establish the frequency of words in a corpus and provide detailed contexts. The most radical case for basing a syllabus entirely on word frequencies is made by Willis (1990) who used the COBUILD corpus as a source for a complete EFL course (Willis and Willis 1988) and a pedagogic grammar (Willis 1991). On a smaller scale, Flowerdew (1993a, 1993b) used a frequency count and concordancing in relation to a small, specialist corpus (biology lecture notes) in order to identify a key vocabulary to help in course design. Comparison of the top twenty nouns in his corpus with those in the COBUILD corpus convinced him of the necessity of this procedure. Moreover, by providing examples of vocabulary in context, the concordances helped to identify which uses of lexical items to teach, to show the syntactic patterns in which they appear and to throw light on the true functions of certain items, especially discourse makers such as *right*, *now*, and *well* (1993a:238). He concludes:

[I have] shown how computerized text analysis can assist in the design and evaluation of an ESP course. The only necessary conditions for such a procedure are possession of the relatively simple frequency and concordance programs, and a corpus of language drawn from the specific area of the targeted students. Armed with these, course designers can ensure that the language they present in their courses corresponds as closely as possible to the language that is actually required by the learners in their specific purpose area (op.cit.:241).

Polezzi (1993, 1994) describes in some detail the use of concordancing in the design of a university *ab initio* course in Italian for special purposes. The corpus used was based on a series of lectures given as part of the course. The software, through a combination of frequency counts and keyword concordances, provided a simplified dictionary, various partial word lists and a corpus-specific lexical grammar. Moreover, the methodology allowed the author to go beyond the purely lexical and grammatical and helped determine the syllabus: "One of the most significant results of this approach is that we are able to stress discourse features encountered in the corpus and stimulate textual strategies in the learners much more and much earlier than usual [...] - introducing work on link words

and phrases, for instance, or on anaphora and other mechanisms of textual cohesion” (1994:18).

Mindt (1996, 1997) uses concordance style frequency counts and text samples to evaluate syllabus and course material design in the teaching of English in Germany in order “to bring textbooks for teaching English as a foreign language into closer correspondence with actual English”(1996:247). Among the areas he examines are modal verbs and the formation of the future and conditional clauses. Johns (1997) focuses on a specific aspect of course design, the development of remedial grammar exercises. Because one of his aims is “to exclude invented examples of English as far as possible”(op.cit.:100), the concordancer is essential in the recovery of items from his corpus which are then edited and saved in a format that can be used as input for the teaching/testing activities.

4.3.3 GRAMMAR / COURSE BOOK EVALUATION

Whereas the activities described above exploit the quantitative output of the concordancer, those described below make use of the qualitative information obtained from detailed analysis of the concordance output. One outcome is the open challenging of the adequacy of published materials. Stevens (1993), after describing an activity based on *if* constructions in English, goes on to suggest: “As a final exercise, students could be asked to compare their findings with the formulas for conditional sentences often taught in grammar books. The discrepancy is sometimes so large as to render the further use of the grammar books potentially embarrassing”(op.cit.:11: cf also O’Keefe and Farr 2003:390). A somewhat different and possibly sinister example is to be found in Louw (1989). He describes how he compared the contents of the “writings of Zimbabwean language planners and educators with a corpus of English language materials in Zimbabwe”(op.cit.:30). By identifying and carrying out searches on key words such as *Zimbabwe* and *socialism* in order to measure frequencies as well as examine collocations, it was possible to assess manuals and other texts not just in terms of their pedagogic effectiveness but also in terms of their compatibility with the values and goals of the new state

Dodd (1997) describes use of concordancers in the teaching of a foreign language, in this instance German, “where corpus data can be used with the twin aims of improving language competence and sensitizing students to issues in the kind of language description known to them from standard reference works”(op.cit.:135). He identifies two approaches. The first requires students to move from textbook to data by looking at the way a grammatical feature is described before looking for examples. The second, which he sees as being more demanding and which fits better the DDL philosophy, requires students to start with the corpus data before proceeding to formulate their grammatical descriptions. Finally these are checked against those contained in the published grammars. Jones (1997), also working on German, is more interested in the construction of an oral corpus and the differences in word frequencies between written and oral corpora. He provides a number of examples of lexical and grammatical features of modern spoken German, using authentic data to highlight differences between grammatical descriptions and real usage.

4.3.4 TESTING

The use of concordances as sources for gap-filling exercises which can be used in class for testing purposes is discussed by Stevens (1988 and 1991b) in the context of vocabulary reinforcement. He starts from the statement that students generally score badly on gap-filler exercises, either because of the lack of authenticity of examples or because of the domino effect caused when one word from a list is put in the wrong place (1991b:47). In his view, concordances can provide authentic examples with sufficient context to allow identification of the right word, especially if there are multiple contexts for each which reduce the likelihood of one misplaced word leading to others being misplaced (op.cit.:48). He also suggests the use of textbooks as a source for the concordances (ibid). He addresses certain criticisms of concordance extracts that they are unsuitable for this type of activity because of their truncated nature:

I contend that this is actually an advantage lending itself to fruitful exploitation in subsequent class periods, because students can extrapolate the information missing at either end of the concordance output, and the more they extrapolate, the more they can recall and discuss the original context (ibid).

This ability “to extrapolate holistically from fragmentary evidence” is, in Stevens’ eyes, “a vital skill sadly lacking, even at highest levels, in students passing through many educational systems”(ibid), one that is necessary if students are to cope with real data and one that concordances can help develop.

Stevens goes on to describe an experiment designed to challenge those who “doubt that students, left to their own devices, are able to make the mental leaps required to learn for themselves a system as complex as that of a living language”(ibid). The experiment required students to sit parallel tests, one a teacher-constructed gap-filler and the other a concordance-derived gap-filler. The main difference was that the concordance sentences were taken from the students’ textbook and thus perceived as authentic. Stevens admits that the results of the first set of tests were not as conclusive as hoped in spite of a higher level of performance in the concordance-based tests. However, in the second set of tests, there were significantly higher scores in the concordance based tests which were felt to be reasonably conclusive.

Relatively little has been written on the use of concordances for formal testing. Butler (1991) describes how the inclusion of concordance generated cloze tests was proposed for the new University of Cambridge Local Examinations Syndicate Certificate in Advanced English examination but then abandoned. He rejects machine produced cloze tests which combine “the worst aspects of a computer (stupidity) [...] with the worst aspects of the problems of test designers and/or teachers: a lack of time and hence a desire for short cuts with consequent low reliability and validity”(op.cit.:33). Instead he proposes a modified cloze test based on concordance samples. The selection of four examples of the same word combines authenticity with a minimum of ambiguity and leads, in his view, to a more reliable test (ibid). Alderson (1996) examines in some detail the feasibility of using computers to generate and mark tests and suggests the use of concordances as a source of language samples for cloze tests (op.cit.:254) or as a source of material for stylistic and/or linguistic responses (op.cit.:255). His main interest, not surprising given his involvement in the Lancaster University Computer-Based Assessment System (LUCAS), is in developing the software rather than in the pedagogic benefits.

While Wilson's (1997) use of concordancers is to produce exercises rather than tests, an exercise used to assess students' performance is *de facto* a test and so the method could be used for testing, especially if combined with a program such as *Hot Potatoes*¹⁵ which allows for automatic scoring and submission of results to the teacher. Wilson warns however against some of the shortcomings of concordancing as a means for finding examples of test items. The concordancer effectively rips the item, with or without the full sentence, from its context and so may well lose information, such as the referents of pronouns, that are essential for full comprehension. What might be an advantage (see Stevens' views above) as a trigger to discussion in an ordinary class is clearly a disadvantage in a test procedure. Another danger is that the sentences provided are simply too complex for the purpose. This is more problematic in Wilson's case because her aim was to automate the process. However, when using the concordancer in a non-automated manner, it should be possible with the program to recover more of the context and replace pronouns by their noun referents as well as globally simplify the sentence. This may make the sentences less authentic but few foreign language test types could justify the label authentic. At the very least, such sentences are palpably more authentic than invented examples. On this points, her views echo Johns' and Stevens'.

4.4 TEACHING WITH CONCORDANCERS

Before concordancers could be considered as a viable classroom teaching tool, two developments were necessary. Firstly the hardware had to be both small and powerful enough to become easily available not just to teachers but more importantly to students and this did not really come about until computer rooms or laboratories became current in teaching institutions and CALL had become established (see Chapter 3). Secondly, and more critically, it required a sufficiently large mass of teachers to have acquired experience of concordancing as part of their private research or course preparation before its potential as a classroom resource could be perceived (cf Leech 1997).

15 This is a suite of authoring tools available for download from: <http://web.uvic.ca/hrd/hotpot/> (March 2003) which enables, among other things, the construction of gap fill tests.

4.4.1 PEDAGOGIC PRINCIPLES

The hardware, software and corpus together create a potential but the real challenge for the teacher lies in the uses that he or she makes of the concordancer. Since the late 1980s, there have been many articles explicitly referring to use of the concordancer (Bernardini 2000; Butler 1990, 1991; Ching and Wong 1994; Cobb 1997; Flowerdew 1993a, 1993b; Garton 1996; Hanson-Smith 1993; Honeyfield 1989; Hoey 1997; Jackson 1997; Kettemann 1995a, 1995b, 1996, 1997; Lamy and Mortensen (2000); Levy 1990; Owen 1996; Pearson 1996; Polezzi 1993, 1994; Qiao 1995; Qiao and Sussex 1996; Rézeau 1997, 2001; Somogyi 1996; St.John 2001; Stevens 1988, 1993, 1995; Turnbull and Burston 1998; Whistle 1999a, 1999b; Wichmann 1995; Willis 1998; Wong *et al* 1992; Wu 1992) while in others, mainly referring to the use of corpora, use of the concordancer is central to the exploitation of the material (Aston 1995, 1997a, 1997b; Barlow 1995; Bernardini 1997; Biber *et al* 1994; Bowker 1999; Coffey 2002; Fligelstone 1993; Fox 1998; Friedbichler and Friedbichler 1997; Gavioli and Zanettin 1997; Granger 1994; Kennedy and Miceli 2001; Knowles G 1990; McEnery and Wilson 1997a, 1997b, 2000; Murison-Bowie 1996; Rundell 1996a, 1996b; Rundell and Stock 1992; Tribble 1997; Zanettin 1998).

Among the early pioneers in the UK were John Higgins, Timothy Johns, Chris Tribble and Glyn Jones. Tribble and Jones (1990), as the subtitle of their book (*a resource book for teachers*) suggests, set out to persuade teachers by listing and describing a number of activities that appear to them to be highly productive and these will be discussed below in the relevant sections. The importance of their work lies in the range of activities described since they devote relatively little space to discussion of the pedagogical implications and have been accused of being over-optimistic about the level of autonomy that students might achieve (Lamy 1999:38). Most authors give examples of possible activities but, more important as far as the history of concordancing in the classroom is concerned, is the detailed and coherent pedagogic framework for such activities which Higgins and Johns (1984) provide (cf also Johns 1991a, 1991b).

It is this pedagogic underpinning of concordancing that makes Higgins and Johns' (1984) contribution significant and justifies substantial quotation. They

were the first to promote the use of the concordancer in language learning, identifying as its strengths its ability to:

1. produce a list of all the words that occur in a text or corpus of texts
2. establish the frequency of each word
3. establish the range of frequency of words across different texts and text types
4. calculate the relative coverage of different subsets of words
5. identify the collocations within which words are most frequently found
6. isolate the context in which a word occurs (op.cit.:88-89).

They clearly relate these uses to the practice of teaching, based upon “data-based research” which they see as able to “help teachers to gain new perceptions and insights[...]. The computer gives practising teachers the opportunity to do simple on-line research into how the language actually operates in the real world”(op.cit.:90). The first example they give is that of a concordance based on the word *absolutely* and its collocates as part of an investigation into the use of adjective intensifiers in English.

The results confirmed certain of their hypotheses about the type of adjective used with *absolutely* - essentially ‘polar’ adjectives such as *vital*, *certain* and *right* - but challenged the supposition that these would normally be negative rather than positive in value. One question that the data could not answer was whether such uses were essentially British or typical of a specific social class or group. Any data raises questions and these lead the authors to speculate: “One of the most powerful assets of the concordance might be its suggestiveness and its way of opening up questions that might not otherwise be asked at all”(op.cit.:93). Comparisons of Higgins and Johns’ data with data from the BNC Sampler (Table 1 below) shows that *absolutely* is far more common in speech than writing. Because Higgins and Johns do not provide a complete listing, it is not possible to make a full comparison but certain features do stand out: the preponderance of positive adjectives; the consistent appearance of *right*, *sure*, *clear* and, to a lesser extent, *certain*, *essential*, *marvellous*, *true* and *no*. The written examples not only show a much lower frequency but a total lack of any regular collocates, suggesting by implication that the spoken language tends to have recourse to fixed expressions more frequently than the written language.

Higgins & Johns oral words = 1,500,000 n = ?	BNC spoken words = 1,000,000 n = 129	BNC written words = 1,000,000 n = 20
right 11 marvellous 7 sure 7 vital 7 certain 5 clear 4 wonderful 4 brilliant 3 crucial 2 necessary 2 specific 2 splendid 2 true 2	right 7 sure 5 clear 3 essential 3 marvellous 3 massive 2 ¹⁶ beautiful 2 true 2	certain 1 clear 1 confident 1 cruising 1 delighted 1 essential 1 important 1 ravishing 1 right 1 stable 1 sumptuous 1 sure 1
no ?	no 7 awful 3 useless 2	fatal 1 frightful 1

Table 1: Adjectives collocating with *absolutely*

As well as identifying potential applications of the concordancer, Higgins and Johns situate its use within a coherent framework of pedagogical theory. Borrowing their ideas from research published by the University of East Anglia (Kemmis *et al* 1997; discussed in greater detail in Section 3.2 above), they identify a number of paradigms of which the most important in this context are the second or **revelatory** which “is applied to learning which concentrates on discovery and vicarious experience through stimulation and data-handling” and the third or **conjectural** paradigm which “emphasises active knowledge, manipulation and hypothesis testing”(op.cit.:18-19).

Johns (1991a,1991b) develops these ideas further and situates the use of the concordancer within the framework of Data-driven Learning (DDL) in which “we simply provide the evidence needed to answer the learner’s questions, and rely on the learner’s intelligence to find answers”(1991a:2). The concept of Data-driven Learning has since become one of the most quoted and arguably most potent concepts in concordance based CALL (Farrington 1996; Garton 1996; Polezzi 1993; Leech 1997; Turnbull and Burston 1998; Bernardini 2000; Rézeau 2001; St.John 2001; Hadley n.d. and 2002; Beatty 2003).

16 *massive* actually occurs four times but three of these are repetitions within the same sentence

What emerges from the literature is above all a reappraisal of the roles of the learner and teacher. Frequently the emphasis is on exploratory learning with the learner represented as a researcher acquiring a large measure of independence over his or her learning : “The role of the learners changes significantly ; in acquiring the ability to explore the foreign language and to evaluate their own learning processes, they will become autonomous language learners who are not dependent any more on their teacher and [...] their teacher’s personal constructs”(Wolff 1984:141). Barlow (1996) also sees students as becoming largely independent language researchers. The advantages of such a change for McEnery and Wilson (1997b) are that students are allowed to discover things for themselves, they can take divergent paths and are required to interact with the linguistic data (op.cit.:6:cf also Leech 1997). In this case the teacher is not as redundant as Wolff above seems to suggest: “[L]earning from a corpus is directed by the teacher, but not led. The teacher may point the general direction for the students, that is, tell them what points to examine and what questions to ask. But it is the students who lead themselves and one another through the learning process”(McEnery and Wilson op.cit.:loc.cit.).

4.4.2 LITERATURE

The earliest uses of concordancers were in the field of literature but the hardware and software limited these to research in specialist departments for the production of word lists or the stylistic analysis of vocabulary (Crosland 1975; Koubourlis 1975; Akehurst 1981; Burton 1981). Once, however, both software and machines became more widely available, it was not long before concordancers were being used in the teaching of literature. In many cases literary analysis is just one among a number of activities referred to (Davidson 1990; Jackson 1990; Kettemann 1995a, 1995b, 1997; Last 1984). Tribble (1990) identifies the following applications in the study of literature: character studies, recurrent images in fiction, rhyme patterns in poetry, narrative development in short stories (op.cit.:474). These ideas are developed further in a collaborative publication (Tribble and Jones 1990) already mentioned which, in its final chapter, demonstrates how concordancers, through a combination of frequency

counts and concordances focusing on key words, can give insights into text structure and character description¹⁷.

Louw (1991) argues the case for using concordances in the teaching of literature, both to throw light on the literary process itself but also, indirectly, to develop students' understanding of a foreign language and especially its lexis. One aspect which he considers to be particularly problematic is "progressive delexicalisation". He uses a concordance based on the verb *take* to demonstrate how the majority of occurrences found in a set text do not conform to the lexical meaning of the word as given in the textbooks or dictionaries available to learners. By focusing on these delexicalised meanings, the teacher can help raise students' linguistic awareness as part of the process of understanding the text. His application of the concept to the figurative uses of *ripple* or the personification of *Earth* is perhaps open to challenge since both can only be understood in relation to the root or "full" meaning of the word and as such are not really delexicalised. He also sets out to demonstrate how concordances, linked to frequency counts, can support insights into the meanings of texts through highlighting collocations but weakens his case by not always making these insights sufficiently explicit. The main point that Louw succeeds in making is the ease with which concordances and counts can be performed and hence their flexibility as tools in the teaching of literature. In a later contribution (Louw 1997), he demonstrates how concordance based evidence can be used to verify intuition or, in the case of L2 speakers, compensate for lack of native-speaker intuition. Many of his examples taken from poetry bear upon irony and delexicalisation or metaphor. The corpus enables students to compare central and figurative meanings in order to gauge effect. An important feature of this is the use of words divorced from their usual collocations, giving to what would otherwise be clichés stylistic significance.

Kowitz and Carroll (1991) concentrate on integrating concordancing into the study of literary texts. The teacher first has to make a preliminary reading/scan of the texts to identify key elements and words. These then serve as the basis for a set of concordances which, when edited to remove irrelevant or misleading

17 Other suggestions for concordances performed on literary texts can be found on-line at Catherine Ball's Tutorial webpage: <http://www.georgetown.edu/faculty/ballc/corpora/tutorial.html>.

examples, are given to the students. The focus can be on the multiple meanings of key words (*blood* in *Julius Caesar*) or on character (collating adverbs with the names of the characters in *The Great Gatsby*). The authors conclude by itemising those elements which they see as encouraging individual participation: tasks are individualised; in addition they are problem-centred and vary in size and difficulty; work is checked as part of the class activity rather than by the teacher outside the class; students learn from each other and have to assume responsibility (op.cit.:149). Inkster (1997) describes both the construction of a French literary corpus and the production of concordance extracts as part of a student project on psychological terms in *Manon Lescaut*.

In most of the examples quoted, concordances are used in class to test intuitions about a text's content or structure that are formulated in advance by the teacher. Jackson (1997) provides a detailed description not only of the type of projects that students can undertake but also of the types of software and investigative procedures involved, among which are concordances. The difference is that the students both choose the texts that they wish to study and formulate the questions they wish to investigate. Jackson describes the process in great detail but summarises the outcomes rather briefly. However, examples he gives are comparative frequencies of personal pronouns to identify active and passive participants in narratives, the frequency of verbs of speech or adverbs of manner in novels and, more subtle perhaps, searching for patterns of alliteration in a selection of Elizabethan sonnets.

Analysis of the comparative frequencies of selected words can indeed throw light both on the style and content of texts, provided that the benchmark frequency is valid. A frequency count of *Macbeth* will produce high frequencies for 'witch', 'fear', 'sleep', 'nature' and 'blood'. But, in order to assess the significance of these results, what should they be compared to? A contemporary or a modern corpus? Should this comprise written or spoken texts, or simply plays? How significant would the comparisons be if the meaning of the words had changed radically? 'Fear' is much more frequent these days in its weakened meanings (*I fear [that], without fear of*) as is, to a much greater extent, 'witch' (*the wicked witch*). More valid would be a comparison (1) against the whole Shakespeare canon, to see to what extent *Macbeth* differs from the other plays,

(2) against the whole body of contemporary drama, and (3) against a control corpus of all known Elizabethan texts (the wider issue of comparability has already been discussed in Section 4.2.3). Akehurst (1981) found, when he compared southern and northern French mediaeval poetry, a significantly higher frequency of *dever* in the northern sample, suggesting a greater preoccupation with duty. Comparison of literal and metaphorical uses of key words can also be revealing. In many of the cases quoted, in contrast to the Akehurst example where the researcher discovered something new, the concordancer simply confirmed the researcher's intuition. Some of the literary examples cited, such as the instances of delexicalisation given by Louw (1991), are more questionable, especially the case of the verb 'take'. However significant delexicalisation may be as a feature of English and however useful it might be for students to be aware of it, it is difficult to see just how this throws light on the specific text or how literature is different from non-literature in this respect.

Although the concordancer seems more adapted to confirm (or refute) the researcher's intuitions, there is always the chance, as the Akehurst example shows, that examination of frequency lists will reveal the unexpected. It is quite simple to routinely conduct searches on the main protagonists in novels (for example Adolphe and Ellénore in *Adolphe* or quite simply Emma in *Madame Bovary*) to discover what verbs or adjectives are associated with them. Ellénore in *Adolphe* is more often than not the object or beneficiary of verb constructions and, when she is the subject, it is frequently of verbs like *être* and *paraître*. These facts are easily established with a concordancer and reflect Ellénore's role first as object of Adolphe's gaze and consequently as object of his passion and then his neglect. Whether a concordancer could lead to the insights that Halliday (1973) derives from a comparison of the verbal constructions used by Neanderthal man and *homo sapiens* in Golding's *The Inheritors*, is open to question. The absence in the novel of transitive constructions in Neanderthal man's speech is interpreted here as his inability to see the relationship between cause and effect and, consequently, his inability to manipulate and control the physical world, hence the ultimate triumph of *homo sapiens*. Certainly an ordinary concordancer would speed up the search enormously, once the relevant verbs had been identified, and exemplify Kemmis' emancipatory paradigm. The

availability of a tagged text would allow comparative searches to be performed on transitive and intransitive verbs and, possibly, on verb + noun object constructions. However, the original intuition could only come from a reading of the text.

4.4.3 LEXIS

One aspect of lexis is the development of vocabulary in the traditional sense, that is the acquisition of what might be termed “full lexical items” in that their meaning belongs to the world of experience and is usually represented by a noun, verb or adjective. These are the words that are most easily looked up in a dictionary. However these are the words that are most topic-specific and hence likely to be excluded from general language coursebooks and dictionaries. An early and detailed description of the use of concordances in vocabulary development in a specialist field is to be found in Baten *et al* (1988). The pedagogic context was the teaching of business English to second-year university students of economics. The aim was to make students “explicitly [pay] attention to how concepts are [...] verbalized in their own field of study so as to be able to gradually internalize and use the specific wording in their daily practice”(op.cit.:453). The method was to identify a number of key words and give students concordances based on these. Students then had to select their own keywords from a small number of relevant texts and consult the concordances in the library, copying relevant examples which then served as the basis for productive work. One initial problem was students’ lack of familiarity with concordances and the way they present material. At the end of the year, the authors measured students’ performance in written and oral modes to assess acquisition of vocabulary and especially of collocations. Although there was some variation between modes and above all between students, the authors were positive about the experiment but felt that, in future, contexts should be longer and each student should have copies of the concordances or the ability to print them out from a PC in the library.

A more “hands-on” approach is described by Ilse (1991) who used Micro-Concord in EFL classes with students of banking in Germany. Students were already using networked computers to familiarise themselves with new banking

systems and the concordance work was designed to complement this but with the emphasis on specialist vocabulary acquisition. The basic procedure was for students to perform concordances on a selection of banking texts looking for all the uses of the keywords *share**, *debenture**, *bond**, *unit**, *broker**, *jobber**, and *securit** which were allocated according to group. Students saved the results and then had to load them into a wordprocessor and then, with their partners or neighbours, check meanings, weed out the irrelevant examples and complete fragmented sentences. Finally, they had to report their findings back to the other groups. Although the students may have left the class with less factual information and may have used more German in class than might have been the case with a teacher-led session, Ilse felt that the procedure employed helped them acquire useful skills in addition to the pedagogic advantages of text as opposed to content-focused work. She concludes: "The motivation is undoubtedly there: my students found the concordance program a fascinating piece of software and appreciated its potential for investigating and extracting information whether on facts and figures or linguistic questions"(op.cit.:107).

Another example of specialist vocabulary acquisition is to be found in Mparutsa *et al* (1991) who report on their use of the concordancer with first-year university students in Zimbabwe and who see the acquisition of specialist vocabulary as central to improving inadequate reading skills in the first year. "In Economics, [the methodology] was used to support lexical familiarisation and to sensitise students to an awareness of the developing use of terminology"(op.cit.:131). Wu (1992) demonstrates a broader approach, exploiting the concordancer to extract a wider sample of recurrent features in authentic engineering texts. These can be verbs (*consider*, *describe*) used as discourse signals, tenses (present, present perfect, future) used to describe problems and solutions, syntactic structures (especially nominalisation) and collocations/multi-word units such as: *as described in*, *as follows*, *carry out a test*, *achieve an objective*, *adopt an approach*, *problems occur*. The purpose was to help learners recognise regular features of engineering texts and thus speed up their learning.

Similarly, Johns (1991a) uses concordances as a basis for working out with the students a rule or set of rules to contrast and explain usage of less specialist verbs like *convince* and *persuade*. St.John (2001) focuses specifically on German

modal particles. Flowerdew (1993a, 1993b), Goodfellow (1994) and Gavioli (1997) all use concordance searches in order to facilitate vocabulary learning and lexical information-collecting.

What differentiates Cobb (1997) from most of the others is that, like Stevens (1991b), he tests the efficacy of using concordances to improve lexical acquisition under experimental conditions using two versions of purpose-built software, one using concordances to contextualize lexical items and the other example sentences and descriptions. The group exposed to concordances acquired “12% more transferable word knowledge”(op.cit.:312). This leads Cobb to conclude that concordancing can compensate for the fact that students do not have enough time “for rich, natural, multi-contextual lexical acquisition to take place”(p.314) in the way that it does in their first language. He thus sees concordancing as having the potential to “mimic the effects of natural contextual learning”(ibid) in a more concentrated and efficient manner. The case for the linking of concordancing to reading is further made in Cobb *et al* (2001), Gavioli (1997) and Qiao (1995).

Underlying all these attempts to teach vocabulary via concordancing is the assumption that L2 vocabulary is best learnt in context reflecting the way that L1 vocabulary is acquired. This is certainly Cobb's (1997) view. Groot (2000) suggests however that this may not be true of all learners and that some may perform better from working with bilingual lists but he also suggests that concordancing could be a valuable follow-up activity. One point that Groot does not address is the inadequacy of many dictionary definitions and the need to test them against real use, a question that recurs frequently in the literature on vocabulary acquisition (cf Wichmann 1995; Pearson 1996).

A slightly different perspective is given by Knowles G (1997) who uses concordances in a diachronic study of English to tackle a variety of problems: changes of meaning of individual words; distribution of *-eth* and *-s* verb endings; comparison of degrees of formality and origin of words (mainly Germanic, French and Latin). In many respects, the author is more interested in the problems that arise from a student centred approach, partly as a result of lack of various types of background knowledge (metalinguistic, historical), but does

highlight the computer's ability to provide breadth of coverage. Without the concordancer's ability efficiently and effortlessly to search the source texts, it is questionable whether the task would have been undertaken.

4.4.4 GRAMMAR

This is the area of use that is the most directly relevant to the present study. Many users of concordancers saw them and the learning modes they enabled as a means to radically reform the teaching of grammar. Among these was Johns (1991a) whose theory of Data Driven Learning has previously been discussed in Section 4.4.1 in the context of a discovery-based approach to language learning which gives pride of place to the concordancer. Johns also advocates a radical re-evaluation of the place of grammar in language-learning and language-teaching.

Traditional grammar-based methods are vitiated by assumptions about *how* grammar is learned, and *what* is to be learned. The *how* usually involves presenting the student with a set of "rules" or "patterns" that are then applied in "constructing" text in the foreign language (op.cit.:3).

The reasons why this, in John's view, fails are that existing grammatical descriptions are neither full nor reliable (ibid). Large areas are often omitted and what is presented is often based on the "armchair intuitions" of grammarians and can be misleading. One consequence of this state of affairs is a reaction against grammar teaching. For Johns teaching is to be replaced by consciousness-raising:

The DDL approach [...] makes possible a new style of 'grammatical consciousness-raising' [...] by placing the learner's own discovery of grammar at the centre of language learning, and by making it possible for that discovery to be based on evidence from authentic language use (ibid).

This can lead to a much deeper understanding on the part of the student: "when grammatical description is the product of the learner's own engagement with the evidence, that description may show a far greater degree of abstraction and subtlety [...] than would normally be allowed for in the type of pedagogic description that is presented as 'given'"(op.cit.:3-4). He does however warn against the dangers of "over-generalising on the basis of negative evidence"(ibid) which can happen when the concordancer only provides examples of the most frequent uses of a word.

Johns describes his students as having a preference for grammar teaching. "Most of them come from an educational background in which language learning is equated with formal grammar learning, and against such a background problems with English are likely to be identified as 'not knowing enough grammar'"(1991b:28). The request for grammar is also seen as a response to the limitations of acquisition in context (in Krashen's sense of the term): "acquisition tends to be haphazard and leave them feeling confused"(op.cit.:29). Use of concordances, therefore, enables Johns to cater for students' expressed needs and play upon their strengths: "They are intelligent, and respond to challenges to their intelligence: and most of them are accustomed to the idea of research and finding things out for themselves"(ibid).

In a slightly later treatment of the subject, Stevens (1995:3) adopts a similar line, quoting McDonough in support:

First, it is unlikely that the students can understand the rule statement until they have tested it against the various examples [...]. Secondly, [giving a rule first] imposes a rule formulation rather than encouraging the student to make one up in his own terms. In cases where the discrimination necessary is relatively simple, an imposed classification is usually less easy to remember and therefore less efficient than one invented for oneself (McDonough 1986:31).

In another quotation, Stevens (1995:4) describes the kind of insights that become possible as linguists "find explanations that fit the evidence rather than adjusting the evidence to fit a pre-set explanation"(Sinclair 1986:202). In his contribution, he provides examples of others' activities to gain linguistic insights and discusses role of teacher as helper, especially with students in early stages, referring to his own work with *if... then* constructions.

Wong *et al* (1992) provide a thoughtful and detailed description of use of the concordancer in the EFL classroom. What makes their contribution particularly useful as a model in the present context and justifies extensive quotation is that they set out to test its effectiveness, "focusing on (1) whether the use of this method of teaching/learning helped [students] better grasp some of the 'rules' of grammar, and (2) their reaction to the use of concordance data to help them discover for themselves the patterns of English which they had not been sure of"(op.cit.:41).

The basic methodology was to use MicroConcord for KWIC searches, the advantages being that the keyword is obvious, the disadvantage being truncation of the context. The authors' preference was for students to perform the searches themselves but, because of class sizes, students were provided with print-outs. One advantage of these was that these could be edited to make them suitable for elementary classes. However, "[t]he ultimate aim, of course, is to have [the students] graduate to the point where they can use the concordancer for themselves, for self-accessed and self-paced learning"(op.cit.:43). The students were consequently familiarised with concordances and then split into small groups. The authors found that the students particularly appreciated the involvement of the instructor, not as someone who knew the answers but as a guide and helper, and conclude: "The discovery process being a shared one seems to be one of the strengths of an approach using concordancing in the classroom"(op.cit.:44). In this they echo and would seem to confirm McEnery and Wilson's (1997b) belief in the necessary role of the teacher (cf 4.3.1 above).

The main focus of the article is the authors' experimental study concentrating on two weak areas of grammar and two types of exercise, the first based on the inflectional forms of the main verb in various types of verb phrase and the second based on the use of the article (definite, indefinite, zero) with specific noun heads. The first was chosen because it was felt to be "a grammatical item that was quite systematic and whose behaviour could be easily discerned from the data given"(op.cit.:44), the second because it was "a grammatical item that has caused many problems to learners of English and foiled many attempts by even the most skilful teachers to teach it"(ibid).

The method adopted was to split students into three experimental groups: the first worked on inflectional verb forms, the second on articles and the third worked on both. When it came to evaluating the experiment, the authors found that the majority (69% with 18.8% neutral) found the approach helpful, however this varied between groups: group 1: 78%, group 2: 54%, group 3: 75%. As far as the first exercise was concerned, responses were positive. 100% of students liked it and had no problems with the truncated sentences. They actually found the exercise too simple but that may have been because they had previously worked on the grammatical point. There was a marked contrast in the responses

to the second exercise. Here the majority of students disliked the truncated sentences and felt that complete sentences were necessary for an understanding of articles. Not surprisingly, the authors report a marked improvement between pre-test and post-test scores in the case of the first exercise but only a marginal improvement in the case of the second.

The authors found that the concordance-based approach was “very effective when shorter contexts (eg., verb phrases, prepositional phrases, and active passive voice) are sufficient for the discovery of the ‘rules’ which govern the use of a particular word or morpheme” but less so for grammatical items like articles, tense and modals “which require far more context and are more meaning based”(op.cit.:46). Globally, however, their message is positive:

[O]ur experience with using this approach [...] has made us firm believers that there are many other possibilities which bear further experimentation. We therefore urge language teachers to try out the use of concordancing for various other purposes and in various other roles in their own contexts. We believe that the approach is a beneficial one, not only from the language teacher’s point of view but, perhaps more importantly, from the students’ too (op.cit.:46).

In another contribution from the Far-East, Ching & Wong (1994) describe a concordance based project to teach English verb inflections to Chinese-speaking students. Concordances are first prepared from two corpora, one composed of engineering texts and the other a collection of students’ work in which the errors had been tagged. Searches on the former are made according to keywords such as *has*, *have*, *will* and *would* and on the latter using the tags. Students have a few minutes to study the concordances and then have to perform activities such as gap-filling or correcting errors. The examples provided are an exercise with gaps designed to work on infinitive/past participle distinction and another in which they had a printout of their own work on which they had to ring errors and insert the correct form. The students’ response was positive (op.cit.:53) and the authors identified fewer errors (op.cit.:54).

One of the few experiments where the students themselves formulated the questions is described by Qiao and Sussex (1996) who were not only working with advanced students but were also using tagged corpora in order to base concordancing upon grammatical constructions and not lexis. Their examples ae

taken from EFL: the pre-qualifier *such* (“ABL”), to + infinitive clause (“Ti”), the + adj + noun (“*_ATI*_JJ*_NN”). Tagged corpora allow frequency counts of different constructions and comparisons across texts and searches for grammatical patterns (to + gerund, look + prepositions, infinitive without to) without the interferences that one gets in non-tagged corpora. They describe this method as “open-ended, student-centred and heuristic” (op.cit.:55) and in their approach encourage students to generate hypotheses first and then find answers in corpus. Their justification is similar to that of others working with corpora: “This approach can help students think, compare, judge and learn in a more autonomous way than in traditional teacher-centred, theory-oriented grammar learning”(ibid). They claim that the available syntagmatic information in the tagged corpus is an advantage for the learner while recognising that such corpora can be difficult for students to read.

Not all activities need be so complicated. Tribble (1990) describes a simple searching on the suffix “*ly” designed to examine both the formation and position of adverbs. He concludes with the statement quoted on p 51 in 3.4: “Because it is relatively simple to use, does not require authoring before it can be deployed, and has an exceptionally wide range of applications, the concordancer will perhaps be the pre-eminent software tool in this next stage in the development of computer-assisted language learning”(op.cit.:475).

4.4.5 TRANSLATION

Relatively little was written on the application of concordances to the teaching of translation, at least until the mid 1990s. This stands in stark contrast to the interest in machine translation (Nirenburg 1987)¹⁸ and may in part have been because translation became unfashionable as a teaching activity as communicative methodologies came to dominate (see Section 2.2.4 above). Another factor may well have been the practical difficulty of finding comparable texts in more than one language, especially in areas other than literature, although the expansion of the Internet has since made this task much easier. Even though parallel corpora have been constructed and successfully exploited to highlight stylistic differences between languages such as existential sentences

18 For details of the European Community’s EUROTRA project see Nirenburg 1987:pp.114-135.

and passives in French and English (Guillemin-Flescher 1996), there remain the problems of selecting and aligning the content (McEnery and Wilson 1996:57-59; Kenny 1997; Piotrowska 1997; Ulrych 1997; Peters and Picchi 1998; Bowker 2002:55-64) and creating software capable of producing parallel concordances (King and Woolls 1996; Ulrych 1997). As a result, much of the activity in the field has focused upon the use of monolingual corpora or non-aligned bilingual comparable corpora (for a fuller discussion of the differences see 4.2.3 above).

Wichmann's (1995) contribution is particularly useful because it highlights the problems that learners have in translating. She begins by making the point that "concordancers are not widely used for the teaching of other languages in this country" and suggests that "they have the potential to address some of the problems faced by undergraduate learners of foreign languages. These include a lack of awareness of different levels of meaning in language, particularly contextual meaning and speaker meaning, leading to problems of lexical choice and lack of skills in using reference works, particularly dictionaries"(op.cit.:61). Because of the lack of available German corpora, she was obliged to use an English corpus and thus take English as her starting point, focusing on translation difficulties. Part of the problem is that "the dictionaries do not provide enough information on meaning in context to guide students' choice"(ibid). As examples she cites the noun *accident* and the verb *behave* and shows how a dictionary and a grammar can differ in the information they provide with neither being comprehensive. Students were given copies of concordances of eight English words selected from the grammar which cause translation problems and were asked to use whatever reference sources they had available to decide how to translate the examples in the concordances. The outcomes were clearly successful and revealing:

The most striking effect of this approach was the sudden awareness on the part of the students of shades of meaning in their own language. When one explains the need to choose carefully from dictionary entries, one makes the assumption that learners are aware of different referential, connotative and contextual meaning in their own language. This seems not always to be the case, and students claimed that the exercise made them think about English for the first time (op.cit.:63).

She also reports a high level of motivation among her students in spite of difficulties which arose.

The limitations of dictionaries are also recognised by Pearson (1996) and Bowker (1999) who describe the use of concordances in training student translators at City University Dublin. Like Wichmann (1995), both resort to corpus based concordance searches as a means of developing students' linguistic sensitivity. Pearson uses as her corpus commercially available CD-ROMs to help students with specific translation problems where dictionaries are of little help: neologisms, including acronyms; meaning shifts (an example given is *sleaze*) and collocates. Bowker (1999) is more concerned to overcome similar attitudes in her students to those identified by Wichmann which in her view act as obstacles: (1) their resistance to correction in their mother tongue because they feel they are experts; (2) their unwillingness to research their own language; (3) their lack of awareness of the distinction between general and specific varieties of language (op.cit.:161). The main obstacle in her view is students' over-reliance on bilingual dictionaries and their unwillingness to use other sources. After comparing the results of two experimental groups working in parallel, one using a wide range of normal reference materials and the other using nothing but a general bilingual dictionary and a concordancer linked to a specialist corpus, Bowker concludes that use of the concordancer not only improved the accuracy of the translations but also made students more aware of the importance of context and collocation when looking for equivalents (op.cit.:168-170). Varantola (1997), working in Finland, came to similar conclusions.

The examples quoted above are essentially terminological in their focus. In addition they are either based on a monolingual corpus or on two comparable but not parallel corpora in that the target language texts are not translations of the source language texts but are chosen because they belong to the same genre or have the same subject. Friedbichler and Friedbichler (1997:2) declare their preference for working with comparable corpora made up of "representative domain-specific corpora of authentic peer-reviewed texts in the target language" because "they provide reliable evidence from TL experts rather than more or less adequate translations"(ibid; cf also Peters and Picchi 1998:95-96; Zanettin 1998:2). In addition "corpus analysis provides a highly efficient tool for all

language queries that are beyond the scope of dictionaries, particularly with such issues as collocations and contextual usage”(Friedbichler and Friedbichler op.cit.:3). Gavioli and Zanettin (1997) and Zanettin (1998) describe activities with their students based upon parallel word searches on two comparable corpora and designed to highlight differences between English and Italian in areas as disparate as terms designating political leaders, nominal as opposed to verbal constructions (*la caduta dei prezzi* / *prices fell*), adjectival as opposed to noun constructions (*biopsie epatiche* / *liver biopsies*) or conventionally divergent references to the same phenomenon (*il gradino più alto del podio* / *the gold medal*). The method of working is for students to look for comparable texts and to search first for obvious cognates and then to compare the attendant collocations and constructions in the two languages.

While comparable corpora seem to be the preferred instrument of those involved in the professional training of translators, aligned corpora and the associated software seem to have been developed mainly as an aid to L2 language acquisition. Roussel (1991) was one of the first to make the case for Cross-Linguistic Parallel Concordances (CLPCs) as a means to highlight contrasts between languages. Because of the lack in the early 1990s of concordancers capable of carrying out a simultaneous search of parallel texts, her procedure required consecutive searches of two corpora, the second being an *ad hoc* translation of the first. Her aims also required prosodic features to be read onto these texts by native speakers since at that time there were no real oral corpora available, at least not in French or in a form readable by existing concordancers. As a substitute, she identified certain items such as auxiliaries in English which can possess certain intonational features and used the concordancer to search for these. This enabled her to examine with her students how these features are translated into French.

As a result of this experience, Roussel was active in the development of MULTICONCORD (cf King and Woolls 1996), a multilingual parallel concordancer used by Ulrych (1997) with a parallel corpus of Italian texts and their English translations to highlight specific translation problems, for example the translation of the conjunction *ossia* (430-434). Rézeau (2001), using the same concordancer, describes an activity based on *although*, including a printout of

results (op.cit.:pp.164-166). Müller and Groß (1999) compare the performance of two groups translating a German text into English, one working with a bilingual dictionary and the other working on a parallel corpus using MULTICONCORD. They found that the dictionary group performed better but the result was largely explained by the fact that they were more experienced. This group however expressed frustration at the absence of contexts and especially of help with following prepositions (eg *fable about* or *fable on*), something which the concordancer group found especially helpful. The authors nevertheless felt that the concordancer was useful and suggest its use by the teacher to prepare handouts of concordances focusing on difficulties. In fact, the activities described could have been performed equally as well using a monolingual corpus and standard concordancer.

Barlow (1996) uses ParaConc, another multilingual concordance program, together with several parallel corpora, to address translation difficulties mainly in order to sensitise students to similarities and differences between languages rather than to train them as translators. A search for reflexive constructions in an English text and their equivalents in the parallel French text serves to bring out the different patterns (or “schemas”) of usage in the two languages, showing how some correspondences were regular while others were totally irregular. A similar exercise was performed on the words *head* and *tête*. Whereas there was a consistent equivalence between the words when referring to the human body, there was only partial equivalence when the words are used metaphorically and zero equivalence when the English word *head* meant “chief”. He suggests that a search on the word *for* and its equivalents in Spanish would help English speakers towards an understanding of the choices that translation requires. The advantage for Barlow of using concordances is that they provide a much richer selection of examples for analysis than can a dictionary and thus can often highlight uses that would otherwise be ignored. In addition, the combination of the syntagmatic (the horizontal presentation of text)) and the paradigmatic (the vertical presentation of the search term) makes concordances especially conducive to inductive learning.

4.5 CONCLUSION

Considering the variety of applications that have been described above, it may come as a surprise to be reminded by Stevens (1995) that concordancing was only perceived as a useful tool in language learning towards the end of the 1980s. In explanation, the author gives reasons for his and others' use of this tool:

Many teachers [...] feel that concordancers are the type of software that most closely approaches fulfilling the potential of computers in language learning. In a sense they are working approximations of expert systems: they bring students' cognitive and analytic skills to bear on the manipulation of comprehensive databases for the purpose of solving real-language problems (op.cit.:2).

The applications described above reflect the two different types of data that the concordancer can produce. On the one hand there is the essentially quantitative data in the form of information about word frequencies which is used as the basis for dictionary compilation and course and syllabus design as well as in certain aspects of literary investigation. On the other, there is the qualitative, the actual concordances of keywords whose detailed analysis provides the foundation for work in lexico-semantics and grammar. In the earlier contributions, much was often speculative pointing up ways that the concordancer might be used but increasingly the literature has provided examples of actual practice.

It is not surprising that the majority of articles describing the applications of concordancers are written by those who are convinced of their usefulness. Aarts (1991), Owen (1996) and Widdowson (2000) warn against over-reliance on corpus evidence and assert the role of the teacher's intuitions in assessing the evidence and deciding what is or is not correct or what should and should not be taught. Nevertheless, they remain convinced advocates of the value of concordances. Flowerdew (1993b) expresses the following caveats: learners may well become frustrated if there are too few or too many examples; they must be sufficiently trained in the software and methodology; above all teachers must not see the concordancer as a panacea. Pienemann and Jansen's (1992) reservations are more technical in that they question the need for specialist concordancers: "With the arrival of the microcomputer and the wide range of commercially available software, such programs were superseded by standard features of

ready-made software packages [...] Similarly, the function of a concordance program is included in standard database systems, which contain additional features” (op.cit.:208). One could argue that such a solution requires switching from program to program, some of which may require text to be in a special format incompatible with the other programs, as well as familiarisation on the part of the user with several applications of varying complexity. In addition, it would be difficult to maintain the syntagmatic and paradigmatic axes that Barlow (op.cit.) sees as being so valuable.

What also emerges from much of the research cited above is the centrality of the role of the teacher. The general use of specialist corpora may require the teacher to judge the representativeness and applicability of search findings. Students are often ill-prepared to manage the learning process, lacking the experience and linguistic knowledge to adopt appropriate search strategies and formulate useful questions, let alone interpret and use the results. The concordancer must not therefore be seen simply as a means of replacing the teacher and hence reducing staff costs in the name of student autonomy. This will at best be a gradual process. Meanwhile the presence of the teacher remains essential if the concordancer’s full potential is to be realised.

CHAPTER 5

5. THE RESEARCH PROJECT

5.1 THE RESEARCH QUESTION

5.1.1 PEDAGOGICAL ISSUES

The pedagogical aim of this study is to test the effectiveness of concordances as a means to improve students' knowledge of grammar. This was seen as the key to improving their reading and listening skills, which in turn would be the foundation for improved writing and speaking skills. There are those who for a long time have insisted that grammar is something that students best learn for themselves through cognitive processing of language input (Krashen 1981, 1982; Corness *et al.* 1997). A similar view is expressed in Whistle (1986). However, learning a foreign language is radically different from acquiring a first language where the processes are largely inductive and unconscious (cf Hawkins and Towell 1992; Klapper 1998; Mitchell and Miles 1998; Vygotsky 1986:159-161). However different the process may be, the pedagogic question is essentially one of how teachers can create conditions in which students can learn. The question of how grammar and lexis are acquired presupposes a set of learning assumptions. Recent reviews of the research (Chapelle 1997; Wright 1999) suggest that there is now a general consensus that some form of explicit grammar teaching is necessary if learners are to acquire a meaningful level of competence. "In the acquisition versus learning debate, naturalistic theories of acquisition need to be set against research which indicates that focus on form does help the acquisition of linguistic competence" (Wright 1999:38).

Johns (1991a & 1991b) provides a clear rationale for the use of the concordancer in this context.

[T]he language learner is [...], essentially, a research worker whose learning needs to be driven by access to linguistic data [...] The use of the concordancer can have a considerable influence on the process of language learning, stimulating enquiry and speculation on the part of the learner, and helping the learner also to develop the ability to see patterning in the target language and to form generalisations to account for that patterning (1991a:2).

More recently, other authors have emphasised the value of concordancing in grammar teaching (Fligelstone 1993; Flowerdew 1993b; Kettemann 1997; McEnery *et al.* 1997; Hadley 2002). However, most of the published work describing classroom practice focuses on teaching English as a foreign language. There is relatively little on teaching foreign languages (Dodd 1997; Farrington 1996; Polezzi 1993, 1994; Wichmann 1995) and here, as in most of the EFL examples (Tribble 1990; Davidson 1990; Ilse 1991; Wu 1992; Flowerdew 1993a; Goodfellow 1994; Gavioli 1997; Fox 1998; Willis 1998), the emphasis is mainly on lexis (often linked to syllabus design) and semantics rather than grammar.

An important intended outcome of the activity was greater autonomy on the part of the students, to be achieved partly by providing them with a tool which they could access independently and partly by encouraging greater awareness of the language learning process, something seen as essential by Lamy and Goodfellow (1999):

For learners with limited opportunity to interact with other target language users, the promotion of learner autonomy via critical reflection [...] has come to be regarded as of equal importance as, say, the provision of comprehensible input and the opportunity for productive practice (op.cit.:43).

This is all the more important when one considers that most foreign language learning in this country takes place with a relatively limited range of opportunities for exposure to the target language outside the increasingly fewer timetable hours allocated.

5.2 THE EXPERIMENT

5.2.1 THE CORPUS

Before the experiment could start, it was necessary to find or construct a corpus. A teaching corpus is not the same as a research corpus and as such must be viewed “from within an applied rather than a corpus linguistics perspective”(Bernardini 2000:226). Its purpose is to provide normative examples and not to test the limits of the given language. As Owen (1996) remarks, the needs of an L2 learner are very different from those of an L1 learner and can be held to require a corpus that is not simply descriptive but prescriptive:

[I]t is certainly true that most L2 learners fully expect prescription. After all, what is the business of a language teacher if not to enable the learner to speak the language as accurately as possible? This is not an issue of accuracy versus fluency. It is simply that all L2 learners, deep down, believe it to be part of their task to improve their accuracy, whereas most native-speaker learners do not (op.cit.:220).

While this statement may be more true of foreign as opposed to British L2 learners, it did clearly recognise the responsibilities of the teacher.

The corpus did not therefore need to be representative as conceived by Biber (1993a:243): “Representativeness refers to the extent to which a sample includes the full range of variability in a population” (for a more detailed discussion of this point see section 4.2.3 above). It does however have to “have value as a point of normative reference”(Sinclair 1991:17). In addition the norms were to be those of written French since the main objective of the degree course is to improve students’ reading and writing skills (cf Kennedy and Miceli 2001). Exclusive reliance on journalistic texts was excluded because these often contain a variety of stylistic problems that act as barriers to students’ understanding¹. First and second year undergraduates need a linguistic model or models on which they can base their own French confident that they have wide acceptance. Examiners and employers tend after all to be linguistically conservative. The need was therefore for a “didactic language corpus” (Polezzi 1993) which could provide examples of accepted and acceptable use and satisfied Polezzi’s three parameters:

- 1) it must be based on the learners’ needs;
- 2) its size should be determined by the nature of the course and the level of the learners;
- 3) it must be flexible allowing addition and modification (op.cit.:16).

Rather than convert written materials into electronic form, it seemed more sensible and efficient to search for an existing machine-readable source of texts which could be used to constitute an *ad hoc* corpus.

The decision was taken to base the corpus on *Label France* published quarterly by the French Ministère *des Affaires étrangères* and now accessible on the

1 Sinclair (ibid) makes a similar point about English language journalism.

Internet². As the magazine is also published in paper form, it does not suffer from the shortcomings associated with many Internet-specific texts (see Section 4.2.3 above). At the time the experiment began, the resulting corpus contained 232 texts totalling 312000 words. It was decided to treat it as a *monitor* corpus (cf Sinclair 1991) which could be regularly expanded³ to keep track of changes in the language. It was also important to include only whole texts rather than samples in order to preserve the integrity of the source documents and to retain the natural distribution of key discourse features that structure the argument through the length of the text. Although each issue has a theme, the texts are varied in subject, covering cultural, social, political and economic topics, but maintain an essentially French focus, apart from the occasional special issue on subjects such as the 50th anniversary of the Universal Declaration of Human Rights. As a result they provide the breadth of lexis necessary if a corpus is to be of real use. In addition the language is fairly conservative which was felt to be an advantage, the disadvantage being that it is at times syntactically quite complex. This is in part explained by the fact that many of the articles are written by journalists but far fewer than would be the case with a corpus drawn exclusively from newspapers and magazines. The texts also contain a number of interviews which, while remaining within the prescribed formal norms, broaden the range of available linguistic structures. Because *Label France* is aimed at a non-French audience, it does not assume native-speaker cultural and linguistic knowledge on the part of its readers. As a result, references tend to be more explicit. An important additional factor in the choice of corpus materials was the fact that copyright clearance⁴ was readily given by the French Embassy, on the understanding that texts would be used exclusively for pedagogic purposes and the source acknowledged. Finally, in order to speed up searches (see 5.2.2. below), all texts were converted from HTML to text format.

5.2.2 THE CONCORDANCER

When choosing a concordancer for use with students, two sets of parameters must be met. The first set is purely pedagogic. The concordancer must:

- 1) enable meaningful rather than mechanical interaction with the language

2 This can be accessed at <http://www.france.diplomatie.gouv.fr/>

3 The *Label* corpus now contains 696 texts totalling 863584 words (July 2004).

4 For a fuller discussion of this point see section 4.2.3 of Chapter 4.

- 2) present language in context
- 3) present grammatical features both within the sentence and the paragraph to throw light on their communicative function.

To a large extent, these parameters derive from existing, familiar traditions of textual analysis and text-based grammar derivation. The second set is technical.

The concordancer must:

- 1) be fast enough to allow productive searches to take place
- 2) be intuitive enough to be used without unnecessary loss of time
- 3) allow meaningful manipulation and analysis of the results of any search
- 4) allow results to be saved to disk and/or printed
- 5) be robust enough to cope with user errors.

The concordancer must not only be technically adequate but also be sufficiently intuitive for students with little or no experience of concordancing to use.

The choice of the concordancer might have been problematic had there been funding available to buy one. The generally preferred program for research purposes is WordSmith Tools⁵. However students often find it difficult to use (Zanettin 2002:6). The Longman Mini-Concordancer was another possibility. However, both cost money. Micro-concord is available now for free download from its author's website⁶ but has two major disadvantages. The first was its DOS interface which makes it relatively user-unfriendly, at least by modern standards, and the fact that it requires texts to be written using the ASCII (or MS-DOS) as opposed to ANSI (or Windows) code set. This may not matter with texts in English but is of crucial importance when dealing with texts in a foreign language. Reading in texts produced using Windows programs results in accented characters behaving in a bizarre way, as shown in the following example:

C'est Ó partir des annÚes trente que le tourisme devient un phÚnomÞne de masse.

The results are to say the least disconcerting. Pedagogically, instructing students to enter an upper case 'u' acute every time they want a lower case 'e' acute, not to mention the other characters, would be both confusing and potentially demotivating.

5 For details see <http://www.lexically.net/wordsmith/>

6 For details see <http://web.bham.ac.uk/johnstf/timconc.htm>

Practical and financial considerations forced the development for student use of a program which had originally been written for personal research purposes. This was a basic concordancer without the frequency-count facility that more sophisticated programs have and which at that time was not felt to be necessary since students would not be required to perform any statistical analyses. The advantage of this program, baptised KWIC-CONCORD, was that it had been written specially to search for words containing accented characters in ASCII, ANSI and HTML formats. It allowed a search to be performed on single or multiple files, on a whole directory or subdirectory or on a whole disk. It is, in Tribble and Jones' (1990) taxonomy, a streaming concordancer in that it reads texts line-by-line (actually in 31k chunks) and sends concordanced text to screen and, if required, to disk. It also matched Stevens (1995) parameters:

A concordancer intended for use with students should be fast and responsive. It must load quickly, so as not to waste valuable class time [...]. Sorts on output should be instantaneous and permitted both one-over and proximally left and right of the key word. Because second-language learners frequently make mistakes with string specifications, loading and query should be interruptible at any point, and the student able to work with the portion of the corpus or concordance processed up to that point (op.cit.:3-4).

Although streaming can never be as comprehensive as text-indexing, it is faster and on modern machines almost as fast as in-memory concordancing. The main modification made to the program before using it with students was to port it from MS-DOS to Windows 3.1.

The technical criteria which provided a framework for both the development and evaluation of the concordance program were originally derived from Hubbard (1992) and Stevens (1995) but were further refined as the project advanced.

- 1) **Speed:** Searching must obviously be quick, otherwise it becomes impractical. Searches on ASCII and ANSI encoded texts are fast because of the lack of embedded formatting codes. Searches on HTML texts take longer because of the sheer number of character and formatting codes that they contain.
- 2) **Ability to cope with diacritics and inflected forms:** The need to cope with diacritics required the writing of a code routine that would allow for

searches using different character sets. A search based on 'après' requires the word to be broken down into three parts: 'apr', 'è' and 's'. One possibility would be to break the word into its three elements, replace the 'e grave' by the respective code (ANSI, ASCII or HTML), recombine the three elements and perform the search. This would work but would not find upper case examples. HTML is even more complicated. There are officially two separate means of encoding most diacritics (and certain other non-text characters) and three for the 'oe ligature' digraph. To complicate the issue, HTML texts are occasionally found which use standard ANSI codes and, increasingly, UTF-8 codes. The solution adopted was to perform a search in three stages: first find 'apr', then to check that this is followed by 'è' by cycling through the different code sets, both lower and upper case if the search is not case sensitive, and then, if successful, to check for 's'. This solution maximises the chances of finding search items and, in addition, allows for the future addition of other code sets.

- 3) **Flexibility:** To achieve maximum flexibility, the program had to allow the following:
 - a) *The searching of texts of any length and in any number.* Some of the earlier concordancers were designed for literary concordancing based on a single large text. Later programs were designed for use with corpora and so allowed for the searching of large numbers of texts. To allow thematic organisation of the corpus, it was decided to design the present program to allow searches on single or multiple texts in a single directory, on a single directory and all its subdirectories or on a whole disk, including all directories and subdirectories. Finally the possibility of creating a file list combining files from different directories and even disks was added.
 - b) *A variety of keyword options.* It should be possible to enter keywords as single words or as phrases, in upper or lower case, with the option of a case sensitive search. Wildcards should be permissible and, to reduce the possibility of unwanted forms, the inclusion of forms to exclude from the search should also be possible. A later addition to the criteria was the possibility of specifying a position to the right of

each term in the keyword / key phrase at which or before which the following term had to be present (proximity searches).

- c) *Both automatic or user selection of default code set.* HTML texts can be automatically identified by their header codes and the appropriate code set selected. This is not possible in the case of ASCII and ANSI documents. It was therefore essential to allow the user to change the default from ANSI to ASCII where it was known or suspected that the texts had been created under MS-DOS and not Windows.
- d) *User to select the target disk and / or directory.* It would have been possible, to make the programming simpler, to save search results to the root of the C: drive, especially since one can be sure that all modern PCs will have this if no other drive. However, such a course of action would have precluded wildcard searches of files on the C: drive and so reduced the flexibility of the program
- e) *User to determine the left and right horizons of the keyword search.* A relatively small horizon (or context) makes for faster searches and reduces the chances of memory problems. At the same time, the amount of contextual information attached to the keyword is reduced. By providing the option of varying both horizons, users are allowed more control over the way the results will be saved. This is important if one wants to be able to go beyond the sentence as immediate context.
- f) *Searches with or without the results being saved to file.* One use of the concordancer is to search for the presence within a text or corpus of a certain term or terms. With a small corpus or with a rare word or phrase there is a strong possibility that the search will be unsuccessful. In addition, searches carried out by students are often unsuccessful because they have wrongly entered the keyword. In either case, it seems pointless creating an empty results file.
- g) *Users to view the results while the search is being carried out.* Being able to see the search as it takes place leads to greater involvement

and is more motivating. Also, it is possible to identify success or failure immediately and take appropriate action.

- h) *Users to be able to abandon the search.* If users realise that they have made a mistake in entering keywords or selecting files or if it becomes clear that the search has already produced sufficient examples or will take much longer than anticipated, there should be the option of abandoning the search at any point. It is nevertheless important that the results should be saved and not lost.
- i) *Viewing of the results in concordance format.* One of the essential features of a concordance is the way the results are presented with the keyword in the centre of the screen/page. This feature is likely to be lost if the results have to be read using a text-editor or word-processor.
- j) *Manipulation of the search results:* The user should be able to:
 - 1) scroll up and down
 - 2) view the keyword in its full context as well as in concordance format
 - 3) delete any unwanted examples
 - 4) sort on the keyword and on designated positions to the left and right
 - 5) save to memory and / or print some or all of the examples
 - 6) save any modifications to file.

Not all of the last few criteria were identified initially. Some, such as sorting on the results, were added as a result of using other programs. Others, such as deleting unwanted examples, were added as and when there was sufficient time to write and test the necessary program code.

Initially, the program was conceived purely as a concordancer and so the option of performing word-frequency counts was not included. However, as the project progressed, the usefulness of such a facility made itself felt and so the original conception was extended to include this.

- 4) **User-friendliness:** This may not appear necessary in the case of software written for a single user for a specific purpose. However, in the event of breaks in development, it is surprising how even the programmer can forget how to use his / her own program. Selfishness therefore required building in a minimum level of user-friendliness.

Once however the program was to be used by students, many of them relatively ill at ease with computers and none of them familiar with concordancing, it became essential to give much greater weight to the issue of user-friendliness. One of the first decisions was to move from QuickBasic 4.5 to Visual Basic (progressively versions 1, 3, 4 and 6) in order to create a Windows style interface and mode of operation.

As the project advanced a number of conceptual rather than technical problems arose due to the fact that the program had not been tested with students. These are discussed below.

5.2.3 *THE EXPERIMENTAL GROUPS*

The choice of student groups was ultimately determined by the availability of the only computer room which had the language software installed. As a result the first phase of the experiment was conducted with first and second year Combined Honours students. All students involved were post 'A' Level entrants. The target learner groups involved in the experiment were first and second year degree students. The most important variables that categorise the students are:

- | | | |
|----|--------------------|---|
| 1) | Age: | 18+ |
| 2) | Native language: | predominantly English |
| 3) | Proficiency level: | post 'A' Level (7+ years of secondary French) |
| 4) | Sex: | Year 1: male 4 female 9 |
| | | Year 2: male 4 female 14 |

5.2.4 *IMPLEMENTATION OF THE CONCORDANCER: STAGE 1*

Materials developed by the TELL Consortium⁷ were already being used in the Year 1 and Year 2 French language elements of the College's Combined Honours degree. These were essentially suites of grammar exercises and

⁷ The Technology Enhanced Language Learning project funded by the UK Higher Education funding bodies. Materials are free to UK Higher Education institutions but otherwise commercially available from Hodder & Stoughton.

explanations, or programs to provide practice in translation. Student reaction had been positive and so it seemed a logical step to add to the range of activities. Existing activities were felt to be useful but largely duplicated what was available in books, with however the added advantage of interactivity. Nevertheless, the grammar programs still present rules to students in traditional pre-packaged form. Since the intention was to get students to formulate rules for themselves, it was necessary to present them with the raw data, through use of the concordancer. Certain advocates of concordancing use the software as a way of resolving student uncertainties and queries (Higgins 1991; Johns 1991a). The concordance is created as a result of a problem identified in class and the results used in a subsequent class as a basis for analysis and discussion. The linguistic items that are the object of the concordances are not chosen by the teacher in advance. Conversely, there are instances where grammatical points have been identified in advance (Ching & Wong 1994, Wong *et al* 1992) and then successfully exploited in class, even if student responses were mixed. Because the main aim of the experiment was to make grammar learning more learner-centred, ideally the former procedure should have been adopted. However, the Combined Honours course requires detailed syllabuses and schemes of work to be in place at the beginning of the academic year⁸. There is also the question whether most students, certainly in the first year, are grammatically aware enough spontaneously to identify problems and needs.

Therefore specific linguistic items in the revision syllabus which were identified as lending themselves to concordance based analysis were slotted into a new grammar scheme of work constructed around the scheduled fortnightly computer session of one hour's duration. As a result, there was a high level of teacher direction. The specific points chosen to begin with were:

- | | |
|---------|--|
| Year 1: | determiners and noun phrase structure
position of adjectives
noun gender |
| Year 2: | third person clitic and disjunctive pronouns
demonstrative pronouns
relative pronouns
subjunctive mood. |

The first year language course begins with a thorough revision of noun phrase structure since experience has shown that this is where students not only make most mistakes in reading comprehension and translation (noun / adjective order and expanded adjectival constructions) but also in writing (noun genders and adjectival agreements). Nouns are relatively easy to search for since the only alternative form is the plural. Adjectives are more complicated but normally limited to four forms at most. By concentrating on nouns and adjectives, it is possible to perform searches quickly and reliably.

The first term of the second year course contains a thorough revision of pronouns including forms, position and functions. Apart from the direct object pronouns, these are fairly distinctive and thus relatively easy to perform searches on. Verbs are generally more difficult not only to search for, given the number of endings, but also to derive adequate rules of use, especially from a concordance. However, revision of the subjunctive is part of the syllabus and it was decided to include this since searches can be made on a small number of distinctive forms and the rule can usually be derived from the concordance, especially where forms follow conjunctions.

Worksheets were prepared in advance and handed out at the beginning of the lesson (examples can be found in Appendix 2). These had two main purposes:

1. to identify the grammatical point to be investigated and the specific lexical items to be searched for,
2. to guide students through the search process, taking into account the fact that some would have little or no experience of computers while none would have used a concordancer.

As a result the worksheets for the first few sessions were quite detailed. To make the concordances manageable, searches were restricted to specific directories and these were indicated on the sheets. Instructions for each stage of the process were provided. At the end of each sheet there were two boxes: the first asked students to draw conclusions from their observations while the second asked them to formulate a rule.

In order to investigate noun phrase structure and the role of determiners, **Year 1** students were asked to produce a concordance on the keyword *France* (activity 1a) and then to:

1. list the words that immediately preceded it,
2. translate examples into English,
3. compare French and English usage.

The second stage was to repeat the process with *Paris* (activity 1b), carrying out the three stages of analysis and then comparing the structures surrounding the two keywords. The intention was to repeat the exercise, but without translating into English, based on the keyword *pays* (activity 1c) but this time with the focus on the accompanying adjectives. Because the first two phases took longer than expected, mainly because of students' lack of familiarity with the software but in some cases also with Windows 3.1, the third phase had to be carried out in class using a prepared concordance. The aim was to make students more aware of the necessity of the determiner and to think about the position of adjectives. The latter point was followed up with a new session, once again based on a worksheet (activity 1d), in which students worked as groups on three sets of adjectives: the first set, *français, rouge, blanc*, students were supposed to identify as coming after the noun, the second set, *bon, mauvais, petit, grand*, as coming before and the third set, *certain, dernier, long, ancien, curieux, vif/vive*, as coming both before and after. The choice of keywords was determined by what was in the corpus.

The last session on the noun phrase focused on gender, with groups of students being given sets of mixed masculine and feminine noun suffixes to search for and asked to identify regularities and classify them. From the teacher point of view, the sessions seemed to proceed well with the students active all the time and without any major hardware or software failures. In the follow up discussions, there was a gap between students who had a reasonably good grasp of grammatical concepts and could therefore identify patterns and those, admittedly a minority, who failed to see anything. Having students discuss results first as a group can help with this latter category. There were some problems with noun gender, usually caused by the possessive adjectives *mon* and *son*, which highlighted how easily students can base false conclusions on rational interpretations. There were even more difficulties with the third set of adjectives

where the student-formulated rule was at times the reverse of the normally accepted one. This was even after some vetting since certain adjectives, especially *ancien* and *dernier*, can generate concordances which appear to be contradictory. Generally, however, students reached conclusions that were rational and in accordance with linguistic reality, especially when there was sufficient time for group discussion and comparison of results.

As part of the general revision of the pronoun system, **Year 2** students were asked to begin by focusing on *elle* and *lui* in order to (1) revise the notion of direct and indirect objects; (2) highlight the fact that indirect object *lui* is both masculine and feminine and (3) extend the revision of pronouns begun in Year 1 to include the disjunctive pronouns. The computer session was preceded by a classroom session on subject pronouns and a detailed worksheet was prepared (activity 2a). Once again it became clear that students with a sound conceptual base could perceive patterns and formulate rules while the few who still had no clear idea of the difference between subject and object still saw pronoun use as purely arbitrary. Revision of demonstrative pronouns required students to search for *celui* and *celle* (activity 2b) and while the search produced potentially useful results, the students failed to formulate any clear rules for themselves. As a result it was decided not to proceed with the possessive pronouns worksheet (activity 2c) but to use a less prescriptive format with the remaining pronouns. Work on relative pronouns focused on *dont* and *lequel/laquelle*. Students seemed to remember the basic rule for *dont* quite well, without always being able to apply it, and so it is uncertain whether the concordance added to their knowledge.

The *lequel/laquelle* concordance did however highlight the use of the pronouns after prepositions and was thus valuable. Ideally, as part of a total revision of the relative pronouns, *qui* and *que* should have been included but these highlight the limits of a non-tagged corpus. Any concordance based on these, especially on *que*, and intended for teaching purposes requires extensive editing to remove all the non-relative uses. Pronouns were followed by revision (or, for many students, introduction) of the subjunctive. This had been held over to Year 2 to allow a comprehensive revision of the indicative tenses in the first year. Once again, the question of the limits of a non-tagged corpus raises itself. The forms *ait*, *soient*, *puisse*, *sache*, *fasse* were selected as keywords since prior

investigation had shown that these would produce sufficient examples to work on. *Soit* was excluded because it produced 230 examples, relatively few of which were examples of its use as a subjunctive form in opposition to the indicative *est*.

Students were able to identify examples conforming to the rules that they already knew but were poor at deducing any new rules from the linguistic data. The rules that they did advance were often based on a false interpretation of the data. In the case of the subjunctive, some students thought that *que* alone was sufficient cause, others that the subjunctive was required after “verbs of believing”. Both of these are over-generalisations and can be put down to the methodology which only required students to interpret clauses containing a subjunctive without being able to compare these with other clauses containing an indicative. This was especially so with verbs like *croire* and *penser*. In addition, students seemed more ready to salvage from memory partially remembered rules that they had been given in the past rather than formulate a new rule for themselves. This was in part confirmed by the results of the follow-up questionnaire (see section 5.2.5 below). Group discussions were less successful in this lesson because students who thought they knew the rule tended to dominate and, all too often, misled the others, effectively undermining the discovery-based methodology underlying the experiment. However, on a positive note, the concordances did lead to a more sustained discussion of the point with greater student participation than would normally happen in a formal grammar lesson.

It is tempting to assume that concordancing is not appropriate with similar undergraduate groups but the difficulties arose largely from the physical circumstances. What is clear is that the tutor needs to be closely involved in the discussions, steering students away from false conclusions without imposing his or her own interpretation. This can be exceedingly difficult in a large computer room where students are split into two main groups by a central bank of PCs, where there is no single point of focus, where half of the group will have their backs to the tutor for most of the session, and where the poor acoustics are further degraded by the humming from the machines. Time available for discussion can be reduced by the amount of time wasted walking from one building to another, clearing other students from the room and waiting for members of the group who have fitted in one or two other activities on their

journey across campus. Follow-up in the next session a week later may be resented by some students who complain that the scheme of work is not being respected. Two hour sessions would have avoided most of the problems but would have used all the language module's weekly timetable allocation.

5.2.5 *THE STUDENTS' RESPONSE*

For both Years 1 and 2, the above activities all took place in the Autumn Term. The intention was to continue into the Spring Term but feedback at the end of the first term suggested that the activity was fairly unpopular with a majority of students. The main reason given was that it took too long to gather the information. Students could not see why the concordances could not be prepared in advance and handed out at the beginning of the lesson. This would allow more time for what they saw as more useful activities, essentially practice of translation in Year 1 and summary in Year 2. This highlights the potential for different perceptions between teachers and students. Their attitude is in many respects logical but disappointing as they seemed not to accept the premise that extracting the concordances themselves was intended to make them more involved in and thus more conscious of the process. When it was explained to them that one of the objectives was to give them a useful learning tool which they could use independently, those who responded suggested that the idea sounded nice in theory but in reality they would not have the time, or the inclination.

5.2.6 *PRACTICAL PROBLEMS THAT AROSE*

The students' frustration can in part be explained by inadequate printing facilities - 10 to 15 minutes at the end of a lesson could be spent waiting for print-out and often no hard-copy appeared at all. However, some of the difficulties that arose related to the software. One significant problem was that, in writing the program, the programmer is forced to perceive language as a computer does, that is simply as a string of characters. A space is no different from any other character, except for the few which are reserved for use as line feeds, carriage returns, etc. For the computer, words only become words if the boundaries are explicitly marked. To search for *qui* with the earliest versions of the program it was necessary to enter it as '*qui*'. If the slashes were omitted then the search would come up with

everything from *quinconce* to *maquisard*. For the programmer, who has to learn to “think” like a computer, such a procedure is not a problem. However, from the students’ point of view it was counter-intuitive. If they left a slash off or put the wrong one in, then the results were unpredictable. Early on then in the experiment, it proved necessary to include an interpreter stage to translate human concepts into machine concepts.

Other modifications were prompted by the specific difficulties thrown up by working in a language other than English. Nearly all the literature on concordancing is based on practice in English, mainly in English as a Foreign Language. English lacks many things such as accents. Changes in French verb stems, usually permutations of ‘e’, ‘é’ or ‘è’, require a search for ‘*espérer*’ to include both ‘*espér*’ and ‘*espèr*’. Other difficulties are caused by the more inflected nature of French. To search for an adjective in English, one has only to enter one form. However, most if not all adjectives in French have distinctive masculine and feminine singular and plural forms making four in all. A simple solution is to use a standard DOS wildcard such as ‘?’ or ‘*’. Thus, a search for ‘*français**’ will find all forms of the adjective and nothing else. However, a search on ‘*blanc**’ will produce, in addition to the desired adjectival forms, *blanchâtre*, *blanchir* (and all its inflections), *blanchissage*, *blanchisseur*, etc. With verbs, the situation becomes even more complicated. The form ‘*esp?r**’ will find all forms of the verb (plus *espérance* but this is not too much of a problem since it is predictable and the program allows the inclusion of words to exclude). A search however for all the forms of the imperfect of *être* based on ‘*ét**’ will, in addition to finding *étant* and *été* (including of course *été(s)* meaning ‘summer’), also find totally unrelated forms ranging from *établi* via *étranger* to *étudiés*. The corpus produced a total of 2590 responses of which only 377 were forms of the imperfect. Most of the unwanted forms are difficult if not impossible to predict and exclude.

The simplest solution, and the one first adopted, was simply to enter each word separately. However, in the case of the imperfect, this produced five separate files which made follow-up analysis difficult. This required the addition of a new option allowing results to be saved to a single file. Ultimately, it proved desirable to introduce two new search procedures based on the use of special characters.

The slash '/' allows a search for incremental endings. Thus the single form 'blanc/he,s' will find all masculine and feminine singular and plural forms and nothing else. The combination 'esp?r* qu/e' will find all forms of *espérer* (plus *espérance*) before both 'qu'' and 'que'. The sign '+' allows searches for a stem plus endings. Thus the single form 'ét+ais,ait,ions,iez,aient' will find these and nothing else. While making the program more flexible, these changes did however make it more complicated. The progress of development was faster than that of updating the manual and so students were not always able to use the new features.

Other features that were progressively incorporated were the ability to print or copy to memory the concordance results; to sort them and save/copy/print them in sorted form; to delete unwanted forms (eg *été* meaning 'summer' instead of 'been'); to delete the results files on exit to avoid cluttering up the hard disks; and finally to perform word counts. Another modification was to review all program instructions that appeared on screen to make them more user-friendly to non-specialists. This meant replacing, for example, the phrase "source files" by "files to search" and "target directory" by "save results to..".

Ultimately, though, however much effort one puts into the software, it cannot guard against hardware problems or students' lack of forethought. If the software is copied daily from a central server and something goes wrong in the process, this can sabotage a session with students. If the printers are not working or the printer switches point to the wrong number, students can become quickly frustrated and blame the activity. If they do not bring a floppy with them to save the results to, they cannot work on these afterwards.

5.2.7 PRELIMINARY CONCLUSIONS

In the light of the student responses, it was important to take stock and decide what lessons could be learned. The feedback indicated that the students did not enjoy the computer work and especially the concordancing. This was especially true of the weaker students. This was in marked contrast to the previous year when reaction to the computer had been generally favourable. The most obvious new variable was the concordancer. However, from a careful reading of the feedback returns, other factors became apparent and one of these was frequency.

There was a general feeling among the students that the computer should not be timetabled too often. One session every three weeks, it seemed, maintained interest whereas once a fortnight killed it. Lack of familiarity with computers is another factor which was mentioned, mainly but not exclusively by the mature students. A few students could not find the corpus texts because they were not familiar with directories and sub-directories. Some students had difficulty in typing the keywords in accurately, more so in the early stages when they had to enter the slashes. Some would keep “losing” the program, usually when they clicked outside the program window on the background Program Manager screen, and would start from scratch again, leading to a situation where they had two or three instances of the program running concurrently. This would have been less of a problem with Windows 95 and NT and did not happen with the other CALL programs which filled the whole screen. Many of these problems would not happen with present cohorts of students who are noticeably more familiar and confident with computers.

The main problem, however, that the experiment threw up was the inability of many students to process the results, confirming Kenning’s (1996:131) observation: “Even with prepared activities, some learners will need considerable prompting, missing what are for the teacher obvious clues.” A first year activity based on noun endings and gender worked well with most students making the right connections. So did one focusing on the environment surrounding, and especially preceding, the words *France* and *Paris* and contrasting the results with English. This was designed to focus on use of determiners and in addition the position of adjectives. Articles and adjectives are familiar to native-speakers of English and do not require any conceptual leaps. The same is true of the nouns involved. Therefore the contrastive analysis of the structure of noun groups in the two languages could yield the results desired.

Searching on *elle* and *lui* with a second year group as part of revision of personal pronouns merely highlighted the fact that if students have no clear idea of the difference between subject and object they cannot make sense of the pronouns. Feedback from students suggests that many resist having to learn concepts that

they have been told do not matter⁹. Another exercise with the same group investigating the subjunctive showed that their interpretation of the data was mediated through their existing misunderstandings, resulting in various distortions of the evidence. As Manning (1994) found, “learners build their own rule system and misapply the rules [...] or they fail to establish the right connections between the important elements”(op.cit.:11). This point is echoed in Kennedy and Miceli 2001:81): “lack of rigor in observation and reasoning contributed greatly to [students’] difficulties.”

In many respects, the experiment reinforces what is known about most self-access language learning. If students have a good conceptual basis, they will be able to learn or, in other words, teach themselves the language (this point will be discussed in greater detail in Chapter 6). It is also necessary for students to accept the principle of discovery learning and formulation of their own rules. The feedback demonstrated that many students were unwilling to accept this basic principle and preferred a more passive role.¹⁰

5.2.8 *IMPLEMENTATION OF THE CONCORDANCER: STAGE 2*

The intention was to continue the experiment in the academic year 1999-2000. The only major changes envisaged were to be made to the concordancer thus removing some of the problems outlined above. This would require some minor changes to the worksheets. Otherwise the only difference planned was to timetable classes every three rather than two weeks. It was at that moment that the College chose to upgrade its IT provision in anticipation of the new millennium, necessitating the total replacement of all PCs and a change of operating system. This meant that the concordancer and support text files were not reinstalled until half-way through the second term, too late to be incorporated into the timetable. The delay did however make it possible to rewrite the program so that it was fully functional on a 32-bit operating system. This

9 The lack of grammatical awareness among many students has already been discussed in some detail in Chapter 2, especially in the context of the dominant communicative methodology.

10 In the survey carried out in October 1998, when asked how they thought they learned best, 27% of Y1 students and 35% of Y2 thought it was by working things out for themselves, 87% of Y1 and 85% of Y2 thought it was by having things explained by the teacher, 87% of Y1 and 92% of Y2 thought it was by doing exercises and only 60% of Y1 and 50% of Y2 thought it was by talking to other students

allowed the use of long file and directory names and also removed difficulties arising with networked drives.

In the circumstances, the choice was either to abandon the experiment, at least on a temporary basis, or to adapt to circumstances and use the concordancer in a different way. Students could still work with concordances but would not be directly involved in the gathering of data, something which many had felt was a waste of time. In practice, this led to a situation where work with students was in many respects an extension of personal research. Where possible, students would be invited to join in the process of interpreting the data. At the very least they would be exposed to the raw data rather than having this hidden from them in an interpreted form. It would also be necessary to bear in mind the results of the previous year's questionnaire and the attitudes of students to rule formulation. One new feature of the concordancer was the ability to make word-counts of single or multiple texts which would permit greater focus on lexis. An extra advantage is that this facility makes it much easier to produce rapid, complete concordances of single texts based on an alphabetical word list.

In the first unit of the First Year French Grammar and Usage module, the first use of the concordancer was to extract the most frequent words from the first listening text to highlight the notion of a "core vocabulary" and to test students' knowledge of these (see Figure 1 below). The text was a recording taken from the *Voix d'Auvergne* (1977) archive, the gapped form of which can be found in Appendix 3. The text was not seen as being in any way representative beyond the fact that it is an authentic example of an unprepared interview with a young Frenchwoman of roughly the same age and educational background as the students. Its main function was to provide both a model for students talking about themselves and their course and a source of relevant vocabulary. As can be seen, most of these items in Figure 1 can be described as function or grammatical words and these were the focal point of the lesson. Furthermore, the items have been categorised as one of the objectives of the activity was to develop awareness of parts of speech and to introduce the appropriate terminology.

Below is a list of more core vocabulary items. Insert the English equivalents:				
Word		Frequency	%	norm
6	on	13	3.3592%	03.4 per 100
25	ça	3	0.7752%	00.8 per 100
139	tout	1	0.2584%	00.3 per 100
7	à	12	3.1008%	03.1 per 100
13	en	8	2.0672%	02.1 per 100
42	avec	2	0.5168%	00.5 per 100
46	dans	2	0.5168%	00.5 per 100
77	après	1	0.2584%	00.3 per 100
85	comme	1	0.2584%	00.3 per 100
131	pour	1	0.2584%	00.3 per 100
132	sans	1	0.2584%	00.3 per 100
134	sur	1	0.2584%	00.3 per 100
8	et	11	2.8424%	02.8 per 100
115	mais	1	0.2584%	00.3 per 100
121	ou	1	0.2584%	00.3 per 100
122	où	1	0.2584%	00.3 per 100
123	parce que	1	0.2584%	00.3 per 100
61	peu	2	0.5168%	00.5 per 100
81	beaucoup	1	0.2584%	00.3 per 100
38	très	3	0.7752%	00.8 per 100
78	assez	1	0.2584%	00.3 per 100
84	combien	1	0.2584%	00.3 per 100
9	alors	10	2.5840%	02.6 per 100
93	donc	1	0.2584%	00.3 per 100
136	surtout	1	0.2584%	00.3 per 100
50	ensuite	2	0.5168%	00.5 per 100
62	puis	2	0.5168%	00.5 per 100
95	encore	1	0.2584%	00.3 per 100
96	enfin	1	0.2584%	00.3 per 100
112	là	1	0.2584%	00.3 per 100
143	voilà	1	0.2584%	00.3 per 100
57	même	2	0.5168%	00.5 per 100
67	tous	2	0.5168%	00.5 per 100
79	autre	1	0.2584%	00.3 per 100
16	(il) y a	5	1.2920%	01.3 per 100

Figure 1: “Below is a list of core vocabulary items...”

This was followed up later in the unit by two similar exercises, one focusing upon high frequency lexical items such as *année*, *dire*, *chose* and the other upon the more restricted vocabulary belonging to the text’s semantic field (education and the training of physiotherapists) and the clues they give to the text’s topic¹¹. The next activity was to present a concordance extract containing a small sample of functional words in order to highlight the difficulties of assigning equivalent meanings to them and the importance of context (see Figure 2 below).

¹¹ It must be remembered that the purpose of the activity was not to make a linguistic statement but to provide students with vocabulary relevant to the tasks they had to do.

Some very common words can only be understood and translated in context. How would you translate the following examples ?

is étudiante en kinésithérapie à Clermont-Ferrand hein euh alors voilà c'est surtout euh indispensable c'est-à-dire l'anatomie la physiologie bon alors ça ie la physiologie bon alors ça à fond et puis ensuite à la fin de la première lors ça à fond et puis ensuite à la fin de la première année on commence à a la première année on commence à aller tous les matins à l'hôpital - et à tr mmmence à aller tous les matins à l'hôpital - et à travailler avec des gens s us les matins à l'hôpital - et à travailler avec des gens sur des malades et oup de de concours pour entrer à cette école bien euh c'est - enfin ça devie t alors y a y a eu un concours à partir de cette année c'est la première ann ent avec un baccalauréat c'est-à-dire euh un certificat de fin d'études au l un enseignement très théorique au départ - on apprend bon ce qui est euh sur e l'école qui part en vacances au mois de juillet et l'autre moitié au mois un certificat de fin d'études au lycée mais euh l'année dernière donc il y

tout et vous faites des études de combien d'années alors là mes études duren alors parlez-moi un petit peu de la formation de - kinésithérapeute oui c'e i un petit peu de la formation de - kinésithérapeute oui c'est (rire) alors fond et puis ensuite à la fin de la première année on commence à aller tous n est à l'hôpital on a un mois de vacances pendant l'été on on intercale y a on on intercale y a la moitié de l'école qui part en vacances au mois de ju e qui part en vacances au mois de juillet et l'autre moitié au mois d'août e ller est-ce qu'il y a beaucoup de de concours pour entrer à cette école bien 'est - enfin ça devient disons de plus en plus difficile et alors y a y a eu a y a eu un concours à partir de cette année c'est la première année où il c'est-à-dire euh un certificat de fin d'études au lycée mais euh l'année de

lors euh moi je suis étudiante en kinésithérapie à Clermont-Ferrand hein euh en organisé comme enseignement en France la kinésithérapie alors parlez-moi la moitié de l'école qui part en vacances au mois de juillet et l'autre mo moitié au mois d'août et puis en troisième année c'est la même chose on con nfin ça devient disons de plus en plus difficile et alors y a y a eu un conc 'est la première année où il y en a un bon moi j'y suis rentrée je suis en t

oiselle est-ce que vous voulez bien me dire ce que vous faites dans la vie o ue vous faites dans la vie oui bien sûr (rire) alors euh moi je suis étudia est c'est pas encore très très bien organisé comme enseignement en France la ute oui c'est (rire) alors euh bien la première année on commence euh on a u ours pour entrer à cette école bien euh c'est - enfin ça devient disons de p

e on passe un diplôme national qu'on a ou pas et (alors) ensuite on peut tra uite on peut travailler est-ce qu'il y a beaucoup de de concours pour entrer ue vous voulez bien me dire ce que vous faites dans la vie oui bien sûr (rir nc il y a eu un concours parce que ça devient de plus en plus difficile et i

au départ - on apprend bon ce qui est euh surtout euh indispensable c'est-à rcale y a la moitié de l'école qui part en vacances au mois de juillet et l'

en sûr (rire) alors euh moi je suis étudiante en kinésithérapie à Clermont-F ée où il y en a un bon moi j'y suis rentrée je suis en troisième année cette

Figure 2: "Some very common words can only be understood..."

The next unit followed up on the distinction between "grammatical" and "content" words and presented a new concordance extract once again focusing on high frequency items whose meaning was highly dependent on context (*à, de, des, était, le/la/l', leur, lui, ne*). The text originated from a magazine for French teenagers and dealt with the marital circumstances of the pop-star Michel Sardou. By presenting a concordance extract (see Figure 3 below) focusing exclusively on the two participants (Michel and Babette Sardou) and their roles as grammatical subjects and objects, it was possible not only to work on these two functions but also to demonstrate how the grammar revealed the "power

relationship” within the couple. Once the text had been analysed in this way, it became possible to discuss the “message” or “promise” contained in the text and to see how this might appeal to young adolescent girls’ fantasies about marrying pop stars, thus enabling discussion of the purpose of the text and the target audience.

What are the dominant grammatical roles respectively of Babette and Michelle? Why are there several examples of *il* and none of *elle* ?

rait sur son éventuel divorce. Babette était d'ailleurs venue le rejoindre à age, Michel a décidé d'emmener Babette pour une seconde lune de miel à Venise llets pour Venise, en disant à Babette: "Prépare tes valises, nous partons d frit une montagne de cadeaux à Babette. Des cadeaux qui auront toujours une aleur sentimentale aux yeux de Babette, puisqu'ils lui rappelleront à jamais el était rayonnant de bonheur, Babette riait aux éclats. Ils se sont amusés étente, d'amour pour Michel et Babette. Michel apprécia d'autant plus cette l réserve une autre surprise à Babette: passer tout l'été à ses côtés en Sar

ait une interview exclusive de Michel dans laquelle il niait fermement la ru ur amour n'a subi aucun nuage, Michel a décidé d'emmener Babette pour une se ne de miel à Venise. Ce voyage Michel l'a décidé sur un coup tête, il le fai rent ensuite du lèche-vitrine. Michel offrit une montagne de cadeaux à Babet ce merveilleux voyage d'amour. Michel était rayonnant de bonheur, Babette ri epos, de détente, d'amour pour Michel et Babette. Michel apprécia d'autant p 'amour pour Michel et Babette. Michel apprécia d'autant plus cette escapade

lusive de Michel dans laquelle il niait fermement la rumeur qui courait sur l l'a décidé sur un coup tête, il le fait pour tout. Un matin il est arrivé il le fait pour tout. Un matin il est arrivé chez lui, avec en poche deux bi escapade que, dès sa rentrée, il devait affronter le public parisien au Pal en au Palais des Congrès. Mais il réserve une autre surprise à Babette: pass

How is this reinforced by *lui* in the examples below?

t. Un matin il est arrivé chez lui, avec en poche deux billets pour Venise, ux yeux de Babette, puisqu'ils lui rappelleront à jamais ce merveilleux voya

Figure 3: “What are the dominant grammatical roles...”

The units that followed focused upon pronouns and verb tenses. In the case of the former, the concordancer was used merely to provide examples of authentic use of pronouns that could be incorporated into the teaching materials. Pronouns also point up the limits of concordancing using “raw” as opposed to “tagged” texts. A search on the direct object pronouns will also find all examples of the definite article and the only way to separate the two, in the absence of parts of speech tags, is to go through the concordance list item by item. In the case of the original 273 files, such a search produced over 37000 examples. By restricting the search to a single directory of 13 files, this number was reduced to a mere 1859, still large but more manageable. Searches on random single files often produce few useful examples. The other pronouns are more distinctive and searches produced sufficient examples that could be incorporated into course

materials, mainly exercises requiring substitution (in both directions) or tracing pronouns back to their referents.

The first units of the **First Year** language module constitute a language awareness course with the focus on word order and grammatical categories. The rest of the year is devoted to a systematic revision of the verb system from the bottom up. Because of the sheer number of endings and the fact that many of these are not distinctive, French verbs do not lend themselves easily to concordancing unless one treats them as a single lexical item, e.g. *faire*, *prendre*. Hence little use was made of concordances in the first year to investigate verb forms. An exception was the subjunctive¹² and a similar procedure to that described in 5.2.5 above was followed. Generally, it was felt to be both easier and more efficient to use selected texts containing contextualised and contrasting examples of tense usage.

The first use of the concordancer in the **Second Year** Language module was to re-use the small subjunctive concordance mentioned above as a revision activity. It was not until the topic of relative pronouns was introduced, following several weeks' work on verb forms such as modals and passives, that the concordancer was used to find examples of *qui*, *que*, *dont*, *lequel*, *le mien*, *celui* that could be incorporated into descriptions or exercises (Appendix 3). The main emphasis in Year Two is on the development of summary skills requiring more work on reading and interpreting texts which leaves less time for grammatical work.

In **Year Three** students only have one hour a week for written language and the emphasis is mainly on translation. No real attempt was made to use concordances systematically in the teaching and revision of grammar. However, they were regularly used to throw light upon errors that occurred in homework. Examples of the sort of errors that were addressed were confusion of *à cause de* and *parce que*, *change* and *changement*, *avant*, *avant de* and *avant que*. At first, in order to use the results in class, they had to be saved by first capturing the full screen and then copying the image to file before printing it out as an overhead transparency as in Figure 4 below. As this procedure proved to be rather

12 This has since been moved to the Second Year as it has proved necessary to devote more time in Year One to the revision of the indicative tenses.

complicated, the program was partially rewritten to permit the saving to memory or the direct printing of sections of the concordance results. As a result students could be given individual copies. These could even be produced in the break between two one hour sessions.

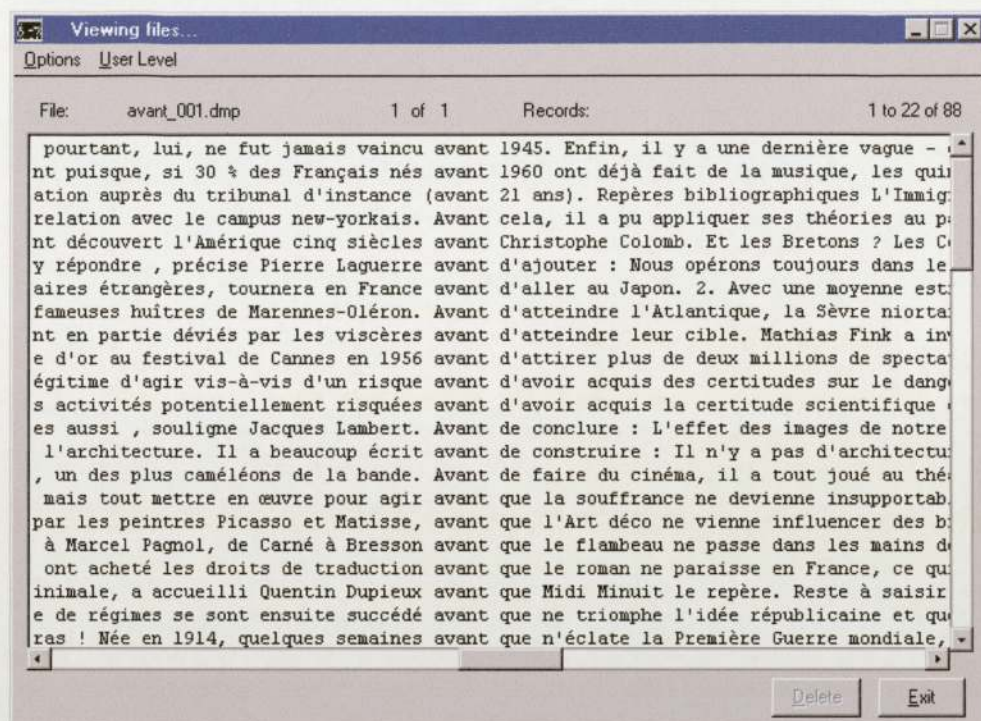


Figure 4: concordance of *avant*

Concordances were also used in the French Business Context module, mainly to highlight problems arising from interference with English such as the semantic limits of *manager*, *cadre* and *dirigeant*; *management* and *gestion*; *marketing* and *commercialisation*.

The situation should have changed in the academic year 2000-2001. The software and corpus files had been installed in the computer rooms but technical and timetabling difficulties delayed their use with students until the Spring Term. This made it impossible to conduct the experiment as originally designed. At a practical level, the best solution seemed to be to extend the use of the concordancer in the preparation of course materials and especially in Years 2 and 3. Concordance search results, in both extract and expanded form, were incorporated into the language units dealing with modal, passive and reflexive verb forms to enable identification and analysis of the forms themselves. More authentic examples of usage derived from the corpus were then used to produce

exercises designed to develop awareness of the functions and to highlight differences between French and English, as in the case of reflexive pronouns¹³. However, students are still introduced to the concordancer during the first term's skills week as part of a general induction to CALL (see Appendix 4).

5.3 CONCLUSIONS

The most obvious conclusion is that any attempt to use modern technology is dependant not just upon the flexibility and reliability of the equipment itself but also upon the way that the technology is managed as a resource. Problems arose from the beginning in the experiment due to the inadequacies of the printing facilities. Switches were wrongly labelled, connections were sometimes missing and printers were often without paper. The frustration that this caused among the students was one factor explaining their early negative reaction to using the concordancer. Delays in installing the software and difficulties of access meant that use of the concordancer had to be opportunistic rather than programmed.

Experience suggests that, given the circumstances in which the experiment took place, the concordancer is best used as a tool by the tutor in the preparation of course materials. At a purely practical level, the preparation time thus falls to the tutor and does not engender resentment on the part of the students. Such a procedure also recognises two other important factors. One is the highly prescriptive nature of the syllabus which, whether pedagogically desirable or not, constitutes an unavoidable constraint. It is therefore the tutor's responsibility to identify the linguistic items to be studied and provide the basic materials. The advantage of the concordancer is that it is no longer necessary to resort to contrived examples that reflect the tutor's intuitions, which may be wrong or simply incomplete. Instead it is possible to build up a bank of real or authentic examples which initially obliges the tutor to see how the linguistic reality matches up to his or her intuitions. Only then should the process of selection and presentation, including decisions about the amount of accompanying contextual information, be undertaken. By then inviting the students to explore the data given to them and formulate hypotheses, the principle of discovery learning can be preserved.

13 Other example of concordance based exercises can be found on-line on Tim Johns' webpage at: http://web.bham.ac.uk/johnstf/ddl_lib.htm

The other important factor is the students' previous experience. A general lack of familiarity with discovery methods and a lack of knowledge of grammatical terminology would seem to constitute a major obstacle to the type of learning that concordances are designed to encourage. In the early stages at least, the tutor should ideally be present to guide and above all anticipate and challenge false lines of reasoning. One advantage of this collaborative process is that it is easier for students to ask the sort of questions that Johns (1991a and 1991b) and others see as the preferred method of working. At this point, the tutor can decide whether to encourage the student to perform his or her own searches or whether the point is of such general interest to justify the tutor making the searches and basing new materials upon these. Similarly, the tutor may use student errors as a starting point for performing searches and producing materials which can be used in a raw or processed state. Such materials have the advantage that they more clearly reflect students' needs.

The methodology described here could be criticised on the grounds that it is teacher-centred and does not necessarily make the learners autonomous. It does however require the learners to work in a focused way on examples of real usage of language drawn from authentic texts. Even if it does not meet the more optimistic hopes of certain enthusiasts of concordance based learning, it does reflect the practice described in the majority of sources cited and is seen by Leech (1997:10) as ensuring "that the maximum number of students are able and willing to participate in this kind of learning experience". It also provides committed students with a set of tools and methods which they can use as part of their personal process of language learning and thus enhances their autonomy. For the majority of students, for reasons that will be discussed in more detail in the final chapter, a teacher-centred approach would seem to be the preferred way of working with concordancers in contexts such as the one described here.

CHAPTER 6

6. CONCLUSION

6.1 INTRODUCTION

For the reasons described and discussed in the previous chapter, the outcome of the experiment left largely unproven the primary hypothesis:

working with a concordancer would make students more autonomous.

This does not mean that the opposite is true or that students should not work with a concordancer. In fact, some students did become more self-directed as a result and more questioning in their approach to language work. In general, however, it does mean that certain conditions must be met before students can work confidently with a concordancer and take greater control of their learning. These will be discussed in greater depth below.

There remains the second hypothesis:

working with a concordancer would reduce the number of errors that students make and thus improve their end of year performance.

The relatively short period of time that students spent working with the concordancer during the experimental year and the loss of a control group makes it difficult to assess with any confidence the effect, whether positive or negative, that using the concordancer may have had. Comparison with earlier and later cohorts was in theory possible but the numbers in the experimental group were relatively small and differed in composition from following year groups who had a much higher even though variable proportion of mature students. In addition, later groups were clearly more familiar with IT, having had hands-on experience of computers throughout their schooling or subsequent career, and thus more able to work independently. These factors suggested that any conclusions drawn would be problematic and open to challenge. What did come out of the experiment was that using concordances, as opposed to concordancers, with students encouraged much closer study and analysis of the facts of language. As such it chimes with Lewis' (1996:13) Observe – Hypothesize – Experiment paradigm.

The outcome of the experiment highlighted a number of issues, some of them pedagogic and some of them purely operational such as the adequacy of the hardware and software. These will be discussed in turn below before coming to any conclusion.

6.2 OPERATIONAL ISSUES

All teaching makes use of resources but teaching using CALL makes greater demands on both physical and virtual resources over which teachers often have little or no control since their availability and quality are dependent upon institutional policies. As far as the **hardware** is concerned, most institutions now have computer rooms or laboratories equipped with reasonably up to date PCs which will normally be connected to a network. The position of the facility may result in time lost between classes, especially if the room has to be cleared of students who are already encamped and need time to close their programs. Language departments are often not perceived as being IT dependent and may well find themselves low on the list of timetabling priorities. More problematic may be the operating system with institutions sometimes having a combination of several. Changes to the OS may bring about a situation where programs which have run perfectly well for years will suddenly either not work at all or will produce unexpected errors. Some programs may also require sound cards and speakers that have not been included in the hardware specification. At a more basic level, using the program effectively might simply require a functioning printer and printer connection. However efficient the support mechanism, it will rarely be able to rescue a class that has ground to a halt. The main difficulties that initially beset the experiment arose from malfunctioning printers and changes of operating system and network structure. Support was often ineffective or delayed because the management of the IT facilities was split between the faculty and the institution.

The nature and ambitiousness of the activity will therefore have to take all these factors into account and it may prove necessary to delay or abandon its implementation. Teachers may have greater control when it comes to choosing the **software**, provided it relates to CALL, but must take care to familiarise themselves not just with its workings but also with its limitations so that they do

not set themselves or their students impossible targets. The software must have been tested so that it does not unexpectedly malfunction or simply crash. It must also be intuitive and user-friendly and have support documentation, preferably in the form of a help-file accessible from the program itself. Many of these problems arose during the experiment and are in part responsible for negative reactions among the students.

Above all **students** must understand what is required of them. A surprising proportion of the students were still unfamiliar with computers (many had never even word-processed) when the experiment began and lost a lot of time looking for the files, switching between windows or entering text on screen. Some kept on “losing” the program and ended up with two or more instances of it running in parallel. Even those who were at home with computers were unfamiliar with concordancing and needed more explanation in advance. This alone, for reasons which are discussed below, would not be enough to guarantee success. Any suggestion that students are guinea pigs taking part in an experiment also arouses suspicion and fears on their part that their progress may suffer as a result of their being involved in an attempt to prove what is in effect an unproven hypothesis.

6.3 PEDAGOGICAL ISSUES

While the circumstances of the experiment cannot be seen as a fully blown self-study system, the **student attitudes** reported by Beeching (1996) and Broady (1996) parallel closely those described above and suggest that resistance to greater autonomy and independence is wide spread and deep rooted. Beeching found that “[s]tudents would like the teacher to be the fount of all knowledge and learning to be as effortless as possible”(op.cit.:92). In her research, Broady found that 83% of students surveyed agreed with the statement: “I want language classes where the teacher explains grammar and vocabulary in detail”(op.cit.:221). This might seem to be contradicted by the fact that 61% felt that a lot of grammar can be learnt without a teacher and only 30% felt that grammar has to be explained by an expert and cannot be learnt on one’s own (op.cit.:221-222). Broady suggests the following explanation: “it may be that students are distinguishing between grammar practice, often associated with gapped-sentence exercises which are straightforward to undertake independently,

and grammatical explanation, where they may feel they need detailed guidance”(op.cit.:222). Working with a concordancer would clearly fall into the latter category.

In her discussion of learners’ attitudes towards autonomy, Wenden (1991) identifies a number of highly relevant features. When our students arrive, they come with **expectations** about the learning process as well as with opinions of themselves as language learners which have been conditioned by their previous experience. Many will have acquired beliefs that “encourage dependence rather than independence”(op.cit.:55). Their lack of metacognitive knowledge may well have inculcated a “lack of willingness and self-confidence when it comes to taking on responsibility”(op.cit.:57) leading to a situation where many “regress to a state of ‘learned helplessness’, believing themselves incapable of learning without a teacher”(ibid). Confirmation of this is provided by Broady (1996) who found that “a significant proportion of students seem to lack confidence in their own learning ability, even if they recognise that self-study in language learning is important”(op.cit.:228). As a result, students “consider the ‘teacher-class’ framework is important for guiding and motivating their learning”(ibid). In Broady’s words, “[t]o weaken the classroom framework for language learning would seem to be counter productive in such a context”(ibid). Towndrow (1999) found a similar response in his students with 85% identifying the teacher as the preferred source of information when they could not understand some aspect of the CALL activity they were engaged in as opposed to 19% who felt that the computer could provide all the help needed. Those students who were unable or unwilling to experiment and solve their difficulties became frustrated and lost interest.

Hsu *et al* (1993) report on a similar experiment in which “[t]he software was intended to heighten students’ awareness of and improve their ability with elements of grammar”(op.cit.:8) using the type of exploration “that would demonstrate an active, interested, intelligent approach to the task, because it would require students to come up with multiple hypotheses about a given grammatical form in order to test them”(op.cit.:12-13), and express disappointment that none of the students observed behaved as predicted. One explanation they advance is that students are simply more product oriented

(focusing on correct forms) than experiment oriented. With CALL this often means simply clicking on the answer button (cf Kenning 1996:127). More telling, and more discouraging, is their observation that students “simply [do] not share the linguists’ interest in discovering morphosyntactic rules of language”(op.cit.:13). Stevens (1991a) identifies another negative factor, citing research in the Middle East which suggests that a “deep-seated exam fixation” (something which can be observed increasingly in the UK) can militate against “students perceiving the real purpose of the tasks they were asked to perform”(op.cit.:291).

One simple explanation for our students’ responses could be that the activity of concordancing was in itself not suitable if only because of its newness. Rubdy (1998) suggests that exploratory language learning is best developed by “exploiting **familiar pedagogic activities**”(op.cit.:2) and clearly concordancing does not fall within that category. However, at some stage students have to be introduced to new activities and methods. What matters then is the amount of preparation, a point which is also emphasised by both Littlemore (2001) and Wenden (1991). In this case, students needed to be prepared for both a new activity (concordancing) and a new mode of working (exploratory). However, it only needs one of these to fail (students’ refusal to accept the premises of exploratory or independent learning on the one hand, the occurrence of software or hardware problems on the other) for the pedagogic outcomes to be negated. Hadley (2002) too emphasises the importance of careful preparation and even reports a generally positive response to using concordances among one class of Japanese students, in spite of warnings from colleagues that the activity was unsuitable..

Kenning (1996), in addition to highlighting the importance of developed **cognitive and metacognitive skills**, states that, in order for students to benefit from using a concordancer, “[t]hey must be prepared to be proactive and able to identify fruitful lines of enquiry, to form and test hypotheses and to spot similarities and differences”(op.cit.:131). She recognises that there are huge differences between learners in the same class with some requiring a great deal of help. As a result, she warns: “Teachers should not be misled into thinking that because they can extract meaning from concordance output their students will

necessarily be able to do the same,” and goes on to recommend: “self-directed interactive use of concordancers by learners must be seen as primarily suited to advanced students with a propensity for autonomy”(ibid), a sentiment echoed by Roussel (1991:80). The fact that St.John (2001) reports a successful outcome with a beginner in German does not invalidate this recommendation. It must be remembered that St.John’s student had not only completed a year of German but was studying linguistics as his main subject. His success was only possible because he already possessed cognitive and metacognitive skills of a high order as well as a high level of motivation. Likewise, Turnbull and Burston’s (1998) experience would seem to confirm this polarisation between learners. The student who responded well and who was perceived by them as having succeeded was “familiar and comfortable with independent learning strategies which require the analysis of language data” and as a result “quickly adapted to the new inductive strategy of analysing concordance data”(op.cit.:7).

Another important element in the achievement of autonomy is **student control**. It can be argued that students are more likely to adopt the methodology espoused here if they themselves define the problem and can see how the concordancer can help in its solution. Kirk (1994) perhaps better than anyone sums up the main reason why concordancing with students may be less than successful: “Experience has shown repeatedly that concordances refuse to reveal anything of significance until the user has created a context into which, or a purpose for which, the generated information will be relevant”(op.cit.:260), a point which is echoed in Kennedy and Miceli 2001:82 and Littlemore 2001:48). For concordancing to succeed, the students must be in control not only of the software but of the investigation itself. They must be both willing and able to formulate the questions. Experience suggests that these conditions are most likely to be realised with final year dissertation students (some if not all at least) who are by and large used to working independently, and postgraduate students. They would however require some familiarisation with both the software and the methodology and this could be achieved in the second year by setting assessed concordance tasks linked to predictable errors in student writing.

It is perhaps the inability to formulate the questions rather than lack of motivation that constitutes the biggest barrier to success. It is easy to assume

from the work of Bernardini (1997), Gavioli and Zanettin (1997), Maia (1997), Zanettin (1998) that the successful implementation of the concordancer can be explained by the nationality of the students (Italian and Portuguese) and their presumed higher level of motivation and grammatical awareness compared with English native-speaker students. A more significant factor may well be the fact that the students were training to be translators rather than their nationality and as a result were more skilled in the way they approached their task. The most detailed study of the way that students approach working with a concordancer was carried out in Australia with intermediate students of Italian by Kennedy and Miceli (2001). They found five categories of problems:

Category of problem		Example
1	asking questions that are too general, regardless of context	"Can you say <i>per il corso</i> ?" rather than "Can you say <i>il orario per il corso</i> ?"
2	asking closed questions	"Do you say <i>il orario per</i> ?" rather than "What do you say after <i>orario</i> ?"
3	asking insufficiently explicit questions	when trying to say "Have a good weekend", searching on <i>auguri</i> as a translation of "have a good" and trying to link it to the idea of weekend rather than searching for expressions linked to the idea of weekend
4	treating prepositions in isolation	in the sentence <i>il cane è troppo stanco... continuare il gioco</i> asking what preposition is used before <i>continuare</i> rather than focusing on the <i>too <adjective> to do something</i> construction
5	neglecting lexical considerations in favour of grammatical ones	accepting the sentence <i>non mi sorprenderebbe imperare che ho fatto molti errori</i> without asking whether the verb <i>imperare</i> is appropriate in this context

Table 2: student approaches

Bernardini's (2000) findings are similar in many ways. Her students were given almost total autonomy, apart from sharing a source stimulus phrase, and worked enthusiastically and energetically searching on the BNC. She found that the students were good at noticing things but rarely went beyond the obvious. They failed to use whole texts to expand their findings or resolve difficulties and preferred lexical analyses to grammatical ones. They were also poor at using alternative strategies by modifying their questions or inventing new ones when their searches either produced nothing or too few examples to come to any viable

conclusion. This inability to correctly formulate the question led to the students adopting inappropriate strategies which often resulted in either false conclusions or frustration.

It is precisely this strategic competence that distinguishes St John's (2001) single student and Turnbull and Burston's (1998) successful student from the mass. Such a competence may be inherent but it may well be the product of training that focuses less on the manipulation of the software, however complex this may appear to the students, and more on formulating the right strategies and asking the right questions:

[T]he most important [conclusion from our study is] that language learners are likely to experience varying levels of success with concordancing strategies and that concordance data alone does not necessarily stimulate inductive learning strategies. A continuum of abilities and familiarity with such strategies exists, and is very much influenced by previous experience of language learning and individual learning styles (Turnbull and Burston 1998:1).

In the light of such variation in learners' ability and experience, the authors recommend much more extensive training than is often envisaged and which is not simply technical:

On the basis of these findings, we argue that extensive pre-training and guidance in concordancing strategies will be necessary for some learners, and desirable for others, if optimal outcomes from independent concordance investigations are to be achieved (ibid).

In fact, this is essentially the approach that the Italian and Portuguese examples mentioned above adopt. The focus is very much on teaching the students how to formulate the question and then implement the strategy. It could be argued that professionally orientated translation classes are a better medium for such an approach because students are more focused on the "end product" and are thus more willing to invest more in the process of achieving their goal. In this situation, the learner becomes a "research worker" in the sense that Johns (1991a:2) envisaged. Such classes are radically different from "run-of-the-mill" language classes where students' priority seems all too often to be to achieve the maximum gain for the minimum effort knowing that the end product is essentially ephemeral and will never be seen by anyone other than themselves and one or two marking tutors and who, like Turnbull and Burston's (op.cit.:8)

unsuccessful student, see “no compelling need” to improve their linguistic accuracy.

6.4 CONCLUSION

In any final evaluation of the experiment, it is useful to bear in mind Stevens’ (1991a) assertion that “[i]t is difficult to establish proof of hypotheses in qualitative studies”, especially since “such studies are often undertaken not to rigorously establish causal relationships but to gain valuable insight into the phenomena under study”(op.cit.:298). What can be concluded from the experiment? The first seems to be that for discovery methods to work, at least in the area of grammar acquisition, students need a good conceptual base to start from. Students who do not have this appear unable to benefit. In their case, careful presentation and structured practice still seem the best approach, as recommended by Manning (1996). This may in part be explained by the fact that, at least in the UK, students have little experience of formulating their own rules and tend to be passive as opposed to active learners. It is important therefore to prepare them gently and to put in place suitable learning strategies as well as continued support from teachers (Littlemore 2001).

In the light of this, what role should the concordancer play? In the first years of undergraduate study, there is a good case for arguing that the concordancer’s main role is as a tool for the teacher, allowing the testing of rule formulations and hypotheses and the isolation of authentic examples of usage for incorporation into teaching materials. This can be supplemented by requiring students to perform occasional and clearly defined searches on specific lexical items and their constructions (for example *savoir* and *connaître* or *encourager à* and *persuader de*), after a short period of familiarisation with the software. Analysis of the results should be carried out as a follow-up activity in class with the tutor present to guide the process and head off widely inaccurate interpretations. Once students are sufficiently familiar with both the software and the learning strategy, though this may not be until their final year, it should be possible to give them a number of problems and ask them to use the concordancer to search for answers. At the same time, the tutor should be prepared for a range of responses, ranging from those who cannot see the point

or for whom the result is not worth the effort, to those who are keen to take responsibility for their own language development and who are genuinely excited by what they find in the data, even though they may well still need help in interpreting what they have discovered. It is probably preferable that such activities should be an extra and not a core part of the learning experience unless the conditions are similar to those identified above in the case of advanced translation classes where students see their language development as being the key to their future career success and are willing to invest more of themselves in the acquisition process.

One area where the concordancer would seem to have a valuable part to play is in advanced translation classes, especially but not uniquely in the context of the professional training of translators. Although there do seem to be differences in motivation between students in Britain and Ireland on the one hand and students on the continent on the other (see Section 4.4.5 above), there is sufficient evidence to suggest that students are willing to accept responsibility not just to perform searches and analyse the results but also to construct their own subcorpora as part of the process. In order to convince them of the necessity of this, it might be advisable to begin by demonstrating the limitations of conventional dictionaries and reference tools. In addition, it would seem desirable to adopt a task-based approach in which the translation is conceived as an outcome which is either successful or not. One would expect this to be the case where concordancing is part of a professional training course but, in traditional language classes where translation is seen as a means to improve general language acquisition and not as an end, this might well require more rigorous marking criteria than may currently be the norm. To encourage students, it would also seem desirable to base a substantial part of the mark upon the effort that students put into corpus construction and the care with which they process and analyse the results. One important advantage of a more student-centred approach is that it avoids the “hostage to fortune” situation described in Section 6.2 above where the failure of a single item of technology can sabotage a lesson and undermine student commitment. Adequate training in the software and proper use of the Internet as well as in the methodologies would still be essential but much of this could be carried out by technicians or librarians,

leaving academic staff to concentrate in class on the linguistic issues (cf Littlemore 2001:48-49).

6.5 POSTSCRIPT

As a postscript, the author would like to cite his experience of working in a Chinese sixth-form college/senior high school (Yi Zhong Xue, Shijiazhuang, Hebei Province) in early 2004. Part of his teaching duties required working with teachers of English in the school, none of whom had ever been to an English-speaking country. Much of the teaching was based on multiple choice questions, ranging from the simple (for example *send by fax* vs *send on fax*) to the more complex (*make it easy to do something* vs *make easy to do something*). Many of these came from text-books and official worksheets. Often the answers proposed were not acceptable to the native-speaker British, American and Australian teachers employed, something that the Chinese teachers often suspected but did not have the confidence to state openly. As a result, they constantly asked the English native-speakers for their opinion, making it all but impossible for the latter to work in the staff room. Another problem that Chinese teachers had was with semantically related words such as *aims* and *goals*, *belief* and *faith*. Although this reflects the experience of any teacher of a foreign language, the problem is more acute for the Chinese. The reasons are partly cultural but more often than not linguistic and relate mainly to notions of parts of speech and especially the comparative size and nature of the vocabularies, above all in the difference between concrete and abstract (witness the frequent use of *civilisation* in translations rather than *monument*).

The teachers were introduced to *ad hoc* concordances, based upon their queries and drawn from a slightly edited version of the BNC sampler¹, in the in-service sessions. Analysis of the examples generated fruitful discussions of the semantic differences and the teachers involved clearly found the activity a valuable extension to their usual way of working. The situation described was remarkably similar to that described by Johns (see Section 4.4.1 above) where motivated learners formulate the question and the concordancer, with the guidance of the teacher, provides the data on which answers can be based. The key word is

1 All articles referring to events in Tiananmen Square and Tibet had to be removed.

“motivation” and this can be expected to be higher where the foreign language is being learnt for professional purposes. This is true of both teachers and translators. Provided that the technology functions as expected and the students are willing participants, the use of a concordancer is a useful and stimulating tool to gain insights into specific language features and the way that language functions in contemporary life.

REFERENCES AND BIBLIOGRAPHY

- Aarts Jan (1991), Intuition-based and observation-based grammars, in Aijmer & Altenberg (eds)(1991), pp.44-62
- Abraham Roberta (1985), Field Independence-Dependence and the Teaching of Grammar, *TESOL Quarterly*, 20:4, pp.689-702
- ABU (Association des Bibliophiles Universels), on-line French corpus :
<<http://abu.cnam.fr/>>
- Adamson Robin (ed)(1996), *Ça m'inspire: New Directions in French Language Studies*, Dundee: AFLS
- Ager D. E. (1986), Computer-Assisted Language Learning and Communicative Competence: a Contradiction?, in Bate & Hare (eds)(1986), pp.190-199
- Ager D.E., Knowles F.E. & Smith Joan (eds)(1979), *Advances in Computer-aided Literary and Linguistic Research*, AMLC, Birmingham: University of Aston
- Ahmad Khurshid, Corbett Greville, Rogers Margaret & Sussex Roland (1985), *Computers, Language Learning and Language Teaching*, Cambridge: Cambridge University Press
- Aijmer Karin and Altenberg Bengt (eds)(1991), *English Corpus Linguistics*, Harlow: Longman
- Akehurst F.R.P. (1981), Concordances of Troubadour Poetry in Old Occitan, in Patton and Holoien (1981), pp.155-158
- Alderson J Charles (1996), Do corpora have a role in language assessment?, in Thomas & Short (1996), pp.248-259
- Aston Guy (1995), Corpora in language pedagogy: matching theory and practice, in Cook and Seidlhofer (eds) (1995), pp.257-270
- Aston Guy (1996), The British National Corpus as a Language Resource, in Botley *et al* (eds) (1996), pp.178-191
- Aston Guy (1997a), Enriching the Learning Environment: Corpora in ELT, in Wichmann *et al* (1997), pp.51-66
- Aston Guy (1997b), Small and large corpora in language learning, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.51-62
- Aston Guy (2000), Corpora and language teaching, in Burnard & McEnery (eds)(2000), pp.7-17
- Aston Guy (ed)(2001), *Learning with Corpora*, Houston: Athelstan
- Aston Guy and Burnard Lou (1998), *The BNC Handbook*, Edinburgh: Edinburgh University Press
- Atkins Sue, Clear Jeremy & Ostler Nicholas (1992), Corpus design Criteria, *Literary and Linguistic Computing*, 7:1, 1-16
- Atlan Janet (2000), L'utilisation des stratégies d'apprentissage d'une langue dans un environnement des TICE, *ALSIC*, 3 :1, pp.109-123,
<<http://alsic.univ-fcomte.fr/Num5/atlan/default.htm>>, accessed 14 June 2000
- Backer Jimmy (1995), *Teaching Grammar with Call: Survey of Theoretical Literature*, <<http://ietn.snunit.k12.il/gramcall.html>>, accessed 22 April 99
- Baldegger Markus, Müller Martin & Schneider Günther (1980), *Kontaktschwelle*, Strasbourg: Council of Europe
- Barlow Michael (1995), Corpora for Theory and Practice, *International Journal of Corpus Linguistics*, 1:1, pp.1-37
- Barlow Michael (1996), Parallel texts in Language Teaching, in Botley *et al* (eds)(1996), pp.45-56

- Barnbrook Geoff (1996), *Language and Computers*, Edinburgh: Edinburgh University Press
- Bate Michèle and Hare Geoffrey (eds)(1986), *Communicative Approaches in French in Higher Education*, AFLS Occasional Paper 1, Association for French Language Studies, Salford: AFLS
- Baten Lut, Cornu Anne-Marie, Engels L K (1988), The Use of Concordances in Vocabulary Acquisition, in Laurén & Nordman (1988), pp.452-467
- Bax Stephen (2003), CALL-past, present and future, *System*, 31:1, pp.13-28, on-line at <[http://dx.doi.org/10.1016/S0346-251X\(02\)00071-4](http://dx.doi.org/10.1016/S0346-251X(02)00071-4)>, accessed 12 March 2003
- Beatty Ken (2003), *Teaching and Researching Computer-assisted Language Learning*, London: Longman
- Beeching Kate (1996), Evaluating a self-study system, in Broady & Kenning (eds)(1996), pp.81-104
- Beeching Kate (1997), French for Specific Purposes: The Case for Spoken Corpora, *Applied Linguistics*, 18:3, pp.374-394
- Bernardini Sylvia (1997), A 'trainee' translator's perspective on corpora, paper given at the conference on *Corpus use and learning to translate*, Bertinoro, 14-15 November 1997, on-line at <<http://www.sslmit.unibo.it/cultpaps/trainee.htm>>, 12 pages, accessed 11 December 2003
- Bernardini Sylvia (2000), Systematising serendipity: Proposals for concordancing large corpora with language learners, in Burnard & McEnery (eds)(2000), pp.225-234
- Bernardini Sylvia (2003), Designing a Corpus for Translation and Language Teaching: The CEXI Experience, *TESOL Quarterly*, 37:3, pp.528-537
- Biber Douglas (1993a), Representativeness in Corpus Design, *Literary and Linguistic Computing*, 8:4, pp.243-257
- Biber Douglas (1993b), Using Register-Diversified Corpora for General Language Studies, *Computational Linguistics*, 19:2, pp.219-241
- Biber Douglas (1995), *Dimensions of register variation: A cross-linguistic comparison*, Cambridge: Cambridge University Press
- Biber Douglas, Conrad Susan & Reppen Randi (1994), Corpus-based approaches to Issues in Applied Linguistics, *Applied Linguistics*, 15:2, pp.169-189
- Biber Douglas, Conrad Susan & Reppen Randi (1998), *Corpus Linguistics*, Cambridge: Cambridge University Press
- Blanche-Benveniste Claire (1996), De l'utilité du corpus linguistique, *Revue française de linguistique appliquée*, 1:2, pp.25-42
- Blanchet Philippe (n.d.), Une méthode audio-visuelle: Voix et images de France, *Les Amphis de France* 5, on-line at <http://www.canalu.fr/canalu/affiche_programme.php?programme_id=1865002560&vHtml=0>, accessed 27 July 2004
- Block David (2002), Communicative language teaching revisited : discourses in conflict and foreign national teachers, *Language Learning Journal*, 26, pp.19-26
- Bloomfield Leonard (1933), *Language*, New York: Holt
- BNC (British national Corpus), for more information see <<http://info.ox.ac.uk/bnc>>
- BNC Sampler (1999), release 1.1, 2,000,000 word sample of the British National Corpus, distributed by the Oxford University Humanities Computing Unit, see <<http://info.ox.ac.uk/bnc>> for details
- Botley Simon, Glass Julia, McEnery Tony and Wilson Andrew (eds)(1996), *Proceedings of Teaching and Language Corpora 1996*, Lancaster: UCREL
- Bowden Paul and Edwards Mark (1996), Knowledge Extraction from Corpora for

- Pedagogical Applications, in Botley *et al* (eds)(1996), pp.160-170
- Bowker Lynne (1999), Exploring the Potential of Corpora for Raising Language Awareness in Student Translators, *Language Awareness*, 8:3/4, pp.160-173
- Bowker Lynne (2002), *Computer-Aided Translation Technology*, Ottawa: Ottawa University Press
- Bowker Lynne & Pearson Jennifer (2002), *Working with Specialised Language: A practical guide to using corpora*, London and New York: Routledge
- Bowker Lynne, Cronin Michael, Kenny Dorothy & Pearson Jennifer (eds)(1998), *Unity in Diversity? Current Trends in Translation Studies*, Manchester: St Jerome
- Brett Paul & Motteram Gary (eds) (2000), *A Special Interest in Computers: Learning and Teaching with Information and Communications Technologies*, Whitstable: IATEFL
- Broady Elspeth (1996), Learner Attitudes towards Self-Direction, in Broady & Kenning (eds)(1996), pp.215-235
- Broady Elspeth & Kenning Marie-Madeleine (eds)(1996), *Promoting Learner Autonomy in University Language Teaching*, London: AFLS/CILT
- Brown Douglas (1987), *Principles of Language Learning and Teaching*, Englewood Cliffs: Prentice Hall
- Brown Douglas & Gonzo Susan (eds)(1995), *Readings on second language acquisition*, Englewood Cliffs: Prentice Hall Regents
- Brown Eric (1988), No call for CALL? 20th century technology for 19th century language teaching methodology?, in Phillips (ed)(1988), pp.86-96
- Brumfit Christopher (ed)(1983), *Learning and teaching languages for communication: applied linguistic perspectives*, London: CILT
- Brumfit C.J.K & Johnson K. (eds) (1979), *The Communicative Approach to Language Teaching* Oxford: Oxford University Press
- Brumfit Christopher, Phillips Martin & Skehan Peter (eds)(1985), *Computers in English Language Teaching*, Oxford: Pergamon
- Burnard Lou & McEnery Tony (eds)(2000), *Rethinking Language Pedagogy from a Corpus Perspective*, Frankfurt am Main: Peter Lang Verlag
- Burstall C., Jamieson M., Cohen S. & Hargreaves M.(1974), *Primary French in the Balance*, Slough: NFER
- Burton Dolores M. (1981), Automated Concordances and Word Indexes: the Process, the Programs, and the Products, *Computers and the Humanities*, 15:3, pp.139-154
- Butler C. (1987), Computational Text Analysis: a Survey in relation to CALL, in Chesters & Gardner (1987), pp.67-79
- Butler Jonathan (1990), Concordancing, Teaching and Error Analysis: the advantages of discourse level authenticity in testing expectancy grammar, *System*, 18:3, pp.343-349
- Butler Jonathan (1991), Cloze procedures and concordances, *System*, 19:1/2, pp.29-38
- Cameron Keith (ed)(1989), *Computer Assisted Language Learning*, Oxford: Intellect
- Cameron Keith (1996), Computer Assisted Language Learning, *AIS*,
<http://www.dcs.exeter.ac.uk/~masoud/yazdani/editor/ais/ais-1/ais-1-4.htm>,
 accessed 16 March 98
- Cameron Keith (ed), (1999), *CALL and the Learning Community*, Exeter: Elm Bank Publications
- Canale A. & Swain M. (1980), Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics*, 1, pp.1-47
- Carroll J.B.(1964), Words, meanings and concepts, *Harvard Educational Review*, 34, pp.178-202

- Carroll J.B.(1996), The contributions of psychological theory and educational research to the teaching of foreign languages, in Valdman A.(ed)(1996), pp.93-106
- Cavaglià Gabriela & Kilgariff Adam (2001), Corpora from the Web, *Proceedings of the Fourth VLUK Colloquium*, Sheffield, January 2001, pp.120-124, on-line at <http://ftp.itri.bton.ac.uk/reports/ITRI-01-06.pdf>, 5 pages, accessed 25 January 2004
- Celce-Murcia Marianne (1991), Grammar Pedagogy in Second and Foreign language Teaching, *TESOL Quarterly*, 25:3, pp.459-480
- Chambers Angela & Davies Graham (eds)(2001), *ICT and Language Learning: A European Perspective*, Lisse: Swets & Zeitlinger
- Chapelle Carol (1997), CALL in the year 2000: still in search of research paradigms?, *Language Learning and Technology Journal*, 1:1, pp.19-43, <http://lt.msu.edu/vol1num1/chapelle/default.html> >, accessed 7 January 2003
- Chapelle Carol & Jamieson Joan (1986), Computer-Assisted Language Learning as a Predictor of Success in Acquiring English as a Second Language, *TESOL Quarterly*, 20:1, pp.27-46
- Chesters Graham & Gardner Nigel (eds) (1987), *Use of Computers in the Teaching of Language and Languages*, Bath: CITS
- Ching Lai Phooi & Wong Irene, (1994), Using Concordance-based Material for Teaching Verb Inflections, *TESL Reporter*, vol.27, no.2, pp.50-54
- Chomsky Noam (1957), *Syntactic Structures*, The Hague: Mouton
- Chomsky Noam (1959), review of 'Verbal behavior' by B.F.Skinner, *Language*, 35:1, pp.26-58
- CIEL (2001), *Educational Copyright in the context of language teaching and the provision of independent language learning materials*, 3, Curriculum and Independence for the Learner Project funded by HEFCE under the FDTL and TLTP programmes, on-line at <http://ciel.lang.soton.ac.uk/copyright.htm> >, accessed 26 January 2004
- Cobb Tom (1997), Is there any measurable learning from hands-on concordancing?, *System*, 25:3, pp.301-315
- Cobb Tom (2001), An Introduction to Learner Corpus Analysis, revised version of paper presented at AAAL 2000, on-line at <http://www.er.uqam.ca/nobel/r21270/cv/LC.htm> >, accessed 28 June 2002
- Cobb Tom, Greaves Chris & Horst Marlise (2001), Can the rate of lexical acquisition from reading be increased? An experiment in reading French with a suite of on-line resources, <http://vlc.polyu.edu.hk/StudyGuide/rlearn/lexicalacquisition.htm> >, 11 pages, accessed 05 May 01
- Coffey Stephen (2002), Using a Source Language Corpus in Translator Training, *inTRAlinea*, 5, Special Issue, on-line at <http://www.intralinea.it/vol5/cult2k/coffey.htm> >, 11 pages, accessed 22 January 2004
- Coleman James & Towell Richard (eds)(1987), *The Advanced Language Learner*, London: SUFLRA / CILT
- Collins COBUILD Concordance Samplers (1994, 1995), London: Harper-Collins
- Cook Guy & Seidlhofer Barbara (eds) (1995), *Principle and Practice in Applied Linguistics*, Oxford: Oxford University press
- Cook V.J.(1985), Bridging the Gap between Computers and Language Teaching, in Brumfit, Phillips & Skehan (eds)(1985), pp.13-24
- Corder S. Pit. (1973), *Introducing Applied Linguistics*, London: Penguin
- Cornell Alan (1996), Grammar – grinding or grounding?, *German Teaching*, 13, pp.26-

- Corness Patrick, Courtney Kathy & Matthews Sarah (1997), Astcovea: grammar in context, *ReCALL*, 9:2, pp.33-42
- Corréard Marie-Hélène & Grundy Valerie (1994), *The Oxford-Hachette French Dictionary*, Oxford : Oxford University Press
- Coste D., Courtillon J., Ferenczi V., Martins-Baltar M. & Papo E. (1976), *Un niveau-seuil*, Strasbourg: Council of Europe
- Crosland A.T. (1975), The Concordance and the Study of the Novel, *ALLC Bulletin*, 1975 3:3, pp.190-196
- Davidson T.T.L. (1990), Teaching with the Oxford Concordancing Program, *Literary and Linguistic Computing*, 5:1, pp.81-85
- Davies Graham (1985), *Talking BASIC: an introduction to BASIC programming for users of language*, Eastbourne: Cassell
- Davies Graham (1997), Lessons from the past, lessons for the future: 20 years of CALL, in Korsvold and Rüschhoff (eds)(1997), pp.27-51
- Davies Graham (1999), Introduction to Computer Assisted Language Learning, *ICT4LT Module 1.4*, on-line at <http://www.ict4lt.org/en/en_mod1-4.htm>, 33 pages, accessed on 16 February 2001
- Davies Graham (2001), update of Davies (1997) on-line at <<http://www.camsoftpartners.co.uk/coegdd1.htm>>, 20 pages, accessed 07 October 2001
- Davies Graham & Higgins John (1982), *Computers, language and language learning*, Information Guide 22, London: Centre for Information on Language Teaching and Research (CILT)
- Davies Graham & Higgins John (1985), *Using computers in language learning: a teacher's guide*, Information Guide 22, London: Centre for Information on Language Teaching and Research (CILT) (update of Davies & Higgins 1982)
- Davies Mark (2000), Using multi-million word corpora of historical and dialectal Spanish texts to teach advanced courses in Spanish linguistics, in Burnard & McEnery (eds)(2000), pp.173-185
- De Vive voix* (1972), Crédif, Paris: Didier-Hatier
- Decoo Wilfried (1996), The Induction-Deduction Opposition: Ambiguities and Complexities of the Didactic Reality, *IRAL*, XXXIV:2, pp.95-118
- Delcloque Philippe (2000), *History of CALL: the History of Computer Assisted Language Learning Web Exhibition*, <<http://www.history-of-call.org/>>, accessed 07 November 01
- DES (1991), *Modern Foreign Languages in the National Curriculum*, London: HMSO
- Dodd B. (1997) Exploiting a Corpus of Written German for Advanced language Learning, In Wichmann *et al* (1997), pp.131-145
- Dossena Marina (1996), Evaluating corpora: are we asking the right questions?, in Botley *et al* (eds)(1996), pp.244-253
- Drew Fiona & Ottewill Roger (2002), Learning styles and the potential for learning on institution-wide language programmes: an assessment of the results of a pilot study, *Language Learning Journal*, 26, pp.11-18
- Dulay Heidi, Burt Marina & Krashen Stephen (1982), *Language Two*, New York: Oxford University Press
- Dunkel Patricia (ed) (1991) *Computer-Assisted Language Learning and Research*, New York: Newbury House
- Dunkel Patricia (1991), The Effectiveness Research on Computer-Assisted Instruction

- and Computer-Assisted Language Learning, in Dunkel (ed) (1991), pp.5-36
- Dunning Roy (ed)(1983), *French for Communication: The East Midlands Graded Assessment Feasibility Study*, Leicester: University of Leicester
- Ellis Nick (1993), Rules and Instances in Foreign Language Learning: Interactions of Explicit and Implicit Knowledge, *European Journal of Cognitive Psychology*, 5:3, pp.289-318
- Ellis Rod (1986), *Understanding Second Language Acquisition*, Oxford: Oxford University Press
- Ellis Rod (1987), Does Remedial Instruction Work? A L2 Acquisition Perspective, in Coleman & Towell (eds)(1987), pp.55-73
- En avant: stage 1* (c1964), Leeds, England: E.J. Arnold for the Nuffield Foundation
- En avant: stage 2* (c1965), Leeds, England: E.J. Arnold for the Nuffield Foundation
- Engel Dulcie and Myles Florence (eds)(1996a), *Teaching Grammar: Perspective in Higher Education*, London: AFLS and CILT
- Engel Dulcie and Myles Florence (1996b), Grammar Teaching: The Major Concerns, in Engel & Myles (eds)(1996a), pp.9-19
- Evans Joan (1992), Testing the effectiveness of the computer in promoting communication, *ON-CALL*, 7:1, 5 pages, on-line at <http://www.cltr.uq.edu.au/oncall/71evan.htm>, accessed 7 July 1999
- Facques Bénédicte (2002), Passé composé, imparfait et présent dans les récits journalistiques: des alternances aux ruptures temporelles, in Labeau and Larrivée (eds)(2002), pp.105-133
- Færch Claus & Kasper Gabriele (1984), Pragmatic Knowledge: Rules and Procedures, *Applied Linguistics*, 5:3, pp.214-225
- Færch Claus & Kasper Gabriele (1987), *Introspection in Second Language Research*, Clevedon: Multilingual Matters
- Farrington Brian (1996), Data-Driven Learning: a new horizon for CALL, in Adamson (ed) (1996), pp.177-191
- Fay M. and Ferney D. (eds) (2000), *Current trends in Modern Languages Provision for Non-Specialist Linguists*, London: CILT
- Finocchiaro Mary & Brumfit Christopher (1983), *The Functional-Notional Approach*, Oxford: Oxford University Press
- Fletcher William (2002), Making the Web More Useful as a Source for Linguistic corpora, paper given at the Fourth North American Symposium on Corpus Linguistics in Indianapolis on 1-3 November 2002, 10 pages, on-line at <http://kwicfinder.com/AAACL2002whf.pdf>, accessed 25 January 2004
- Fletcher William (2003a), Concordancing the Web with KwiCFinder, revised version of paper given at the Third North American Symposium on Corpus Linguistics in Boston MA on 23-25 March 2001, 16 pages, on-line at <http://kwicfinder.com/FletcherCLLT2001.pdf>, accessed 25 January 2004
- Fletcher William (2003b), Facilitating the compilation and dissemination of ad-hoc web corpora, revised version of paper given at the Fifth International Conference on Teaching and Language Corpora in Bertinoro, Italy, on 27-31 July 2003, 13 pages, on-line at http://kwicfinder.com/Facilitating_Compilation_and_Dissemination_of_Ad-Hoc_Web_Corpora.pdf, accessed 25 January 2004
- Fligelstone Steve, (1993), Some reflections on the question of teaching from a corpus linguistics perspective, *ICAME Journal*, 17, pp.97-109
- Flowerdew John, (1993a), Concordancing as a tool in course design, *System*, 21:2,

- pp.231-244
- Flowerdew John, (1993b), Concordancing in Language Learning, *Perspectives*, 5:2, pp.87-100, also in Pennington (1996), pp.99-113
- Fotos S. (ed.)(1996), *Multimedia Language Teaching*, Tokyo: Logos
- Foucou Pierre-Yves and Kübler Natalie (2000), A web-based environment for teaching technical English, in Burnard & McEnery (eds)(2000), pp.65-73
- Fox Gwyneth (1998), Using corpus data in the classroom, in Tomlinson (ed)(1998), pp.25-43
- Fox Jeremy, (1989), Can Computers Aid Vocabulary Learning?, in Cameron (1989), pp.1-13
- Fox Jeremy, Matthews Anne, Matthews Clive & Rope Arthur (1990), *Educational technology in Modern language Learning in the secondary, tertiary and vocational sectors*, A report for the Training Agency by the University of East Anglia and the Bell Educational Trust, Sheffield: Training Agency
- FRANTEXT, French corpus developed by INaLF (Institut National de la Langue Française) and accessible to subscribers at the following address:
<<http://jupiter.inalf.cnrs.fr/noncateg.htm>>
- Freudenstein Reinhold & Vaughan James C. (eds), *Confidence through Competence in Modern Language Learning*, London: Centre for Information on Language Teaching and Research (CILT)
- Friedbichler Ingrid & Friedbichler Michael (1997), The Potential of Domain-Specific Target-Language Corpora for the Translator's Workbench, paper given at the conference on *Corpus use and learning to translate*, Bertinoro, 14-15 November 1997, on-line at <<http://www.sslmit.unibo.it/cultpaps/fried.htm>>, 5 pages, accessed 11 December 2003
- Fulcher Glenn (2000), Computers in language testing, in Brett & Motteram (eds)(2000), pp. 93-107
- Gardner R C (1985), *Social Psychology and Second language Learning*, London: Edward Arnold
- Garside Roger, Leech Geoffrey & McEnery Anthony (eds)(1997), *Corpus Annotation*, London: Longman
- Garton James (1996), Interactive concordancing with a specialist corpus, *ON-CALL*, 10:1, 7 pages, on-line at <<http://www.cltr.uq.au/oncall/garton101.html>>, accessed 11 December 2003
- Gavioli Laura (1997), Exploring Texts through the Concordancer: Guiding the Learner, in Wichmann *et al*(1997), pp.83-99
- Gavioli Laura & Aston Guy (2001), Enriching reality: corpora in language pedagogy, *ELT Journal*, 55:3, pp.238-246
- Gavioli Laura & Zanettin Federico (1997), Comparable corpora and translation: a pedagogic perspective, paper given at the conference on *Corpus use and learning to translate*, Bertinoro, 14-15 November 1997, on-line at <<http://www.sslmit.unibo.it/cultpaps/laura-fede.htm>>, 6 pages, accessed 11 December 2003
- Ghadessy Mohsen, Henry Alex & Roseberry Robert (eds) (2001), *Small Corpus Studies and ELT*, Amsterdam & Philadelphia: John Benjamins
- Gledhill Chris (1996), Science as a collocation. Phraseology in cancer research articles, in Botley *et al* (eds)(1996), pp.108-126
- Goodfellow Robin (1994), Design principles for computer-aided vocabulary learning, in Thompson & Chesters (1994), pp.53-62

- Goodfellow R. & Metcalfe P. (1997), The challenge - back to basics or brave new world?, *ReCALL* 9:2, pp.4-7
- Gordon K.L & Downes P.J. (1978), *New Objectives in Modern Language Teaching*, Oxford: Oxfordshire Modern Languages Advisory Committee
- Granger Sylviane (1994), The Learner Corpus: a revolution in applied linguistics, *English Today* 39, 10:3, pp.25-29
- Granger Sylviane, Hung Joseph & Petch-Tyson Stephanie (eds)(2001), *Computer Learner Corpora, Second language Acquisition and Foreign Language Teaching*, Amsterdam & Philadelphia: John Benjamins
- Grévisse Maurice (1955), *Le Bon Usage* (5^e edition), Gembloux: Duculot
- Grishman Ralph (1986), *Computational Linguistics*, Cambridge: Cambridge University Press
- Groot Peter (2000), Computer Assisted Second Language Vocabulary Acquisition, *Language Learning and Technology*, 4:1, pp.60-81, also available on-line at <http://llt.msu.edu/vol4num1/groot/default.html>, accessed 09 June 00
- Grosser W, Hogg J & Hubmayer K (eds) (1995), *Style: Literary and Non-Literary: Contemporary Trends in Cultural Stylistics*, New York: Edwin Mellen Press
- Grüner Margit & Hassert Timm (1991), *Computer im Unterricht*, Munich: Goethe-Institut
- Guillemin-Flescher Jacqueline (1996), La traduction humaine : contraintes et corpus, *Revue française de linguistique appliquée*, 1:2, pp.43-56
- Gumperz John & Hymes Dell (eds)(1970), *Directions in Sociolinguistics*, New York: Holt, Rinehard and Winston
- Habert Benoît, Nazarenko Adeline and Salem André (1997), *Les linguistiques de corpus*, Paris : Armand Colin
- Hadley Gregory, (no date), *Sensing the Winds of Change: An Introduction to Data-driven Learning*, <http://web.bham.ac.uk/johnstf/winds.htm>, no date, 16 pages, accessed 07 July 1999
- Hadley Gregory (2002), An Introduction to Data-Driven Learning, *RELC Journal*, 33:2, pp.99-124
- Hägg Åke (1986), Micro-computers in language teaching, in Freudenstein & Vaughan James (1986), pp.146-153
- Halliday M.A.K. (1973), *Explorations in the Functions of Language*, London: Edward Arnold
- Hanson-Smith Elizabeth (1993), Dancing with Concordances, *CAELL Journal*, 4:2, p.40
- Harding Ann, Page Brian & Rowell Sheila (1980), *Graded Objectives in Modern Languages*, London: CILT
- Hassert Timm (1990), *Software für den computer-unterstützten DaF-Unterricht*, Munich: Goethe-Institut
- Hawkins Eric (1994), Percept before precept, in King & Boaks (eds) (1994), pp.109-123
- Hawkins Roger, Lamy Marie-Noëlle & Towell Richard (1997), *Practising French Grammar*, London: Arnold
- Hawkins Roger & Towell Richard (1992), Second language acquisition research and the second-language acquisition of French, *French Language Studies*, 2, pp.97-121
- Hawkins Roger & Towell Richard (1996), Why Teach Grammar?, in Engel & Myles (eds)(1996a), pp.195-211
- Heafford Michael (1993), What is grammar, who is she?, *Language Learning Journal*, 7, pp.55-58
- Hegelheimer Volker & Chapelle Carol (200), Methodological Issues in Research on

- Learner-Computer Interactions in CALL, *Language Learning and Technology*, 4:1, pp.41-59, on-line at <<http://llt.msu.edu/vol4num1/hegchap/default.htm>>, 22 pages, accessed 09 June 2000
- Hewer Sue (1998), *The TELL Consortium Summative Evaluation Report*, Hull: University of Hull
- Hewer Sue, Rendall Heather, Walker Ros & Davies Graham (1999), *Introduction to Computer Assisted Language Learning*, ICT4LT Module 1.4, <http://www.ict4lt.org/en/en_mod1-4.htm>, 33 pages, accessed 16 February 01
- Higgins John (1985), Should Teachers Learn to Program?, in Brumfit *et al* (eds)(1985), pp.69-77
- Higgins John (1986), The computer and grammar teaching, in Leech & Candlin (1986), pp.31-45
- Higgins John (1988), *Language, Learners and Computers*, Harlow: Longman
- Higgins John (1991), Fuel for learning: the neglected element of textbooks and CALL, *CALL Journal*, 2:2, pp.3-7
- Higgins John and Johns Tim, (1984), *Computers in Language Learning*, London: Collins
- Hockey Susan (1998), Textual databases, in Lawler & Dry (eds)(1998), pp.101-137
- Hoey Michael (1997), From concordance to text structure: new uses of computer corpora, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.2-23
- Honeyfield John (1989), A Typology of Exercises Based on Computer-Generated Concordances, *Guidelines*, 11:1, pp.42-50
- Hornsey Alan (ed)(1975), *Handbook for Modern language Teachers*, London: Methuen
- Howatt Anthony (1972), *Programmed Learning and the Language Teacher*, Harlow: Longman
- Hsu Jing-Fong, Chapelle Carol & Thompson Ann (1993), Exploratory Learning Environments: what are they and do students explore?, *Educational Computing Research*, 9:1, pp.1-15
- Hubbard Philip (1992), A Methodological Framework for CALL Courseware Development, in Pennington & Stevens (eds) (1992), pp.39-65
- Hubbard Philip (1996), Elements of CALL Methodology: Development, Evaluation and Implementation, in Pennington (ed).(1996), pp.15-32
- Huot Hélène (1998), Le rôle des corpus dans la recherche linguistique, *Cahiers AFLS*, 4 :1, pp.5-14
- Hurd Stella (2000), Helping learners to help themselves: the role of metacognitive skills and strategies in independent language learning, in Fay & Ferney (eds) (2000), pp.36-52
- Hymes Dell (1964), Toward Ethnographies of Communication, *American Anthropologist*, 66:6:2, pp.12-25
- Hymes Dell (1971), *On Communicative Competence*, Philadelphia: University of Pennsylvania Press
- Hymes Dell (1972), On Communicative Competence, in Pride & Holmes (eds)(1972), pp.269-293
- Ide Nancy, Reppen Randi & Suderman Keith (2002), The American national Corpus: More Than the Web Can Provide, *Proceedings of the Third Language Resources and Evaluation Conference (LREC)*, Las Palmas, Spain, pp. 839-844, on-line at <<http://www.cs.Vassar.edu/~ide/papers/anc-lrec02.pdf>>, 6 pages, assessed 25 January 2004
- Ilse W-R. (1991), Concordancing in vocational training, in Johns & King (1991),

- pp.103-113
- Incorporated Association of Assistant Masters, 1967, *The Teaching of Modern Languages* (4th edition), London: University of London Press
- Inkster Gordon (1997), First Catch your Corpus: Building a French Undergraduate Corpus from Readily Available Textual Resources, in Wichmann *et al* (eds)(1997), pp.267-276
- Jackson H (1990), OCP and the Computer Analysis of Texts: the Birmingham Polytechnic Experience, *Literary and Linguistic Computing*, 5:1, pp.86-88
- Jackson Howard (1997), Corpus and Concordance: Finding out about Style, in Wichmann *et al* (1997), pp.224-239
- Jamieson Joan (1995), The cognitive styles of reflection/impulsivity and field independence/ dependence and ESL success, in Brown & Gonzo (eds)(1995), pp.119-137
- Johansson Stig (1997), Using the English-Norwegian parallel corpus – a corpus for contrastive analysis and translation studies, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.282-296
- Johns Tim (1985), Micro-Concord, Mimeo, Dept. of English, University of Birmingham
- Johns Tim (1986), Micro-Concord: a Language Learner's Resource Tool, *System*, 14:2, pp.151-162
- Johns Tim (1991a), Should you be Persuaded: Two Examples of Data-driven Learning, in Johns & King (1991), pp.1-13
- Johns Tim (1991b), From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning, in Johns & King (1991), pp.27-45
- Johns Tim (1996), If our Descriptions of Language are to be Accurate . . . A footnote to Kettemann, <<http://web.bham.ac.uk/johnstf/ifbeto.htm>>, 3 pages, accessed 31 August 2001
- Johns Tim (1997), Contexts: the Background, Development and Trialling of a Concordance-based CALL Program, in Wichmann *et al* (1997), pp.100-115
- Johns Tim & King Philip (eds)(1991), *Classroom Concordancing: ELR Journal vol.4*, Birmingham: University of Birmingham
- Johnson Donna M. (1992), *Approaches to Research in Second Language Learning*, New York: Longman
- Johnson Keith (1982), *Communicative Syllabus Design and Methodology*, Oxford: Pergamon
- Johnson Keith & Morrow Keith (eds) (1981), *Communication in the Classroom*, Harlow: Longman
- Johnstone Richard (1994), Grammar: acquisition and use, in King & Boaks (eds)(1994), pp.9-13
- Johnstone Richard (1997), Research on language learning and teaching: 1996, *Language Teaching*, 30:3, pp. 149-165
- Jones Randall L.(1997), Creating and Using a Corpus of Spoken German, in Wichmann *et al* (1997), pp.146-156
- Karttunen Frances (1986), A linguist looks at computer-assisted instruction, in Freudenstein & Vaughan James (1986), pp.133-145
- Kemmis Stephen, Atkin R. & Wright E. (1977), *How do students learn?*, Working paper on computer-assisted learning: UNCAL evaluation studies, Centre for Applied Research in Education, Occasional Publications No. 5, Norwich
- Kennedy Claire & Miceli Tiziana (2001), An evaluation of intermediate students' approaches to corpus investigation, *Language Learning and Technology*, 5:3, pp.77-

- 90, on-line at <<http://lt.msu.edu/vol5num3/kennedymiceli/>>, accessed 7 January 2003
- Kennedy Graeme (1998), *An Introduction to Corpus Linguistics*, Harlow: Longman
- Kenning Marie-Madeleine (1996), IT and Autonomy, in Broady & Kenning (eds)(1996), pp.121-138
- Kenning M.J & Kenning M-M. (1983), *Introduction to Computer Assisted Language Teaching*, Oxford: Oxford University Press
- Kenning M-M & Kenning M.J. (1990), *Computers and Language Learning*, London: Ellis Horwood
- Kenny Dorothy (1997), (Ab)normal translations: a German-English parallel corpus for investigating normalization in translation, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.387-392
- Kettemann Bernhard (1995a), Concordancing in stylistics teaching, in Grosser *et al* (eds)(1995), pp.307-318
- Kettemann Bernhard (1995b), On the Use of Concordancing in ELT, *TELL & CALL*, 1995.4, pp.4-15, on line at <<http://gewi.kfunigraz.ac.at/~kettemann/conco.html>>, 14 pages, accessed 07 July 99
- Kettemann Bernhard (1996), Concordancing in English Language Teaching, in Botley *et al* (1996), pp.4-16
- Kettemann Bernhard (1997), Concordancing as input enhancement in ELT, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.63-73
- Kilgariff Adam (2001a), Comparing Corpora, *International Journal of Corpus Linguistics*, 6:1, pp.1-37 on-line at <<ftp://ftp.itri.bton.ac.uk/reports/ITRI-01-15.pdf>>, 29 pages, accessed 24 January 2004
- Kilgariff Adam (2001b), Web as corpus, *Proceedings of the Corpus Linguistics 2001 Conference, UCREL Technical papers :13*, Lancaster: Lancaster University, pp.342-344, on-line at <<http://www.itri.bton.ac.uk/~Adam.Kilgariff/PAPERS/corpling.txt>>, 4 pages, accessed 25 January 2004
- King Lid (1994), Mad slaves and chocolate ice-cream, in King & Boaks (1994), pp.1-6
- King Lid & Boaks Peter (eds) (1994), *Grammar! A conference report*, London: CILT
- King Philip and Woolls David (1996), Creating and Using a multilingual parallel concordancer, *Translation and Meaning*, 4, pp. 459-466, available on-line at <<http://web.bham.ac.uk/johnstf/paracon.htm>>, 7 pages, accessed 07 July 1999
- Kirk John (1994), Corpus-Concordance-database-VARBRUL, *Literary and Linguistic Computing*, 9:4, pp.259-266
- Klapper John (1997), Language learning at school and university: the great grammar debate continues (I), *Language Learning Journal*, 16, pp.22-27
- Klapper John (1998), Language learning at school and university: the great grammar debate continues (II), *Language Learning Journal*, 18, pp.22-28
- Klapper John & Rees Jonathan (2003), Reviewing the case for explicit grammar instruction in the university foreign language learning context, *Language Teaching Research*, 7:3, pp.285-314
- Knowles Frank (1990), Language and IT: Rivals or Partners, *Literary and Linguistic Computing*, 5:1, pp.38-44
- Knowles Gerry (1990), The Use of Spoken and Written Corpora in the Teaching of Language and Linguistics, *Literary and Linguistic Computing*, 5:1, pp.45-48
- Knowles Gerry (1997), Using Corpora for the Diachronic Study of English, in Wichmann *et al* (1997), pp.195-210
- Korsvold A-K and Rüschoff B.(eds)(1997), *New technologies in language learning and*

- teaching, Strasbourg: Council of Europe
- Koubourlis Demetrius (1975), On Concordances and Their Uses, *Slavic and East European Journal*, 19:2, pp.246-253
- Kowitz Johanna & Carroll David (1991), Using Computer Concordances for Literary Analysis, in Johns & King (1991), pp.135-149
- Krashen, S.D. (1981), *Second Language Acquisition and Second Language Learning*, Oxford: Pergamon
- Krashen S. (1982) *Principles and Practice in Second Language Acquisition*, Oxford: Pergamon
- Krashen S.D.(1985), *The Input Hypothesis: Issues and Implications*, New York: Longman
- Krashen S.D. (1989), We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis, *The Modern Language Journal*, 73, pp.440-464
- Krashen Stephen & Terrell Tracy (1983), *The natural Approach: Language Acquisition in the Classroom*, Oxford: Pergamon
- Labeau Emmanuelle and Larrivée Pierre(eds)(2002), *Les temps du passé français et leur enseignement*, Amsterdam & New York: Rodopi
- Lamy Marie-Noëlle (1999), Analyse de 'Concordances in The Classroom : a resource book for teachers', *Apprentissage des Langues et Systèmes d'Information et de Communication (ALSIC)*, 2 :2, pp.33-40, < <http://alsic.univ-fcomte.fr/Num4/lamy/default.htm> >, accessed 20 June 2000
- Lamy Marie-Noëlle & Goodfellow Robin (1999), 'Reflective conversation' in the virtual language classroom, *Language Learning and Technology Journal*, 2 :2, pp.43-61 < <http://llt.msu.edu/vol2num2/article2/index.html> >, accessed 07 January 2003 (French version available at < <http://alsic.u-strasbg/Num2/lamy/default.htm> >)
- Lamy Marie-Noëlle & Mortensen Hans (2000), *Using concordance programs in the modern foreign languages classroom*, ICT4LT Module 2.4, <http://www.ict4lt.org/en/en_mod2-4.htm>, 48 pages, accessed 10 July 01
- Larsen-Freeman & Long Michael (1991), *An Introduction to Second Language Acquisition Research*, London: Longman
- Last Rex (1984), *Language Teaching and the Microcomputer*, Oxford: Blackwell
- Last Rex (1987), Artificial Intelligence - the Way forward for CALL?, in Chesters & Gardner (1987), pp.61-66
- Last Rex (1992), Computers and Language Learning: Past, Present – and Future?, in Butler (ed)(1992), pp.227-246
- Laurén Christer & Nordman Marianne (eds) (1988), *Special Language: from Humans Thinking to Thinking Machines*, Clevedon: Multilingual Matters
- Laviosa Sara (1998), The English Comparable Corpus, in Bowker, Cronin, Kenny & Pearson (eds)(1998), pp101-112
- Lawler John (1998), The Unix language family, in Lawler & Dry (eds)(1998), pp.138-169
- Lawler John & Dry Helen Aristar (eds)(1998), *Using Computers in Linguistics*, London and New York: Routledge
- Leech Geoffrey (1997), Teaching and Language Corpora: a Convergence, in Wichmann *et al* (1997), pp.1-23
- Leech G. & Candlin C. (eds) (1986), *Computers in English Language Teaching and Research*, Harlow: Longman
- Levy Michael (1990), Concordances and their Integration into a Word-processing Environment for Language Learners, *System*, 18:2, pp.177-188

- Levy Michael (1997), *CALL: context and conceptualisation*, Oxford: Oxford University Press
- Levy Michael (1998), Two conceptions of language learning and their implications for CALL at the tertiary level, *ReCall*, 10:1, pp.86-94
- Lewandowska-Tomaszczyk Barbara & Melia Patrick (eds)(1997), *Practical Applications in Language Corpora*, Lodz: Lodz University Press
- Lewis Michael (1996), Implications of a lexical view of language, in Willis & Willis (eds)(1996), pp.10-16
- Lian Andrew (1992), Intelligence in Computer-Aided Language Learning, in Pennington & Stevens (eds)(1992), pp.66-76
- Lightbown Patsy and Spada Nina (1993), *How languages are learned*, Oxford: Oxford University Press
- Little David (1991), *Learner autonomy 1: Definitions, issues and problems*, Dublin: Authentik
- Little David (1999), Metalinguistic Awareness: The Cornerstone of Learner Autonomy, in Mißler & Multhaup (eds)(1999), pp.3-12
- Littlemore Jeannette (2001), Learner autonomy, self-instruction and new technologies in language learning: Current theory and practice in higher education in Europe, in Chambers & Davies (eds)(2001), pp.39-52
- Littlewood William (1981), *Communicative Language Teaching*, Cambridge: Cambridge University Press
- Longman's audio-visual French. Stage 1* (1966), London : Longman, 1966
- Longman's audio-visual French. Stage 2* (1967), London : Longman, 1967
- Louw Bill (1991), Classroom Concordancing of Delexical Forms and the Case for Integrating Language and Literature, in Johns & King (1991), pp.151-178
- Louw Bill (1997), The Role of Corpora in Critical Literary Appreciation, in Wichmann *et al* (1997), pp.240-251
- Louw W.E. (1989), Computer Assisted Materials Evaluation: Content vs National Policy, *ELR Journal*, 3, pp.29-41
- Lunt Helen (ed)(1982), *Communication skills in modern languages at school and in higher education*, London: CILT
- Lyons John (1977), *Chomsky* (revised edition), Glasgow: Fontana/Collins
- Maia Belinda (1997), Making corpora – a learning process, paper given at the conference on *Corpus use and learning to translate*, Bertinoro, 14-15 November 1997, on-line at <<http://www.sslmit.unibo.it/cultpaps/maia.htm>>, 7 pages, accessed 11 December 2003
- Maia Belinda, Haller Jonathan, Urlrych Margherita (eds)(2002), *Training the Language Services Provider for the New Millennium*, Porto: Faculdade de Letras, Universidade do Porto
- Manning Patricia (1994), Learner data collection and interpretation in the design of intelligent CALL programs, in Thompson & Chesters (1994), pp.3-12
- Manning Patricia (1996), Exploratory teaching of grammar rules and CALL, *ReCALL*, 8:1, pp.24-30
- Martin J.R, Matthiessen Christian & Painter Clare (1997), *Working with Functional Grammar*, London: Arnold
- McArthur Tom (1983), *A Foundation Course for Language Teachers*, Cambridge: Cambridge University Press
- McCulloch Derek (1995), Where has all the grammar gone? An 'accusative' search, *German Teaching*, 12, 13-18

- McDonough Steven (1986), *Psychology in foreign Language Teaching*, London: Allen & Unwin
- McDonough Steven (1995), *Strategy and skill in learning a foreign language*, London: Edward Arnold
- McDonough Steven (1999), Learner strategies, *Language Teaching*, 32:1, pp.1-18
- McEnery Tony & Wilson Andrew (1993), The Role of Corpora in Computer-Assisted Language Learning, *Computer Assisted Language Learning*, 6:3, pp.233-248
- McEnery Tony and Wilson Andrew (1996), *Corpus Linguistics*, Edinburgh: Edinburgh University Press
- McEnery Tony & Wilson Andrew (1997a), *Corpus linguistics*, supplement to McEnery and Wilson (1996),
<<http://www.ling.lancs.ac.uk/monkey/ihe/linguistics/contents.htm>>, accessed 29 August 2001
- McEnery Tony & Wilson Andrew (1997b), Teaching and language corpora (TALC), *ReCALL*, 9:1, pp.5-14
- McEnery Tony & Wilson Andrew (2000), *Corpus linguistics*, ICT4LT Module 3.4, <http://www.ict4lt.org/en/en_mod3-4.htm>, 18 pages, accessed 10 July 2001
- McEnery Tony, Wilson Andrew & Barker Paul (1997), Teaching grammar again after twenty years: corpus-based help for teaching grammar, *ReCALL*, 6:2, pp.8-16
- Metcalf Peter (1992), CALL, the foreign-language undergraduate and the teaching of grammar: a linguistic and political battlefield, *ReCALL* 7, pp.3-5
- Metcalf Peter, Laurillard Diana & Mason Robin (1995), The decline of written accuracy in pupils' use of French verbs, *Language Learning Journal*, 12, pp.47-50
- Metcalf Peter, Laurillard Diana & Mason Robin (1998), 'It's just a word': pupil perceptions of verb form and function, *Language Learning Journal*, 17, pp.14-20
- Mindt Dieter (1996), English corpus linguistics and the foreign language teaching syllabus, in Thomas & Short (1996), pp.232-247
- Mindt Dieter (1997), Corpora and the Teaching of English in Germany, in Wichmann *et al* (eds)(1997), pp.40-50
- Minugh David (1997), All the Language that's Fit to print: Using British and American Newspaper CD-ROMS as Corpora, in Wichmann *et al* (eds)(1997), pp.67-82
- Mißler Bettina & Multhaup Uwe (eds)(1999), *The Construction of Knowledge, Learner Autonomy and Related Issues in Foreign Language Learning*, Tübingen: Stauffenberg Verlag
- Mitchell Rosamond (1988), *Communicative Language Teaching in Practice*, London: CILT
- Mitchell Rosamond (2003), Rethinking the concept of progression in the National Curriculum for Modern Foreign Languages: a research perspective, *Language Learning Journal*, 27, pp.15-23
- Mitchell Rosamond and Miles Florence (1998), *Second Language Learning Theories*, London: Arnold
- Mohammed Abdel-Moneim M.(1993), Towards a learner-centred technique of teaching grammar, *Language Learning Journal*, 7, pp.59-63
- Morgan Carol & Neil Peter (2001), *Teaching Modern Foreign Languages*, London: Kogan Page
- Motteram Gary (2000), Communicating with computers, in Brett & Motteram (eds)(2000), pp. 73-91
- Mparutsa Cynthia, Love Alison & Morrison Andrew (1991), Bringing Concord to the ESP Classroom, in Johns & King (1991), pp.115-134

- Müller Bettina & Groß Annette (1999), Possible Applications of Parallel Concordancers Using MULTICONCORD in Translation, in Mißler & Multhaupt (eds)(1999), pp.65-78
- Mullings Christine, Deegan Marilyn, Ross Seamus & Kenna Stephanie (eds), (1996), *New Technologies for the Humanities*, London: Bowker Saur
- Munby John (1978), *Communicative Syllabus Design*, Cambridge: Cambridge University Press
- Murison-Bowie Simon (1993), *MicroConcord: manual*, Oxford: Oxford University Press
- Murison-Bowie Simon (1996), Linguistic Corpora and Language Teaching, *Annual Review of Applied Linguistics*, vol 16, pp.182-199
- Nirenburg Sergei (ed)(1987), *Machine translation: Theoretical and methodological issues*, Cambridge: Cambridge University Press
- Nott David (1996), Grammar: what is taught and what is learned, *Francophonie*, 13, pp.3-18
- Nunan David (1991), *Language Teaching Methodology*, New York: Prentice Hall
- Oakes Michael (1998), *Statistics for Corpus Linguistics*, Edinburgh: Edinburgh University Press
- O'Keefe Anne & Farr Fiona (2003), Using Language Corpora in Initial Teacher Education: Pedagogical Issues and Practical Applications, *TESOL Quarterly*, 37:3, pp.389-418
- O'Malley J Michael & Chamot Anna Uhl (1990), *Learning Strategies in Second Language Acquisition*, Cambridge: Cambridge University Press
- Ooi Vincent (1998), *Computer Corpus Lexicography*, Edinburgh: Edinburgh University Press
- Ooi Vincent (2001), Investigating and Teaching Genres Using the World Wide Web, in Ghadessy, Henry & Roseberry (eds) (2001), pp.175-203
- Owen Charles (1996), Do concordances require to be consulted?, *ELT Journal*, 50:3, pp.219-224
- Page Brian (1994), Grammar and Communication, in King & Boaks (1994), pp.14-16
- Page Brian & Hewett Derek (1987), *Languages Step by Step: Graded Objectives in the UK*, London: CILT
- Patton Peter & Holoien Renee (eds)(1981), *Computing in the Humanities*, Gower
- Pearson Jennifer (1996), Electronic texts and concordances in the translation classroom, *TEANGA*, 16:March, pp.85-95
- Pearson Jennifer (2000), Surfing the Internet: teaching students to choose their texts wisely, in Burnard & McEnery (eds)(2000), pp.235-239
- Pennington Martha C.(ed).(1996), *The Power of CALL*, Houston: Athelstan
- Pennington Martha C. & Stevens Vance (eds) (1992), *Computers in Applied Linguistics*, Clevedon: Multilingual Matters
- Perrin Geoffrey (1991), It all depends on context?, *German Teaching*, 4, p.9
- Peters Carol & Picchi Eugenio (1998), Bilingual Reference Corpora for Translators and Translation Studies, in Bowker, Cronin, Kenny & Pearson (eds)(1998), pp.91-100
- Phillips David (ed)(1988), *Languages in Schools from Complacency to Conviction*, London: Centre for Information on Language Teaching and Research (CILT)
- Phillips M. (1985), Logical Possibilities and classroom scenarios for the development of CALL, in Brumfit *et al* (1985), pp.25-46
- Pienemann Manfred & Jansen Louise (1992), Computational Analysis of Language Acquisition Data, in Pennington & Stevens (eds) (1992), pp.201-243

- Piotrowska Maria (1997), Criteria for selecting parallel texts in teaching a translation course, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.411-420
- Polezzi Loredana, (1993), Concordancing and the Teaching of *ab initio* Italian Language for Specific Purposes, *ReCALL*, vol.9, pp.14-19
- Polezzi Loredana (1994), Concordancers in the design and implementation of foreign language courses, in Thompson & Chesters (1994), pp.89-96
- Porcher L., Huart M. & Mariet F. (1980), *Adaptation de « Un niveau-seuil » pour des contextes scolaires*, Strasbourg: Council of Europe
- Power Ted (1977), *Motivating children in poorly resourced language learning environments*, on-line at <<http://www.btinternet.com/~ted.power/esl0624.html>>, accessed on 27 July 2004
- Pride J.B. (ed)(1979), *Sociolinguistic Aspects of Language Learning and Teaching*, Oxford: Oxford University Press
- Pride J. & Holmes Janet (eds)(1972), *Sociolinguistics*, London: Penguin, /// pp.269-293
- Qiao Hong Liang (1995), Navigating in corpus space with the Mini-Concordancer: an autonomous language learning perspective, *ON-CALL*, 9:3, 6 pages, on-line at <<http://www.cltr.uq.au/oncall/qiao93.html>>, accessed 11 December 2003
- Qiao Hong Liang and Sussex Roland (1996), Using the Longman Mini-Concordancer on tagged and parsed corpora, with special reference to their use as an aid to grammar learning, *System*, 24:1, pp.41-64
- Quasthoff Uwe & Wolff Christian (1999), Korpuslinguistik und große einsprachige Wörterbücher, *Linguistik online*, 3:2, <http://www.linguistik-online.com/2_99/quasthoff.html>, 10 pages, accessed 25 October 2000
- Ragan Peter (2001), Classroom Use of a Systemic Functional Small Learner Corpus, in Ghadessy, Henry & Roseberry (eds) (2001), pp 207-236
- Reed Alan and Schonfelder J. Laurie (1979), CLOC: a general-purpose concordance and collocations generator, in Ager *et al* (eds)(1979), pp.59-72
- Reppen Randi & Biber Douglas (2001), Habeas corpus, but will you be using it in class?, *EL teaching matters (TESOL Special)*, February 2001, p.1
- Rézeau Joseph (1997), Concordances, Cédérom et Internet au service de l'enseignement du français aux adultes, *The Dong-Eui International Journal*, <http://www.uhb.fr/%7Erezeau_j/articles/dongelui/dong_eui.htm>, 19 pages, accessed 05 January 2000
- Rézeau Joseph (1999), Profils d'apprentissage et représentations dans l'apprentissage des langues en environnement multimédia, *ALSIC*, 2 :1, pp.27-49 ,<<http://alsic.u-strasbrg.fr/Num3/rezeau2/default.htm>>, accessed 06 January 2003
- Rézeau Joseph (2001), Concordances in the classroom: The evidence of the data, in Chambers & Davies (eds)(2001), pp. 147-166
- Richards J. and Rodgers T. (1982), Method: Approach, design and procedure, *TESOL Quarterly*, 16:2, 153-168
- Richards Jack and Rodgers Theodore (1986), *Approaches and Methods in Language Teaching: A description and analysis*, Cambridge: Cambridge University Press
- Richterich R & Chancerel J-L (1980), *Identifying the Needs of Adults Learning a Foreign Language*, Oxford: Pergamon Press
- Rivers Wilga (1968), *Teaching foreign-language skills*, Chicago: University of Chicago press
- Rivers Wilga (1983), *Communicating Naturally in a Second Language*, Cambridge: Cambridge University Press
- Robinson Gail (1991), Effective Feedback Strategies in CALL: Learning Theory and

- Empirical Research, in Dunkel (ed) (1991), pp.133-167
- Rodriguez Maria C.R. (no date), *Using a Concordancer in Literary Studies*, <<http://vlc.polyu.edu.hk/Concordance/Review/programa.htm>>, 9 pages, accessed 05 April 2001
- Rogers Margaret (1996), What's theory got to do with it? Grammar and Communication in L2 Acquisition, in Engel & Myles (eds)(1996a), pp.21-43
- Romary Laurent, Bonhomme Patrice, Bruneseaux Florence & Pierrel Jean-Marie (1999), Silfide: A System for Open Access and Distributed Delivery of TEI Encoded Documents, *Computers and the Humanities*, 33, pp.31-38
- Roulet Eddy (1975), *Linguistic Theory, Linguistic Description and Language Teaching*, Harlow: Longman
- Roussel Francine (1991), Parallel Concordances and Tonic Auxiliaries, in Johns & King (1991), pp.71-101
- Rubdy Rani (1998), Sensitizing Learners Towards a More Productive Classroom Participation: An 'Exploratory' Approach, *RELC Journal*, 29:2, pp.1-26
- Rundell Michael (1996a), Computer Corpora and Their Impact on Lexicography and Language Teaching, in Mullings *et al* (eds), (1996), pp.198-216
- Rundell Michael (1996b), The corpus of the future, and the future of the corpus, paper given at Conference on 'New Trends in Reference Science' at Exeter University, 29.03.96, accessible on-line at <<http://www.ruf.rice.edu/~barlow/futcrp.html>>, 11 pages, accessed 17 June 1999
- Rundell Michael & Stock Penny (1992), The Corpus Revolution, *English Today* 30, 8:2, pp.9-14; *English Today* 31, 8:3, pp.21-32; *English Today* 32, 8:4, pp.45-51
- Rutherford, W.E., (1987) *Second language grammar: learning and teaching*, Harlow: Longman
- Salkie Raphael (1997), Naturalness and contrastive linguistics, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp. 297-312
- Salkie Raphael (2001), *INTERSECT: Multilingual Corpora and Contrastive Linguistics*, INTERSECT homepage on-line at <<http://www.brighton.ac.uk/edusport/languages/html/intersect.html>>, 5 pages, accessed 22 January 2004
- Sampson Geoffrey (1996), From central embedding to corpus linguistics, in Thomas & Short (1996), pp.14-26
- Saunders Keith (1996), Grammatical accuracy: a response to Derek McCulloch, *German Teaching*, 13, pp.30-32
- Saunders Keith, Martin Graham & Richards Brian (1994), The standards debate – responses to Richard Sheppard, *German Teaching*, 9, pp.10-14
- Schäffner Christina (1998), Parallel Texts in Translation, in Bowker, Cronin, Kenny & Pearson (eds)(1998), pp.83-90
- Schulz Reinhard (1988), *Folgerungen fachdidaktischer und programmtechnischer Art und Erfahrungen nach einem mehrjährigen Einsatz des Mediums Computer im Fremdsprachen-unterricht einer Realschule*, Bochum, pp.1-44
- Seedhouse Paul (1995), Communicative CALL: focus on the interaction produced by CALL software, *ReCALL*, 7:2, pp.20-28
- Seidlhofer Barbara (2000), Optionalizing intertextuality: using learner corpora for learning, in Burnard & McEnery (eds)(2000), pp.208-223
- Seliger Herbert (1983), Learner Interaction in the Classroom and its effect on Language Acquisition, in Seliger and Long (eds)(1983), pp.246-267
- Seliger Herbert and Long Michael (eds)(1983), *Classroom Oriented Research in Second*

- Language Acquisition*, Cambridge (Mass): Newbury House
- Sheppard Richard (1993), Getting down to brass syntax: German teaching and the Great Standards Debate, *German Teaching*, 8, pp.2-9
- SILFIDE (Serveur Interactif pour la Langue Française, son Identité, sa Diffusion et son Etude), French corpus developed under the auspices of the CNRS, AUPELF and UREF and accessible at the following address :
<<http://www.loria.fr/projets/Silfide/Index.html>>
- Sinclair John (1986), Basic Computer Processing of Long Texts, in Leech & Candlin (eds) (1986), pp.185-203
- Sinclair John (ed) (1987), *Collins COBUILD English Language Dictionary*, London: Collins
- Sinclair John (1991), *Corpus, Concordance, Collocation*, Oxford: Oxford University Press
- Sinclair John (2001), Preface, in Ghadessy, Henry & Roseberry (eds) (2001), pp.vii-xv
- Skinner B.F (1957), *Verbal Behavior*, New York: Appleton Crofts
- Someya Yasumasa (1999), *A Corpus-based Study of Lexical and Grammatical Features of Written Business English*, on-line at
<<http://www2.gol.com/users/someya/MA/Abstract&Acknowledgements&TableContents.html>>, accessed 03 April 2000
- Somogyi Emma (1996), Using the concordancer in vocabulary development for the Cambridge Advanced English (CAE) course, *ON-CALL*, 10:2, 7 pages, on-line at
<<http://www.cltr.uq.edu.au/oncall/somogyi102.html>>, accessed 11 December 2003
- Spada Nina (1997), Form-Focussed Instruction and Second Language Acquisition: A Review of Classroom and Laboratory Research, *Language Teaching*, 30:2, pp.73-87
- Stern H.H. (1983), *Fundamental Concepts of Language Teaching*, Oxford: Oxford University Press
- Stevens Vance (1988), Studying Vocabulary Using Concordancers on Microcomputers, *TESOL Newsletter*, 22:3, p.27
- Stevens Vance (1991a), A study of student attitudes toward CALL in a self-access student resource centre, *System*, 19:3, pp.289-299
- Stevens Vance (1991b), Concordance-based Vocabulary Exercises : a Viable Alternative to Gap-Fillers, in Johns & King (1991), pp.47-61
- Stevens Vance (1992), Humanism and CALL: a Coming of Age, in Pennington & Stevens (eds)(1992), pp.11-38
- Stevens Vance (1993), Concordancers as enhancements to language competence, *TESOL Matters*, 2:6, pp.11
- Stevens Vance (1995), Concordancing with Language Learners: Why? When? What?, *CAELL Journal*, vol.6, no.2, pp.2-10
- St.John Elke (2001), A case for using a parallel corpus and concordancer for beginners of a foreign language, *Language learning & Technology*, 5:3, pp.185-203, available on-line at <<http://lt.msu.edu/vol5num3/stjohn/>>, accessed 07 November 2001
- Stevens P.D.(1965), *Papers in Language and Language Teaching*, Oxford: Oxford University Press
- Stubbs Michael (1996), *Text and Corpus Analysis*, Oxford: Blackwell
- Szakos Josef. (2000), Producing and using corpora in Chinese language education, in Burnard & McEnery (eds.)(2000), pp.153-157.
- Templeton Brad (n.d.), *A brief intro to copyright*, on-line at
<<http://www.templetons.com/brad/ A brief intro to copyright.htm>>, accessed 26 January 2004

- Thomas Jenny & Short Mick (1996), *Using Corpora for Language Research*, Harlow: Longman
- Thompson June (1993), *Language Learning and Computers: A Survey of Use in the UK Higher Education Sector*, Hull: CTI centre for Modern Languages
- Thompson June and Chesters Graham (eds)(1994), Emancipation through learning Technology: selected papers from the EUROCALL '93 conference, Hull, 15-17 September 1993, *Computers and Education*, 23:1/2
- Tognini-Bonelli Elena (2001), *Corpus Linguistics at Work*, Amsterdam and Philadelphia: John Benjamins
- Tomlinson Brian (ed)(1998), *Materials Development in Language Teaching*, Cambridge: Cambridge University Press
- Towndrow Phillip (1999), Students' reactions to CALL in the Gulf, *CALL EJ Online*, <<http://www.lerc.ritsumei.ac.jp/caliej/2-1/Towndrow.html>>, 6 pages, accessed 14 June 1999
- Tribble Chris (1990), Computers, Corpora, Language Teaching: a practical introduction to the use of concordancing in language teaching and learning, *Die Neueren Sprachen*, 89:5, pp.465-475
- Tribble Christopher (1997), Improvising corpora for ELT: quick-and-dirty ways of developing corpora for language teaching, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.106-117, also accessible on-line at <[http://ourworld.compuserve.com/homepages/ Christopher Tribble/Palc.htm](http://ourworld.compuserve.com/homepages/Christopher_Tribble/Palc.htm)>, 10 pages, accessed 07 July 1999
- Tribble Christopher (2000), Practical uses for language corpora in ELT, in Brett & Motteram (eds)(2000), pp. 31-41
- Tribble Chris & Jones Glyn (1990), *Concordancers in the Classroom*, Harlow: Longman
- Tribble Chris & Jones Glyn (1997), *Concordancers in the Classroom*, new edition, Houston: Athelstan
- Trim J.L.M., Richterich R., Van Ek J.A. A Wilkins D.A. (1973), *Systems Development in Adult Language Learning*, Strasbourg: Council of Europe, republished 1980, Oxford: Pergamon Press
- Turnbull Jill & Burston Jack (1998), Towards independent concordance work for students: Lessons from a case study, *ON-CALL*, 12:2, 11 pages, on-line at <<http://www.cltr.uq.edu.au/oncall/turnbull122.html>>, accessed 11 December 2003
- Ulrych Margherita (1997), The impact of multilingual parallel concordancing on translation, in Lewandowska-Tomaszczyk & Melia (eds)(1997), pp.421-435
- Urdang Laurence (1991), Brian Chandler, *Longman Mini Concordancer* (review), *International Journal of Lexicography*, 4:2, pp.157-159
- Underwood John (1984), *Linguistics, computers and the language teacher: a communicative approach*, Rowley MA: Newbury House
- Valdman A.(ed)(1996), *Trends in Language Teaching*, New York: McGraw-Hill
- Van Ek J.A. (1975), *The Threshold Level*, Strasbourg: Council of Europe
- Van Ek J.A. (1977), *The Threshold Level for Modern Language Learning in Schools*, Harlow: Longman
- Van Els Theo, Bongaerts Theo, Extra Guus, Van Os Charles & Janssen-Van Dieten Anne-Mieke (1984), *Applied Linguistics and the Learning and Teaching of Foreign languages*, London: Edward Arnold
- Varantola Krista (1997), Translators, dictionaries and text corpora, paper given at the conference on *Corpus use and learning to translate*, Bertinoro, 14-15 November

- 1997, on-line at <<http://www.sslmit.unibo.it/cultpaps/varanto.htm>>, 8 pages, accessed 11 December 2003
- Vogel Thomas (2001), Learning out of control: Some thoughts on the World Wide Web in learning and teaching foreign languages, in Chambers & Davies (eds)(2001), pp.133-145
- Voix d'Auvergne* (1977), Transcriptions of recordings made in Clermont-Ferrand and the surrounding region in 1976 by teachers from British Colleges of Education under the auspices of the Institut Français du Royaume-Uni, London: Institut Français du Royaume-Uni
- Voix et images de France* (1966), Crédif, Paris: Didier-Hatier
- Vygotsky Lev (1986), *Thought and Language*, Cambridge, Massachusetts: MIT Press
- Warschauer, M. (1996). Computer-assisted language learning: an introduction, in Fotos (ed.)(1996), pp.3-20
- Warschauer M. (2000), The death of cyberspace and the rebirth of CALL, *English Teachers' Journal*, 53, pp.61-67, <<http://www.gse.uci.edu/markw/cyberspace.html>>, 7 pages, accessed 12 March 2003
- Warschauer Mark & Healey Deborah (1998), Computers and language learning: on overview, *Language teaching*, 31:2, pp.57-71
- Wenden Anita (1991), *Learner Strategies for Learner Autonomy*, London: Apprentice Hall
- Whistle Jeremy (1986) Ends and Means: the Shift from an exercise-based approach to a task-based approach, In Bate & Hare (eds)(1986), pp.127-141
- Whistle Jeremy (1999a), Concordancing and Learner Autonomy: an Experiment with First and Second Year Undergraduates, in Cameron (ed)(1999), pp.443-454
- Whistle Jeremy (1999b), Concordancing with students using an 'off-the-web' corpus, *ReCALL*, 11:2, pp.74-80
- Whitton Kenneth (1994), Deutsch in Großbritannien – a review article, *German Teaching*, 10, pp.22-27
- Wichmann Anne, 1995, Using concordances for the teaching of modern languages in higher education, *Language Learning Journal*, 11, pp.61-63
- Wichmann Anne, Filgelstone Steven, McEnery Tony & Knowles Gerry (1997), *Teaching and Language Corpora*, Harlow: Longman
- Widdowson H.G. (1978), *Teaching Language as Communication*, Oxford: Oxford University Press
- Widdowson H.G. (1979), *Explorations in Applied Linguistics*, Oxford: Oxford University Press
- Widdowson H.G. (1984), *Explorations in Applied Linguistics 2*, Oxford: Oxford University Press
- Widdowson Henry (2000), On the limitations of linguistics applied, *Applied Linguistics*, 21:1, pp.3-25
- Wilkins D.A. (1972), *Linguistics in Language Teaching*, London: Edward Arnold
- Wilkins D.A. (1973), The Linguistic and Situational Content of the Common Core in a Unit/Credit System, in Trim *et al* (1973)
- Wilkins D.A. (1974), *Second-language learning and teaching*, London: Edward Arnold
- Wilkins D.A. (1976), *Notional Syllabuses*, Oxford: Oxford University Press
- Willis Dave (1990), *The Lexical Syllabus: a new approach to language teaching*, London: Collins
- Willis Dave (1991), *Collins COBUILD Student's Grammar*, London: Harper Collins
- Willis Jane (1998), Concordances in the classroom without a computer, in Tomlinson

- (ed)(1998), pp.44-66
- Willis Jane and Willis Dave (1986), *Challenge and Change in Language Teaching*, London: Heinemann
- Willis Jane and Willis Dave (1988), *Collins COBUILD English Course*, London: Collins
- Wilson Eve (1997), The Automatic Generation of CALL Exercises from General Corpora, in Wichmann *et al* (eds)(1997), pp.116-130
- Windeatt S. (1986), Observing CALL in action, in Leech & Candlin (eds) (1986), pp.79-97
- Wolff Dieter (1984), Computers in classroom research, in Thompson & Chesters (1994), pp.133-142
- Wong Irene, Cheung Dorothy & Ching Lai Phooi (1992), Concordancing in the Language Classroom, *Cross Currents*, XIX:1, pp.41-46
- Wright Margaret (1999), Grammar in the languages classroom: findings from research, *Language Learning Journal*, 19, pp.33-39
- Wright Margaret & Craig Olive (1995), L'Enseignement de la grammaire, *Francophonie*, 11, pp.25-28
- Wu Marian (1992), Towards a Contextual Lexico-grammar: An Application of Concordance Analysis in EST Teaching, *RELC Journal*, 23:2, pp.18-34
- Zanettin Federico (1998), Bilingual Comparable Corpora and the Training of Translators, *META*, XLIII:4, on-line at <http://www.erudit.org/revue/meta/1998/v43/n4/004638ar.html>, 14 pages, accessed 22 January 2004
- Zanettin Federico (2002), DIY corpora: the WWW and the translator, in Maia *et al* (eds)(2002), pp.239-248, on-line at <http://www.federicozanettin.net/DIYcorpora.htm>, 8 pages, accessed 22 January 2004

APPENDIX 1

Questionnaire

QUESTIONNAIRE ON USE OF COMPUTER

(please circle the appropriate response: 5 = very; 4 = quite a lot; 3 = fairly; 2 = a little; 1 = not at all)

COURSE: YEAR:

1. In general do you like using the computer for language work? 5 4 3 2 1
2. a) Do you find the grammar programs useful? 5 4 3 2 1
 - 1) if yes, what do you find helpful?
.....
.....
 - 2) if no, what do you find unhelpful?
.....
.....
- b) Do you find the translation programs useful? 5 4 3 2 1
 - 1) if yes, what do you find helpful?
.....
.....
 - 2) if no, what do you find unhelpful?
.....
.....
- c) Do you find using the concordancer useful? 5 4 3 2 1
 - 1) if yes, what do you find helpful?
.....
.....
 - 2) if no, what do you find unhelpful?
.....
.....
3. Should computer sessions be timetabled (1) more often (2) less often (3) the same as this year?
4. Did you use computers before coming to UCN for foreign language work? Yes / No
if yes, please list the sorts of activities the computer was used for:
.....
.....
5. In what areas do you think the computer is most useful?
.....
.....
6. In what areas do you think the computer is least useful?
.....
.....
7. Have you any suggestions for ways to use the computer better?
.....
.....
.....

Thank you for answering these questions.
Jeremy Whistle

APPENDIX 2

Worksheets

FRENCH: LANGUAGE: LEVEL 1

Activity number: 1a

Directory/Files	c:\french\french.txt\fr_dip\label\19
Language structure	noun and noun group
Word to search for:	France
Procedure:	
Select files	<ol style="list-style-type: none"> 1) click the yellow folder icon "source files" 2) click the drive control box and then click "C" to change drive 3) click the appropriate directory to change to "...label\19" 4) click the little triangle in the "Files of type" box and then click *.txt" 5) hold down the shift key with your left hand and then, with your right hand, click first the top name in the files list box and then the bottom name to highlight all the files 6) click "Ok"
Select target directory	<ol style="list-style-type: none"> 1) click the white folder icon ("target directory") 2) click the drive box and then click "C" to change drive 3) click "french" 4) click the "Create new directory" box 5) in the box that appears, enter the word "temp" 6) press "Enter" 7) click "Ok"
Select strings	<ol style="list-style-type: none"> 1) click the keyboard icon ("enter strings") 2) when the new window has appeared, enter the word "France" in the top text box 3) click the "Case sensitive" circle 4) click "Ok"
Perform search	<ol style="list-style-type: none"> 1) click the "Save to File" box under the "Search" button 2) click "Search"
View results	<ol style="list-style-type: none"> 1) click the monitor icon ("view target files") 2) in the files list box click "dumpfile.001" and then click "Ok" 3) use the "Pause/Continue" and "Up" and "Down" buttons to control the movement of the text
Analysis	<p>look at the words that immediately precede the word "France"</p> <p>how many are there in total? what types of word are they?</p> <p>can you identify any pattern(s)?</p> <p>translate the word "France" and those words which belong with it into English</p>
Conclusions	what differences in construction are there between the two languages?
Rule	formulate a basic rule when using the word "France" in French

Activity number: 1b

Directory/Files	c:\french\french.txt\fr_dip\label\19
Language structure	noun and noun group
Word to search for:	Paris
Procedure:	
Select files	no change
Select target directory	no change
Select strings	1) click the keyboard icon ("enter strings") 2) when the new window has appeared, click "Clear" 3) enter the word "Paris\" (with backslash) in the top text box 4) click the "Case sensitive" circle if not checked and finally click "Ok"
Perform search	click "Search"
View results	same procedure as last time except that the name of the file to click is "dumpfile.002"
Analysis	look at the words that immediately precede the word "Paris" how many are there in total? what types of word are they? translate the word "Paris" and those words which belong with it into English
Conclusions	what differences in construction in French are there between "France" and "Paris"?
Rule	formulate a basic rule explaining the different constructions with "France" and "Paris"
Extension	think of a couple of countries whose name in French ends in a consonant repeat this activity with the new names (clear "Paris" first) what do you notice?

Activity number: 1c

Directory/Files	c:\french\french.txt\fr_dip\label\19
Language structure	noun and noun group
Word to search for:	pays
Procedure:	
Select files	no change
Select target directory	no change
Select strings	1) click the keyboard icon ("enter strings") and, when the new window has appeared, enter the word "pays\" in the text box 2) click the "Case sensitive" circle and finally click "Ok"
Perform search	3) click the "Save to File" box under the "Search" button 4) click "Search"
View results	1) same procedure except that the name of the file to click is "dumpfile.003" 2) click "Options" menu at top, then "Sort concordance" 3) enter "1" in top box (to replace "0") 4) click "Ok"
Analysis	look at the words that immediately precede the word "pays" how many are there in total? are there any new ones? what about the words that follow? can you identify any pattern(s)? (translate the word "pays" and those words which belong with it into English)
Conclusions	what differences in construction are there between the two languages?
Rule	formulate a general rule for constructing the noun group in French

Activity number: 1d

Directory/Files	c:\french\french.txt\fr_dip\label
Language structure	position of adjectives
Word to search for:	français, commercial, rouge, blanc bon, mauvais, petit, grand certain, dernier, long, ancien, vif/vive, curieux
Procedure:	
Select files	1) click the yellow folder icon "source files" 2) click the drive control box and then click "C" to change drive 3) click the appropriate directory to change to "french...." 4) hold down the shift key with your left hand and then with your right hand click first the top name in the files list box and then the bottom name to highlight all the files 5) click "Ok"
Select target directory	1) click the white folder icon ("target directory") 2) click the drive box and then click "C" to change drive 3) click "french" 4) click the "Create new directory" box 5) in the box that appears, enter the word "temp" 6) press "Enter" 7) click "Ok"
Select strings	1) click the keyboard icon ("enter keywords") 2) when the new window has appeared, enter the first word above in the top text box 3) click "Ok"
Perform search	1) click the "Save to File" box under the "Search" button if not checked 2) click "Search"
View results	1) click the monitor icon ("view target files") 2) in the files list box click the "dumpfile" with the highest number and then click "Ok" 3) use the scrollbars to control the movement of the text
Analysis	look at the position of the adjective in relation to the noun it refers to is the position consistent? if not, what influence does the position have on the meaning? translate the adjective and nouns into English to test your hypothesis
Conclusions	

FRENCH: LANGUAGE: LEVEL 2

Activity number: 2a

Directory / Files	c:\french\french.txt\fr_dip\label\31
Language structure	personal pronouns
Words to search for	lui, elle
Procedure :	
Select Files	<ol style="list-style-type: none"> 1) click the yellow folder icon "source files" 2) click the drive control box and then click "c" to change drive 3) click the appropriate directories to change to ". . \label\3 1 " 4) click the little triangle in the "Files of type" box and then click *.txt" 5) hold down the shift key with your left hand and then, with your right hand, click first the top name in the files list box and then the bottom name to highlight all the files 6) click "Ok"
Select target directory	<ol style="list-style-type: none"> 1) click the white folder icon ("target directory") 2) click "french" 3) if you do not see 'temp' click the "Create new directory" box 4) in the box that appears, enter the word 'temp' 5) press "Enter" 6) click "Ok"
Select strings	<ol style="list-style-type: none"> 1) click the keyboard icon ("enter strings") 2) when the new window has appeared, enter the word '\lui\' in the top text box (the backslashes are important) 3) click "Ok"
Perform search	<ol style="list-style-type: none"> 1) click the "Save to File" box under the "Search" button 2) click "Search"
View results	<ol style="list-style-type: none"> 1) click the monitor icon ("view target files") 2) in the files list box click "dumpfile.001" and then click "Ok" 3) use the scroll bars to control the movement of the text if necessary
Analysis	<ol style="list-style-type: none"> 1) look at the contexts that immediately surround the word 'lui' (to expand the context, double click on the line of text) and identify the different functions of the pronoun 'lui' 2) repeat the search, starting at "Select strings" but this time click "Clear" before entering the pronoun '\elle\' (once again using the backslashes) 3) repeat the other stages, but this time to view the results click "dumpfile.002"
Conclusions	in what ways does the pronoun 'elle' differ from 'lui'
Rule	how many types of personal pronoun are there? can you list them?

FRENCH: LANGUAGE: LEVEL 2

Activity number: 2b

Directory / Files	c:\french\french.txt\fr_dip\label\31
Language structure	demonstrative pronouns
Words to search for	celui, celle
Procedure:	
Select Files	no change
Select target directory	no change
Select strings	1) click the keyboard icon ("enter strings") and, when the new window has appeared, enter the word 'celui' in the text box
Perform search	1) click the "Save to File" box under the "Search" button 2) click "Search"
View results	1) same procedure except that the name of the file to click is "dumpfile.003 "
Analysis	1) look at the examples of 'celui' in context what does 'celui' stand for in each case? are there any patterns? translate 'celui' and those words which belong with it into English) how many different ways of translating 'celui' are there? 2) repeat the procedure with 'celle'
Conclusions	what differences in construction are there between the two
Rule	what forms are missing in the examples you have for analysis? formulate some rules for use of the demonstrative pronouns

FRENCH: LANGUAGE: LEVEL 2

Activity number: 2c

Directory / Files	c:\french\french.txt\fr_dip\label
Language structure	possessive pronouns
Words to search for	\sien\, siens, \sienne
Procedure:	
Select Files	click "all directories"
Select target directory	no change
Select strings	1) click the keyboard icon ("enter strings") 2) when the new window has appeared, enter the word "\lui\" in the top text box (the backslashes are important) 3) click "Ok"
Perform search	click "Search"
View results	same procedure as last time except that the name of the file to click is "dumpfile.004"
Analysis	translate the various forms of 'sien' into English are there any differences in usage?
Conclusions	what does 'sien' correspond to in English? does it behave in the same way as its English equivalent?
Rule	formulate a basic rule explaining the difference between 'son' etc and 'sien' etc what other pronouns are there in this category?

APPENDIX 3

Course materials based largely on concordances (extracts and expanded examples) and frequency counts

FRE1002 FRENCH GRAMMAR AND USAGE

UNIT 1B: ETRE ETUDIANT PART 2

ACTIVITY 1: LISTENING

ETUDIANTE EN KINESITHERAPIE

Vous allez entendre un entretien avec une jeune étudiante. Dans la transcription de l'entretien que vous avez sous les yeux, on a laissé des blancs. Ecoutez l'enregistrement et complétez les phrases.

alors mademoiselle est-ce que vous voulez bien me dire ce que vous faites dans la vie
oui bien sûr (rire) alors euh moi je suis étudiante kinésithérapie à Clermont-Ferrand hein
euh alors voilà c'est tout
et vous faites des de combien d'
alors là mes durent trois - euh trois dans une c'est
pas une université c'est un petit peu différent c'est assez euh c'est c'est pas encore très très bien
organisé comme en France la kinésithérapie
alors parlez-moi un petit peu de la de - kinésithérapeute
oui c'est (rire) alors euh bien la on commence euh on a un
enseignement très au départ - on apprend bon ce qui est euh surtout euh
indispensable c'est-à-dire l'anatomie la physiologie bon alors ça à fond et puis ensuite à la fin
de la première année on commence à aller tous les matins à l'hôpital - et à travailler avec des
gens sur des malades et après euh la c'est la même chose on continue des
..... théoriques et tous les matins on est à l'hôpital on a un mois de vacances pendant
l'été on on intercale y a la moitié de l' qui part en vacances au mois de juillet et
l'autre moitié au mois d'août et puis troisième c'est la même chose on
continue euh et à la fin de la troisième année on un national qu'on a
ou pas et (alors) ensuite on peut travailler
est-ce qu'il y a beaucoup de de pour cette école
bien euh c'est - enfin ça devient disons de plus en plus difficile et alors y a y a eu un
..... à partir de cette année c'est la première année où il y en a un bon moi j'y suis
..... je suis en troisième année cette année alors je suis sans
concours simplement avec un c'est-à-dire euh.....
..... au mais euh l'année dernière donc il y a eu
un concours parce que ça devient de plus en plus difficile et il faut trier

ACTIVITY 2 : TRADUCTION

Traduisez les phrases suivantes en anglais, en utilisant le texte de l'interview si cela s'avère nécessaire :

I am in my first year of a degree. My course lasts three years. I have chosen a course which allows me to study a foreign language.

.....

.....

.....

ACTIVITY 3: VOCABULARY

One of the key elements in understanding a text is vocabulary. The more words one knows, the easier it is to understand a text. However, it is not just a matter of learning the dictionary. Every language has a relatively small number of core words which are highly frequent. Many of these are so-called “grammatical words” such as articles and pronouns and make up no more than a few hundred in total. They are fixed in number and can be grouped together in “sets”. They are often very different from one language to another and their description in dictionaries often appears inadequate. The meaning of these words cannot be disassociated from their place and function in the grammatical system.

In addition to this core vocabulary, there is a general vocabulary which can amount to tens of thousands of words. Some of these are highly specialised and are unlikely to occur in anything but a specialist text. You may well not know what they are in English, let alone French, unless you are interested or work in in a specialist field (eg a five-bearing crankshaft or a lesser-spotted woodpecker). However, there are between one and five thousand high-frequency words in this category which you will need to master if you are to read efficiently the great majority of everyday texts.

Core vocabulary

In the interview with the student, there are a number of grammatical words which belong to closed sets. One of these sets is complete:

un	des
une	

However, most of the others are not but knowledge of some of the items presupposes knowledge of the others. Fill in the gaps in the boxes below:

	l'	les
la		

	mes

ce /	
cette	

	tous

j' / je	nous
	vous
il /	/

me	
le / la / l'	

moi	

The next two sets are not strictly grammatical words but are so frequent that they are part of the core and must be learnt by heart:

suis	

a	

infinitive :

Below are more words taken from the interview which form a set.

premier / première	deuxième	troisième	dernier / dernière
--------------------	----------	-----------	-------	--------------------

In the box below, fill in the gaps to complete another set

juillet	août				

Below is a list of more core vocabulary items. Insert the English equivalents:

Pos.	Word	Frequency	%	norm
6	on	13	3.3592%	03.4 per 100
25	ça	3	0.7752%	00.8 per 100
139	tout	1	0.2584%	00.3 per 100
7	à	12	3.1008%	03.1 per 100
13	en	8	2.0672%	02.1 per 100
42	avec	2	0.5168%	00.5 per 100
46	dans	2	0.5168%	00.5 per 100
77	après	1	0.2584%	00.3 per 100
85	comme	1	0.2584%	00.3 per 100
131	pour	1	0.2584%	00.3 per 100
132	sans	1	0.2584%	00.3 per 100
134	sur	1	0.2584%	00.3 per 100

8	et	11	2.8424%	02.8 per 100
115	mais	1	0.2584%	00.3 per 100
121	ou	1	0.2584%	00.3 per 100
122	où	1	0.2584%	00.3 per 100
123	parce que	1	0.2584%	00.3 per 100
61	peu	2	0.5168%	00.5 per 100
81	beaucoup	1	0.2584%	00.3 per 100
38	très	3	0.7752%	00.8 per 100
78	assez	1	0.2584%	00.3 per 100
84	combien	1	0.2584%	00.3 per 100
9	alors	10	2.5840%	02.6 per 100
93	donc	1	0.2584%	00.3 per 100
136	surtout	1	0.2584%	00.3 per 100
50	ensuite	2	0.5168%	00.5 per 100
62	puis	2	0.5168%	00.5 per 100
95	encore	1	0.2584%	00.3 per 100
96	enfin	1	0.2584%	00.3 per 100
112	là	1	0.2584%	00.3 per 100
143	voilà	1	0.2584%	00.3 per 100
57	même	2	0.5168%	00.5 per 100
67	tous	2	0.5168%	00.5 per 100
79	autre	1	0.2584%	00.3 per 100
16	(il) y a	5	1.2920%	01.3 per 100

Some very common words can only be understood and translated in context. How would you translate the following examples ?

is étudiante en kinésithérapie à Clermont-Ferrand hein euh alors voilà c'est tout
urtout euh indispensable c'est-à-dire l'anatomie la physiologie bon alors ça à fond
ie la physiologie bon alors ça à fond et puis ensuite à la fin de la première année
lors ça à fond et puis ensuite à la fin de la première année on commence à aller to
la première année on commence à aller tous les matins à l'hôpital - et à travaille
mmence à aller tous les matins à l'hôpital - et à travailler avec des gens sur les
oup de de concours pour entrer à cette école bien euh c'est - enfin ça devient diso
t alors y a y a eu un concours à partir de cette année c'est la première année où
ent avec un baccalauréat c'est-à-dire euh un certificat de fin d'études au lycée
un enseignement très théorique au départ - on apprend bon ce qui est euh surtout eu
e l'école qui part en vacances au mois de juillet et l'autre moitié au mois d'août
un certificat de fin d'études au lycée mais euh l'année dernière donc il y a eu un

tout et vous faites des études de combien d'années alors là mes études durent trois
alors parlez-moi un petit peu de la formation de - kinésithérapeute oui c'est (rir
i un petit peu de la formation de - kinésithérapeute oui c'est (rire) alors euh bie
fond et puis ensuite à la fin de la première année on commence à aller tous les ma
n est à l'hôpital on a un mois de vacances pendant l'été on on intercale y a la mo
on on intercale y a la moitié de l'école qui part en vacances au mois de juillet e
e qui part en vacances au mois de juillet et l'autre moitié au mois d'août et puis
ller est-ce qu'il y a beaucoup de concours pour entrer à cette école bien euh c
'est - enfin ça devient disons de plus en plus difficile et alors y a y a eu un con
a y a eu un concours à partir de cette année c'est la première année où il y en a
c'est-à-dire euh un certificat de fin d'études au lycée mais euh l'année dernière d

lors euh moi je suis étudiante en kinésithérapie à Clermont-Ferrand hein euh alors
en organisé comme enseignement en France la kinésithérapie alors parlez-moi un peti
la moitié de l'école qui part en vacances au mois de juillet et l'autre moitié au
moitié au mois d'août et puis en troisième année c'est la même chose on continue e
nfin ça devient disons de plus en plus difficile et alors y a y a eu un concours à
'est la première année où il y en a un bon moi j'y suis rentrée je suis en troisièm

FRE1002: FRENCH GRAMMAR AND USAGE

UNIT 2B: CONCEPTS AND STRUCTURE 2

SECTION 1: READING: *L'ESCAPADE D'AMOUR DE MICHEL ET BABETTE..*

Read the text below and answer the following questions:

- 1) How would you describe the roles of Michel and Babette?
- 2) What is the target audience of the article?
- 3) What is the hidden message or (in marketing terms) the 'promise' ?

L'ESCAPADE D'AMOUR DE MICHEL ET BABETTE SARDOU

Voici deux mois Podium vous présentait une interview exclusive de Michel dans laquelle il niait fermement la rumeur qui courait sur son éventuel divorce. Babette était d'ailleurs venue le rejoindre à Tours, au Château d'Artigny.

Pour confirmer que leur amour n'a subi aucun nuage, Michel a décidé d'emmener Babette pour une seconde lune de miel à Venise.

Ce voyage Michel l'a décidé sur un coup tête, il le fait pour tout. Un matin il est arrivé chez lui, avec en poche deux billets pour Venise, en disant à Babette: «Prépare tes valises, nous partons demain soir, pour une semaine à Venise, je t'offre cette seconde lune de miel pour nos quatre ans de mariage!»

Ils ont traversé Venise en gondole... noblesse oblige... Ensuite promenade le long des rues. Ils firent ensuite du lèche-vitrine. Michel offrit une montagne de cadeaux à Babette. Des cadeaux qui

auront toujours une grande valeur sentimentale aux yeux de Babette, puisqu'ils lui rappelleront à jamais ce merveilleux voyage d'amour.

Michel était rayonnant de bonheur, Babette riait aux éclats. Ils se sont amusés à danser une valse dans la rue. Le soir, dîner en tête à tête dans les petites «Trattoria» qui bordent le grand canal. Visite des musées, séances de photos au milieu des pigeons. Une semaine de repos, de détente, d'amour pour Michel et Babette. Michel apprécia d'autant plus cette escapade que, dès sa rentrée, il devait affronter le public parisien au Palais des Congrès. Mais il réserve une autre surprise à Babette: passer tout l'été à ses côtés en Sardaigne.

Nous ne leur souhaitons qu'une seule chose: que leur amour soit éternel... WAOW!!!

PODIUM

SECTION 2: LEXIS: *L'ESCAPADE D'AMOUR DE MICHEL ET BABETTE..*

The column on the left contains the 50 most frequent words from the text on page 1 (total: 282 words) and the column on the right the equivalent number from a corpus of 291 texts (total: 381629 words). What are the words that stand out in the left column because they are missing from the right column? What clues do they give to the content of the text?

Word	Frequency	Word	Frequency
1 de	10 03.5461%	1 de	22273 05.8363%
2 à	10 03.5461%	2 la	13168 03.4505%
3 une	9 03.1915%	3 l	9848 02.5805%
4 Babette	8 02.8369%	4 et	9405 02.4644%
5 pour	7 02.4823%	5 le	8422 02.2069%
6 Michel	7 02.4823%	6 des	7573 01.9844%
7 le	6 02.1277%	7 à	7404 01.9401%
8 d	6 02.1277%	8 les	7382 01.9343%
9 en	5 01.773%	9 d	6694 01.7541%
10 il	5 01.773%	10 en	6317 01.6553%
11 Des	5 01.773%	11 du	4993 01.3083%
12 Ils	4 01.4184%	12 un	4431 01.1611%
13 Venise	4 01.4184%	13 une	3708 00.9716%
14 amour	4 01.4184%	14 est	3692 00.9674%
15 tête	3 01.0638%	15 dans	3081 00.8073%
16 a	3 01.0638%	16 qui	2964 00.7767%
17 dans	3 01.0638%	17 A	2699 00.7072%
18 que	3 01.0638%	18 pour	2675 00.7009%
19 leur	3 01.063%	19 au	2633 00.6899%
20 au	3 01.0638%	20 que	2368 00.6205%
21 qui	3 01.0638%	21 par	2309 00.605%
22 lui	2 00.7092%	22 il	2149 00.5631%
23 sur	2 00.7092%	23 plus	1984 00.5199%
24 l	2 00.7092%	24 sur	1888 00.4947%
25 décidé	2 00.7092%	25 France	1801 00.4719%
26 nous	2 00.7092%	26 avec	1521 00.3986%
27 miel	2 00.7092%	27 ou	1488 00.3899%
28 lune	2 00.7092%	28 s	1449 00.3797%
29 semaine	2 00.7092%	29 son	1356 00.3553%
30 seconde	2 00.7092%	30 se	1321 00.3461%
31 était	2 00.7092%	31 aux	1306 00.3422%
32 Ensuite	2 00.7092%	32 qu	1303 00.3414%
33 deux	2 00.7092%	33 ce	1299 00.3404%
34 soir	2 00.7092%	34 pas	1206 00.316%
35 la	2 00.7092%	35 sont	1179 00.3089%
36 Ce	2 00.7092%	36 Français	1098 00.2877%
37 cette	2 00.7092%	37 elle	1065 00.2791%
38 tout	2 00.7092%	38 ne	1000 00.262%
39 Un	2 00.7092%	39 ses	989 00.2592%
40 aux	2 00.7092%	40 Mais	976 00.2557%
41 voyage	2 00.7092%	41 on	962 00.2521%
42 cadeaux	2 00.7092%	42 sa	920 00.2411%
43 vous	1 00.3546%	43 comme	905 00.2371%
44 qu	1 00.3546%	44 c	901 00.2361%
45 Voici	1 00.3546%	45 n	888 00.2327%
46 Palais	1 00.3546%	46 cette	855 00.224%
47 promenade	1 00.3546%	47 ont	818 00.2143%
48 public	1 00.3546%	48 pays	733 00.1921%
49 puisqu	1 00.3546%	49 aussi	722 00.1892%
50 ont	1 00.3546%	50 Paris	707 00.1853%

The majority of the words are so-called “function” or “grammatical” words. They are important because they are very frequent yet they rarely have one-to-one equivalents in English and their meanings are often defined by the other words they are used with. As a result, it is necessary to pay attention to the context in which they are found. Different contexts may well change the meaning. Below are the most common words. If you look them up in a dictionary, you will find them followed by an abbreviation (eg *def art* or *art def*). This tells you what category of word (or “part of speech”) each belongs to and gives important information about how the word is used. Another feature these words share is that they are mostly very short. This means that they are easy to overlook, yet they are often the most important words in the sentence. Ignore them and you may well lose the meaning.

Below are examples of certain words which have multiple meanings or functions and so can be particularly difficult to translate. How would you express the following constructions in English?

à

d'ailleurs venue le rejoindre à Tours, au Château d'Artigny. Pour confirmer que l pour une seconde lune de miel à Venise. Ce voyage Michel l'a décidé sur un coup t billets pour Venise, en disant à Babette: "Prépare tes valises, nous partons demain offrit une montagne de cadeaux à Babette. Des cadeaux qui auront toujours une grand e, puisqu'ils lui rappelleront à jamais ce merveilleux voyage d'amour. Michel était aux éclats. Ils se sont amusés à danser une valse dans la rue. Le soir, dîner en tê la rue. Le soir, dîner en tête à tête dans les petites "Trattoria" qui bordent le g il réserve une autre surprise à Babette: passer tout l'été à ses côtés en Sardaign e à Babette: passer tout l'été à ses côtés en Sardaigne. Nous ne leur souhaitons qu

d' / de

ventuel divorce. Babette était d'ailleurs venue le rejoindre à Tours, au Château d' rejoindre à Tours, au Château d'Artigny. Pour confirmer que leur amour n'a subi au i aucun nuage, Michel a décidé d'emmener Babette pour une seconde lune de miel à Ve à jamais ce merveilleux voyage d'amour. Michel était rayonnant de bonheur, Babette el et Babette. Michel apprécia d'autant plus cette escapade que, dès sa rentrée, il entait une interview exclusive de Michel dans laquelle il niait fermement la rumeur Babette pour une seconde lune de miel à Venise. Ce voyage Michel l'a décidé sur un ne de miel pour nos quatre ans de mariage! Ils ont traversé Venise en gondole... no ne. Michel offrit une montagne de cadeaux à Babette. Des cadeaux qui auront toujour e valeur sentimentale aux yeux de Babette, puisqu'ils lui rappelleront à jamais ce 'amour. Michel était rayonnant de bonheur, Babette riait aux éclats. Ils se sont am al. Visite des musées, séances de photos au milieu des pigeons. Une semaine de repo ilieu des pigeons. Une semaine de repos, de détente, d'amour pour Michel et Babette

des

e... Ensuite promenade le long des rues. Ils firent ensuite du lèche-vitrine. Miche montagne de cadeaux à Babette. Des cadeaux qui auront toujours une grande valeur se bordent le grand canal. Visite des musées, séances de photos au milieu des pigeons. s, séances de photos au milieu des pigeons. Une semaine de repos, de détente, d'amo

était

son éventuel divorce. Babette était d'ailleurs venue le rejoindre à Tours, au Chât eilleux voyage d'amour. Michel était rayonnant de bonheur, Babette riait aux éclats

l' / la / les

iel à Venise. Ce voyage Michel l'a décidé sur un coup tête, il le fait pour tout. U urprise à Babette: passer tout l'été à ses côtés en Sardaigne. Nous ne leur souhait ns laquelle il niait fermement la rumeur qui courait sur son éventuel divorce. Babe amusés à danser une valse dans la rue. Le soir, dîner en tête à tête dans les petit Babette était d'ailleurs venue le rejoindre à Tours, au Château d'Artigny. Pour con 'a décidé sur un coup tête, il le fait pour tout. Un matin il est arrivé chez lui, se oblige... Ensuite promenade le long des rues. Ils firent ensuite du lèche-vitrin danser une valse dans la rue. Le soir, dîner en tête à tête dans les petites "Trat etites "Trattoria" qui bordent le grand canal. Visite des musées, séances de photos a rentrée, il devait affronter le public parisien au Palais des Congrès. Mais il ré

leur

d'Artigny. Pour confirmer que leur amour n'a subi aucun nuage, Michel a décidé d'e es côtés en Sardaigne. Nous ne leur souhaitons qu'une seule chose: que leur amour s

lui

t. Un matin il est arrivé chez lui, avec en poche deux billets pour Venise, en disa ux yeux de Babette, puisqu'ils lui rappelleront à jamais ce merveilleux voyage d'am

n' / ne

Pour confirmer que leur amour n'a subi aucun nuage, Michel a décidé d'emmener Babe à ses côtés en Sardaigne. Nous ne leur souhaitons qu'une seule chose: que leur amou

FRE2001: FRENCH LANGUAGE 2

PRONOUNS : DEMONSTRATIVE PRONOUNS

There is a another type of pronoun, usually called demonstrative pronouns because they are related to the demonstrative adjective *ce*. They only exist in the third person and are gender specific.

The demonstrative pronouns are:

	masculine	feminine
Singular	celui	celle
Plural	ceux	celles

As you can see, they are a fusion of *ce* and the disjunctive pronouns *lui*, *elle*, *eux*, *elles*. In certain circumstances they are combined with *-ci* and *-là*.

Below you will find a number of examples of these pronouns in use. Identify the pronoun and the noun it refers to. How would you translate them into English?

1. Partis parmi les favoris, les Bleus sur sept matchs joués ont remporté sept fois la victoire, même si celle contre l'Italie a été acquise aux tirs au but.
2. La première Eurorégion. La région étant résolument tournée vers l'Europe, ses habitants tendent à ignorer les frontières. Celle avec la Belgique est si peu naturelle qu'il suffit parfois de traverser la rue pour changer de pays !
3. La Déclaration des droits de l'Homme de 1948 s'inscrit dans le droit fil de celle de 1789.
4. La loi du silence s'impose encore le plus souvent aux victimes, presque exclusivement des femmes. Celles qui osent porter plainte s'exposent à une procédure longue et semée d'embûches, sans parler du risque de perdre leur emploi.
5. L'office de tourisme de Marseille et celui d'Aix ont suivi l'itinéraire du peintre. Ils ont tracé, dans chaque ville, un circuit des plus belles lumières.
6. Très attachée à son patrimoine, la France s'intéresse aussi à celui d'autres pays.
7. En 1996, Rhône-Poulenc a réalisé un résultat net de 2,7 milliards de francs (soit 490 millions de dollars), supérieur de 28 % à celui de 1995.
8. ARTICLE 19 Tout individu a droit à la liberté d'opinion et d'expression, ce qui implique le droit de ne pas être inquiété pour ses opinions et celui de chercher, de recevoir et de répandre, sans considération de frontières, les informations et les idées par quelque moyen d'expression que ce soit.
9. Les aéroports de province connaissent également un bon dynamisme, notamment ceux de Nice et de Marseille qui profitent de l'essor des flux touristiques et souffrent assez peu de la concurrence du réseau TGV de la SNCF.
10. Aux Etats-Unis, il y a une très grande flexibilité du marché du travail. On débauche quand cela va mal, on réembauche quand il y a un mieux. Cela permet des taux de chômage inférieurs à ceux que connaissent les Européens, mais la médaille à son revers : de très fortes pressions à la baisse sur les salaires, ce qui n'encourage pas la consommation.

In the following examples the pronoun is combined with *-ci* or *-là*. What difference does this make? Are there any rules?

11. L'engagement européen de la France en est renforcé. En cette fin des années 90, le grand projet d'approfondissement de l'Union européenne est la monnaie unique, l'euro. Celui-ci verra le jour le 1er janvier 1999.
12. Le scrutin majoritaire a été instauré par le général de Gaulle afin de lutter contre l'instabilité gouvernementale qui avait sévi sous la IV^e République du fait notamment du scrutin proportionnel. Celui-ci a été rétabli pour les élections de 1986 par le gouvernement socialiste - qui recherchait une meilleure représentation des petites formations.
13. La IV^e République est alors près de sombrer dans une grave crise engendrée par la décolonisation. Celle-ci commence en Indochine, d'où la France doit se retirer après huit années d'une guerre difficile.
14. La France qui dispose de peu de ressources minières parvient néanmoins à assurer plus de la moitié de ses besoins énergétiques, grâce à sa production d'électricité d'origine nucléaire (celle-ci représente 75% de sa production électrique, au lieu du quart il y a vingt ans).
15. Journaux, télévisions et radios ne pourraient fonctionner sans les agences de presse. Celles-ci sont les grossistes de l'information: elles fournissent aux organes de presse et aux institutions des informations de toutes natures - textes, photos, graphiques - moyennant un abonnement.
16. Près de la moitié des ventes de disques et cassettes en France concernent les variétés internationales, tandis que 70 % des recettes des salles de cinéma proviennent de films étrangers. Ceux-ci sont le plus souvent américains.
17. La France compte 8,9 millions de familles dont 1,2 million de familles monoparentales. Parmi celles-ci, 86 % sont constituées par des femmes seules qui élèvent leurs enfants à la suite d'une rupture d'union.
18. Une telle «convention de la Constitution» découle de la structure de la V^e République, dont on a vu qu'elle repose sur la collaboration étroite du gouvernement et du Président. L'élection de celui-ci au suffrage universel (instaurée par le général de Gaulle en 1962) lui assigne un rôle directeur dans l'orientation de la politique générale du pays.
19. Indéniablement, ce sont les universités qui accueillent le gros du bataillon des étudiants étrangers. A la rentrée 1997-98, ceux-ci étaient au nombre de 125 000, soit près de 9 % des inscrits universitaires.
20. Il y a des artistes qui semblent réunir tous les dons, et Marie-Claude Pietragalla est sans doute de ceux-là : ex-danseuse étoile de l'Opéra de Paris, interprète des plus grands chorégraphes contemporains, elle est aussi chorégraphe et directrice de ballet.
21. Un million de disques de la Mama vendus en 1964. Comment naît un tube comme celui-là?

Now look at the next few examples. Do they behave in the same way? Can they be translated in the same way?

22. Sur le plan de la scolarité, les jeunes d'origine maghrébine prennent souvent de l'avance sur les jeunes Portugais pendant leurs études, mais ceux-ci retrouvent l'avantage au moment de l'insertion professionnelle.
23. La procédure législative française se caractérise par une égalité de principe entre le Sénat et l'Assemblée nationale avec néanmoins un avantage réel en faveur de l'Assemblée nationale puisque celle-ci peut, en cas de conflit entre les deux chambres - c'est-à-dire après deux lectures dans chacune d'entre elles et réunion d'une commission mixte paritaire - voter seule la loi, en dernier lieu.
24. On a pu croire que Jane Birkin, interprète exclusive en même temps que muse de Serge Gainsbourg, serait condamnée au silence à la mort de celui-ci.

Translate the following into French:

1. Leap Years (*années bissextiles*) occur every four years. We usually think of a year as being 365 days long. We call this a calendar year (*année tropique*).
2. A rail tunnel project similar to that abandoned in the 1970s offered the best chance of success.
3. The following titles were amongst those announced for autumn and spring.
4. A small part of the day is spent in educational classes — the children are divided into two groups for these according to age or ability.
5. The most mobile age group in this period were those aged 16 to 34,
6. Extra investment in regional rail networks was announced in addition to that for Cardiff or Swansea.
7. The classifications are the same as those used by the government.
8. One should remember that there are always problems in compiling (*établir*) figures such as these.
9. The judges adopted a different approach to that adopted by their predecessors.
10. If you make a mistake like that again, you'll lose your job
11. The region has various transport weaknesses. Many of these are brought about by the presence of the River Avon.
12. At the first session (*séance*), photos and videos are made. These are continuously updated (*mis à jour*).
13. He opened his bag and took out two small planes. 'These are for my sons,' he said.
14. The Assembly held its first meeting on the 31st September. This was followed by a second meeting on 15 October
15. Education was for the rich or those able (*doué*) enough to get scholarships (*bourses*).

PRONOUNS 3b: POSSESSIVE PRONOUNS

There is a final category of pronoun, called possessive pronouns because they are related to the possessive adjectives *mon, ma, mes*, etc. These pronouns exist for all persons (*je, tu, il / elle*, etc) although they function as third person pronouns in that they are followed by verb forms such as *est* and *sont*. They are gender specific though sometimes the only marker of gender is the definite article rather than the pronoun itself.

The personal pronouns are:

masculine	feminine	plural
le mien	la mienne	les mien(ne)s
le tien	la tienne	les tien(ne)s
le sien	la sienne	les sien(ne)s
le nôtre	la nôtre	les nôtres
le vôtre	la vôtre	les vôtres
le leur	la leur	les leurs

In the sentences below, underline the possessive pronouns and then underline the nouns they refer to :

- 1 « *Un maître d'armes, on le garde pour la vie,* » explique Anne-Lise Touya, la championne du monde de sabre. « *C'est lui qui vous enseigne la technique et façonne votre style. Le mien, je lui dois tous mes succès.* »
- 2 « *Il y a plusieurs façons d'envisager le métier de producteur. La mienne est de ne pas partir de l'idée qu'un film est une affaire commerciale, mais de me fier à mon instinct, à mon envie de travailler avec un metteur en scène et de réaliser un scénario.* » dit Marin Karmitz.
- 3 « *Composer pour des films, précise Petit, me permet de faire les preuves des différentes cultures musicales qui sont les miennes.* »
- 4 « *Et est-ce que tu cr...trouves que c'est difficile d'allier le travail à plein temps comme le tien avec une vie familiale puisque J.-P. travaille aussi beaucoup dans la journée?*»
- 5 « *Pour les gants de toilette, j'te donne tout ce que tu veux. Tu n'as pas besoin d'avoir les tiens.* »
- 6 ARTICLE 13.2. Toute personne a le droit de quitter tout pays, y compris le sien, et de revenir dans son pays.
- 7 C'est par hasard que dans le train tu as entendu quelqu'un parler d'un lieu nommé Lingshan, la Montagne de l'Ame. Cet homme était assis en face de toi, ta tasse à thé était posée à côté de la sienne et les vibrations du train faisaient tinter l'un contre l'autre les couvercles de vos tasses.
- 8 La République peut conclure avec les Etats européens qui sont liés par des engagements identiques aux siens en matière d'asile et de protection des Droits de l'homme et des libertés fondamentales des accords déterminant leurs compétences respectives pour l'examen des demandes d'asile qui leur sont posées.
- 9 L'horlogerie Saint-Honoré a doublé ses exportations en dix ans, la lunetterie a augmenté les siennes de 127 % sur la même période et le jouet les a multipliées par 17.
- 10 Nous devons inventer une nouvelle approche de l'économie capable de penser les interdépendances d'un monde complexe comme le nôtre.

- 11 Alors que la génération précédente peine à utiliser les ordinateurs, la nôtre exprime des sentiments avec.
- 12 Chaque pays véhicule des choses mystérieuses qui s'inscrivent inconsciemment dans les corps. Les danseurs américains ont une culture du mouvement différente de la nôtre.
- 13 Être un créateur français ou vivant en France est, en effet, très bien perçu en Asie ou en Amérique. Cela ne doit pas nous empêcher, bien au contraire, de rester très rigoureux par rapport aux normes de qualité et de livraison qui sont les nôtres.
- 14 Existe-t-il en Europe des journaux comparables au vôtre ?
- 15 Au chapitre religieux, on craignait les progrès de l'extrémisme islamiste, alors que les enfants de Portugais ou d'Espagnols pratiquent plus leur religion que les Algériens la leur : près de la moitié des immigrés algériens se disent sans religion ou non pratiquants.
- 16 Destabilisés par des luttes intestines entre des personnes ou entre des courants, les partis politiques d'alors ne se montrent jamais à la hauteur de la situation. L'échec de la IV^e République est aussi le leur.
- 17 Les éditions *Autrement* tentent de partager avec leurs lecteurs, les approches singulières et problématiques, les intuitions et les découvertes, les interrogations et les colères, les coups de cœur et les rêves qui sont les leurs pour rendre le monde, à l'aube de ce XXI^e siècle, un peu plus « lisible », un peu plus enchanté.

What sometimes has to be added when translating these pronouns into English ?

Now look at the next three examples. What is different about them?

- 18 Lors du procès d'un homme ayant vengé les siens victimes des pogroms en Ukraine, en abattant en 1926 un général qui y avait participé, le journaliste français Bernard Lecache lance une campagne d'opinion. Objectif : faire acquitter l'accusé.
- 19 Pour la Chine, dès la formation de la Première République, en 1912, la France a été une référence dans la mesure où cette nouvelle forme de l'Etat chinois a, pratiquement, fait sien la devise de la République française - Liberté, Egalité, Fraternité.
- 20 Les idées progressistes issues de la Révolution française ont joué un rôle prépondérant dans l'évolution des pays des Balkans. En Albanie aussi, l'intelligentsia a fait siennes les idées antiroyalistes qu'elle a appliquées contre l'empire ottoman.

Translate the following into French :

- 1 There are two coats left which look identical. If you can recognise yours, I shall know which (*lequel*) is mine.

.....

.....

- 2 When she got out of the train, she took my bag and left hers on the seat (*la banquette*).

.....

.....

- 3 The aim of travelling is to put oneself in a culture other than one's own.

.....

.....

4 The advantage of a job like theirs is the freedom.

5 If your car has broken down, why not borrow ours?

FOLLOW-UP ACTIVITY

Identify the demonstrative pronouns (with a circle) and possessive pronouns (with a box) in the following text and the words they refer to :

Notre pays aime les querelles qui tournent autour de sa langue : on l’a vu il y a quelques années avec la « querelle de l’orthographe ». Les débats autour de l’usage du français viennent nourrir discussions et forums et remplissent les pages des journaux. Cela prouve, s’il en était besoin, l’attachement de nos concitoyens à leur langue et le souci permanent du « bon usage » qui nous anime.

Parmi ces querelles prend place celle de la féminisation des noms de métiers, titres, grades et fonctions. Les linguistes le savent depuis longtemps : cette affaire n’est pas seulement la leur. Elle concerne la société tout entière. Elle véhicule nombre de résistances, pour une large part idéologiques.

Le rôle du Gouvernement ne peut certes pas être en la matière d’imposer une norme : la liberté d’expression, une des libertés les plus fondamentales dans une démocratie, suppose le droit pour chacun d’utiliser la langue comme il l’entend. Mais le Gouvernement doit montrer l’exemple dans la sphère qui est la sienne, celle des services publics. Qu’une femme exerçant les fonctions de directeur d’école porte depuis plus d’un siècle le titre de directrice alors que la femme directrice d’administration centrale était encore, il y a un an, appelée « madame le directeur » atteste, s’il en était besoin, que la question de la féminisation des titres est symbolique et non linguistique.

C’est la raison pour laquelle j’ai, par une circulaire en date du 6 mars 1998, invité les administrations à recourir aux appellations féminines pour les noms de métiers, titres, grades et fonctions chaque fois que le féminin était d’usage courant.

À ma demande, la Commission générale de terminologie et de néologie m’a remis un rapport portant au premier chef sur les usages juridiques. Celui-ci montre que lorsque les textes visent une fonction, et non la personne qui remplit cette fonction, l’emploi du masculin est conforme à la règle. La Commission invite à la rigueur dans la rédaction des textes législatifs et réglementaires. Je ne vois que des avantages à mettre en œuvre ces recommandations.

Le présent Guide, rédigé par l’Institut national de la langue française, montre que, contrairement à certaines idées reçues, il n’y a pas de difficulté à féminiser la plupart des métiers, grades, titres et fonctions. Il y en a d’autant moins que le français l’a fait couramment jusqu’au siècle passé. Je suis convaincu que ce guide sera utile à tous ceux qui souhaitent faire avancer la cause de la féminisation. D’ores et déjà, avec l’aide des médias, qui ont assimilé son sens, cette démarche progresse et les querelles sur « le » ou « la » ministre, lorsqu’une femme occupe ces fonctions, appartiendront bientôt au passé.

Notre langue évolue : elle n’est évidemment pas séparée des enjeux du temps. La parité a sa place dans la langue. Je souhaite que ce guide facilite une démarche dont la légitimité n’est plus à démontrer.

Lionel Jospin, Premier ministre

Translate the text into English

APPENDIX 4

Guide Sheets

FRE1002 FRENCH GRAMMAR AND USAGE

SKILLS WEEK

COMPUTER ACTIVITIES IN COTTESBROOKE

1 TELL Consortium software

GramEx French / GramDef French

These are both grammar practice programs. *GramEx* contains a variety of practice activities and tests whereas *GramDef* provides supporting explanations of grammar points.

TransIt Tiger

This is a translation practice program. A **Source text**, visible in the top half of the screen, is provided for you to translate in the bottom half of the screen. There is also a small **Glossary** button which will provide help for the more unusual words and a **Hints** button which provides extra help with comprehension as well as a few warnings. The program does not itself correct your translation but when you have finished, you can see a model translation (once you have asked for the password) to compare your version with.

2 Online via French Webboard

You can access a number of practice activities via the Webboard via the **Level One / Interactive Assignments / French Grammar and Usage** option. These are mainly links to external sites.

3 Concordancing with KWIC-CONCORD

KWIC-CONCORD is a program which enables you to search several hundred texts in a few seconds for words or phrases. This enables you to see words in context and provides the sort of help that you cannot always get from dictionaries.

Select a Folder with a reasonable number of files (\Label \ 25 – 48) and then perform the following searches:

- 1 on genders – enter individually the following forms : **age, *ée, *té, *ment*
then view the results, preferably after sorting on the keyword, looking for any regular patterns or exceptions
- 2 on auxiliaries – enter the patterns: *paru/e,s, apparu/e,s, *venu/e,s*
which are conjugated with *être* and which with *avoir* – any surprises?
- 3 on prepositions – enter : *espérer, réussir, commencer, persuader, encourager*
then enter : *France, Galles, Japon, Etats-Unis, Irak, Israël*
then enter : *en* (sort on first word to right when finished)
- 4 on vocabulary – enter *expérier, implémenter, abilité* - what do you discover ?
enter *change* (select the nouns – how does it differ from English?)

KWIC-CONCORD Text Concordancer

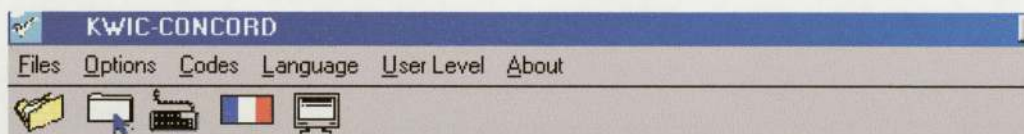
Guide sheet 1

BASIC OPERATIONS

KWIC-CONCORD works like any other concordancer. In order to perform a concordance search, you need to:

- 1) select the files you wish to search
- 2) select the drive and directory you wish to save the search results to
- 3) enter the keywords you wish to search for
- 4) confirm, if appropriate, that you want to save the results
- 5) start the search

To perform a search using KWIC-CONCORD, you need to carry out the following actions, either by clicking on the menus or, more simply, by clicking on the icons below the menu bar shown below.



1) Select source directory and files



These are the files to be searched. Click on the yellow folder icon and a file manager window will appear. This allows you to navigate your way through the disk structure on your machine in order to identify the source files you want to work on. Whether you click on 'All Files/Directories' (this will select all files not just in the directory shown but in all directories or subdirectories below it, thus enabling, if desired, a whole disk to be searched) or whether you select one of the subdirectories and then one or more files within that subdirectory will depend on how extensive you want the search to be.

2) Select target directory



The default is normally C:\TEMP. It's probably better to leave this as it is.

3) Enter keywords



Click on the keyboard icon to open the keyword entering window. This contains boxes for typing and listing your keywords as well as help with accented characters. The word or words you wish to search for are typed into the text box towards the top left of the window. You can enter any accented character by either holding down the 'Alt' key with one hand and typing in the code shown in the language help box with the other, using the numeric key pad on the right of the keyboard, or simply by clicking on the character in the language box. There are also a number of special characters that you can use. For more information, click on the 'Help' menu. Remember to click on 'Enter' after typing in a word.

4) **Performing the search**



Click on the 'Save to file' option before clicking on 'Search'. The default 'save as ANSI' should be left as it is. If you have entered more than one keyword, you have the choice of saving to a single file or not. This can be useful in the case of related items, "bien que" and "quoique" for example.

The results will be displayed on screen as the search takes place. You can pause the search in order to examine the results as they appear and you can also blank the screen by pressing 'Hide' if you find the scrolling movement distracting. Hiding the results will also speed up long searches.

5) **View search results**



Once the results of the search have been saved to file, they can be viewed by clicking on the monitor icon. You will be presented with the file manager screen. Once you have selected the file or files (all named with the extension "dmp"), the screen will change to display the first 22 lines of the concordance with the keyword in the middle. Depending on the number of lines and their length, you will be able to scroll either up and down or left and right to see more. To see the whole extract for a selected line, with details of the source file, either double click on the line or right click and choose 'View context'.

You can also re-sort the lines by choosing the first, second or third (etc) word to the left or right of the keyword. In a concordance based on "que", a sort on the first word to the left would group together all the instances of "alors que", "bien que", "pour que", etc.

There are various options available via the 'Options' menu to copy, delete, print and save all or part of the concordance in sorted or unsorted form (see Guide sheet 2).

KWIC-CONCORD Text Concordancer

Guide sheet 2

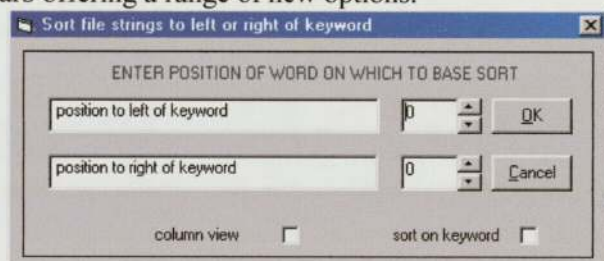
Manipulating search results

There are a number of ways of displaying and manipulating the search results on the viewing screen, all of them accessible from the 'Options' menu. The full range of options (cf contra) is only available if the 'User level' is set to 'Advanced'. Some of the choices are only enabled once a sort or deletion has been made.

Sorting

The most useful option is to perform an alphabetic sort. The norm is sorting based on the beginning of the word. However, it can be useful to group words with the same ending (eg *-age* or *-ait*) and this is done using the 'word end' option. The results of the sort can be saved as a text file for later editing or the status quo ante can be restored using the 'Unsort' option.

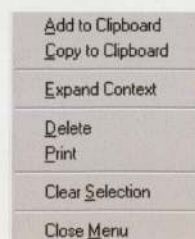
Once 'Sort Concordance' has been selected, a new window appears offering a range of new options.



Sorting can be performed on the keyword itself or on any word to the left or right. To make it easier to identify collocations, the results can be presented in column view, available either from the above window or from the 'Options' menu.

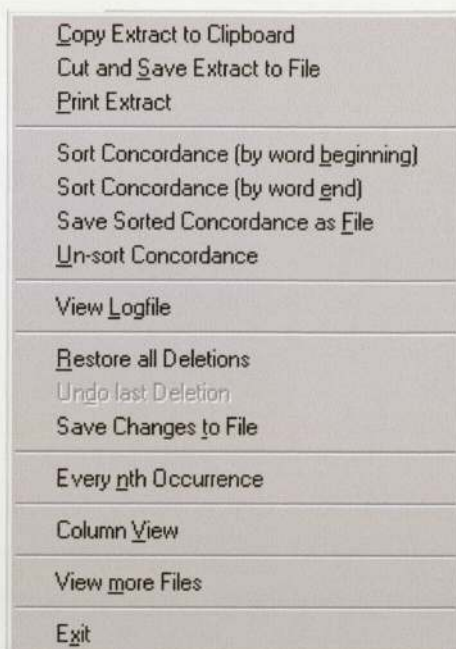
Selecting and copying to clipboard, saving to file, sending to printer, deleting

Selection can be made in the usual way by clicking on individual lines of text. Holding down the Shift key allows block selection while holding down the Ctrl key allows multiple selection. Once lines have been selected, right clicking on the screen brings up a small pop-up menu which allows execution of the operations detailed above. The pop-up menu clipboard options copy the whole line (keyword plus context) to memory unlike the options available from the main menu which allow a choice of either the concordance section only or the keyword plus full context. The 'Expand Context' option (which can also be executed by double-clicking on a selected line) brings up a new window on screen displaying the full line as saved with the number of characters to left and right of the keyword as defined in the default 'Context' panel. One of the new options offered is that of viewing the source file from which the text is taken, provided this is accessible.



Every nth Occurrence

Where a search produces a very large number of results, manipulation can be made more manageable by selecting a sample, eg every 10th, 20th, 100th example. To return to a full listing, reset the number to .



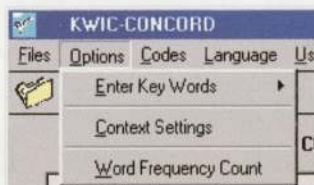
KWIC-CONCORD Text Concordancer

Guide sheet 3

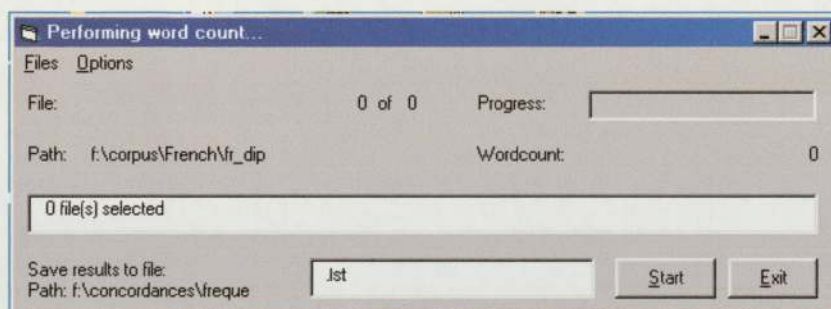
MAKING A FULL CONCORDANCE OF A FILE

1) Performing A Word Count

In order to make a full concordance of a file (or in fact any number of files), it is first necessary to make an alphabetical word list. This will provide a list of keywords to be loaded from file. To do this, select 'Word Frequency Count' from the 'Options' menu:



This will open a new window:



First select the source file or files using the 'File' menu. Be warned, a full concordance of a single large file or many short files can take a long time! Then select a target directory from the same menu by clicking on 'Save to directory...' You may be able to leave this as the default (c:/temp).

Once you have done this, simply enter an appropriate name for the word count file (this can be the same as the source file) and click on 'Start'. When the count has finished, click on 'Continue' and then click on 'Exit'.

2) Loading the word list

To load the list, click on the keyboard icon on the main window. Once the 'Enter keywords' window has appeared, click on the 'User Level' menu to check that this is set to 'Advanced'. Then click on the 'Options' menu and select the 'Load keywords from files' option. This will open the File Manager window allowing you to select the appropriate directory and within this the file with the extension 'lst'. The words will then be loaded and appear on screen. The top line will contain basic statistics about the file. Double click on this line. When it appears in the text box, replace it with a name that will allow easy identification and which, preferably, does not appear in the list of words. Press 'Enter' and then 'Ok'.

3) Performing the search

Perform a search in the normal way. Results will be saved to a file with the extension 'dmp' under the name entered on the top line of the word list.

KWIC-CONCORD Text Concordancer

WHAT MORE DO I NEED TO KNOW?

Types of files - You can perform searches on the following types of files:

- 1) ANSI or Windows text files Most text files these days use the ANSI code set.
- 2) ASCII or DOS text files. These tend to be older files created using MS-DOS programs.
- 3) HTML files. Most files downloaded from the Web use this format. Searches on HTML files tend to be slower because of the high number of formatting codes these contain.

The default is ANSI. KWIC-CONCORD can automatically detect HTML files but not ASCII files so if you are working on an older corpus developed using MS-DOS software you must select ASCII manually from the 'Codes' menu.

Search results should normally be saved as ANSI files as these are more easily read by other Windows programs.

A search will produce three linked results files:

- 1) a file with the extension "dmp" containing the results of the search in the form of the keyword in its immediate context, plus a short filename for identification purposes.
- 2) a file with the extension "log" which contains information on the source directory and file for each concordance entry.
- 3) a file with the extension "sum" which contains details of the distribution of the results.

The results can be edited using any wordprocessor and saved under a different name. Any changes to the "dmp" and "log" files themselves should be avoided as this may well make it impossible for them to be viewed subsequently using KWIC-CONCORD. If the "dmp" file is to be copied to a different disk or directory, the accompanying two files should be copied at the same time.

Wordcounts - Wordcounts can be performed or viewed by making the appropriate selection from the main 'Options' or 'Files' menus. Counts should only be performed on files saved in ANSI format. The results of each count are saved to two files:

- 1) a text file with the extension "lst" containing the word list in alphabetical order
- 2) a binary file with the extension "sta" allowing the results to be viewed and analysed.

Neither of these files should be modified though the word list can be imported into a wordprocessor.

Once it has been edited, it must be saved as a new file. Wordlists are not lemmatised, which means that only the spelling is taken into account. Therefore different forms of the same word (singulars and plurals for example) will be listed separately while different words with the same spelling ("bear" in English and "fait" in French) will be counted and listed as a single word.

Jeremy Whistle

APPENDIX 5

KWIC-CONCORD

KWIC-CONCORD

KWIC-CONCORD is a “streaming” concordancer that has been developed to run on Windows platforms. It has been tested on Windows 95, 98, ME, NT4 and XP and should also run on Windows 2000. Problems may however arise if a particular version of Windows has not been updated or does not contain the Visual Basic 6 run time files, especially msvbvm60.dll. The latter can be downloaded from the Microsoft website.

The program can be run from the enclosed floppy disk but will perform much more quickly if all files are copied to a new folder (e.g. c:\Program Files\KWIC-CONCORD) on the hard disk. It does not make any alterations to Windows registry and so can be removed simply by deleting all files. All the information the program needs is contained in the Concord9.ini file which will be updated each time the program is used. Help with running the program can be accessed in all windows via the ‘Help’ menu.

KWIC-CONCORD will process a virtually unlimited number of files (the only limitation being a maximum of 32767 in any single folder/directory). Because the program reads text in 32k chunks there is no limit on the size of files. It is recommended that word searches and frequency counts are performed only on standard text and HTML files. To read the latter the ‘wp#’ code files must be present in the program folder. Other text files such as Word documents will often produce perfectly acceptable results but searches performed using a wildcard mask such as “*.*” may encounter problems interpreting non-text and especially graphics and sound files. The program itself will filter out many non-text file types, such as EXE files, and newer file types can be excluded by adding the extension (dot plus three letters) in the last line of the Concord9.ini file.