Open Research Online

The Open University's repository of research publications
and other research outputs

A computer-based strategy for foreign-language vocabulary-learning

Thesis

For guidance on citations see FAQs.

Version: Version of Record

oro.open.ac.uk

A COMPUTER-BASED STRATEGY FOR FOREIGN-LANGUAGE VOCABULARY-LEARNING

ROBIN GOODFELLOW B.Ed, M.Sc
Institute of Educational Technology
Open University

Towards the degree of DOCTOR OF PHILOSOPHY
in the discipline of EDUCATIONAL TECHNOLOGY
at the Open University

October 1994

ACKNOWLEDGMENTS:

ABSTRACT:

This work sets out to establish principles for the design and evaluation of a computer-based vocabulary-learning strategy for foreign language learners. The strategy is intended to assist non-beginner learners who are working on their own, to acquire new words in such a way that they will be available when needed in subsequent communicative situations.

The nature of vocabulary-learning is examined from linguistic, psychological and educational perspectives, and a strategy for autonomous learning is derived which emphasizes the processes of: selection of new items from text, mental lexicon-building through the association of items on the basis of their lexical-structural features, and practising productive recall of items by activating the same associations as were used to build the mental network. This strategy is considered from the point of view of the support it would need from a computer-based interaction, and the field of Computer-Assisted Language Learning (CALL) for vocabulary is reviewed for examples of system design which meet the strategic and interactional requirements. Specifications are produced, based on general principles for the design of computer-assisted learning, and on current technological capability to integrate large text-databases and on-line lexical tools such as dictionaries etc., within an interface which facilitates learner control and exploration. Questions of evaluation are considered, in the light of the computer's ability to record interaction data, and a psycholinguistic model of word production is proposed as a basis for assessing the learner's performance in terms of processes as well as quantitative 'end product'. A general model of deep and surface approaches to learning is then adduced to provide a way of interpreting learner subjective data, and an independent means of evaluating the quality of the learning outcome.

A system implementing the strategy is tested with learners of Spanish and English, and the quantitative and qualitative data on learning process and outcome is analyzed in depth. The system is shown to support the learning objectives for learners who adopt a deep approach, or whose approach complements the assumptions of the design in some way, and the general design principles are therefore considered as validated. Some aspects of the strategy related to lexicon-building, however, are shown to be inadequately supported, as is the capability of the system to help learners remediate surface approaches. The main conclusion of the study is that, whilst learner exploration of powerful lexical information resources is essential for autonomous vocabulary-learning, on-line tutorial help of the kind that will encourage deep rather than surface approaches, is needed to optimise the quality of the learning outcome.

INTRODUCTION: BACKGROUND, RESEARCH QUESTIONS, STRUCTURE OF THE THESIS

0.1 Origins of the Research

The research described in this thesis has its origins in work that I pursued in the development of CALL (computer-assisted language learning) at the London School of English[1] between 1986 and 1988. Language-learning pedagogy at the LSE at that time was strongly influenced by principles of learner autonomy, collaboration and productive learning, and the CALL development effort was largely geared to the design and building of programs intended to support students, working alone or together, in producing their own learning material. Although the state-of-the-art in CALL at the time reflected a more traditional kind of teacher-student interaction (ie: what Underwood 1984 calls the 'wrong-try-again' model), there was already evidence, in programs like Wordstore[2], of a design philosophy based on the role of the computer as a learning, rather than testing, tool. Wordstore is a dictionary-building program, which enables the user to enter words and definitions, browse and search through them, and test themselves on their ability to remember a word from its definition. The learner is required to do the bulk of the work (selecting and looking up words and writing definitions) by themselves, with the program functioning mainly as an organiser. This approach fitted well with priorities at LSE regarding vocabulary as well as productive learning, as contemporary ideas on methodology in vocabulary teaching, (eg: Gairns & Redman 1986), stressed the importance of individualisation of content, and the processes by which learners built up their L2 mental lexicons. The idea of getting learners to work together to select and define vocabulary items, building a collaborative lexicon which could then be stored in a computer system easily accessible on a class or autonomous work basis, was seen as a desirable expression of the School's educational philosophy. There were, however, perceived inadequacies in Wordstore itself. An important characteristic of the mental lexicon was thought to be the existence of associative links between words (following Cruse 1986, Miller 1986 etc. - this issue is

---

[1]London School of English, 15 Holland Park Gdns, London W14 8DZ, is an EFL (English as Foreign Language) school specialising in Financial, Legal and Business English. The Principal is Timothy Blake.

[2] By Christopher Jones, published by WIDA Software

discussed more fully in the next chapter), as well as between words and their definitions. This implies a network model of the lexicon, rather than an ordered list of L1-L2 equivalents, but this is not reflected in the design of Wordstore. I consequently attempted to design and build a system which would be an advance on Wordstore, taking into account current theory on the structure of the mental lexicon, and the need to support more associative types of word-searching.

The system which I subsequently produced (called Lexis), turned out to have two distinct avenues of potential development. The first was as a general tool for collaborative language learning - its original purpose. An example of the way this might have been developed can be found in a description of a 'shared information' vocabulary system proposed in System magazine in 1988:

> "..The initial task of allocating vocabulary items to be researched by learners is one that the computer itself can handle by maintaining a transient file of items chosen by consecutive students, on a first come basis. The learner first checks to see if an item from the text is already in the database...if the item is not being researched by anyone else, the learner may 'book' it for research...learners enter their information into the transient file (meanings, grammatical details, examples of usage and so on)...the transient file is then merged in with the database, which is being built up on a weekly basis into a valuable store of vocabulary knowledge.." (Kukulska-Hulme 1988)

This broad conception of the computer as facilitator of group, productive vocabulary learning has remained current, with developments in the technology (eg: on-line lexicographic tools and resources) giving the idea even more potential, eg:

> "..the computer's main role would be to provide files and batteries of lexical information, each organised according to different principles...windows could provide simultaneous displays, giving sense relations, lexical set, collocations, word class, semantic features...learners could 'triangulate' on word meaning in a discourse context...they could withdraw or 'bank' information, and the support system of the computer could be used to assist writers as well as readers.." (Clarke 1993)

Lexis, however, did not develop into this kind of system (as Wordstore itself has not, nor has either of the systems envisaged in the above quotes yet been built). The main reason for this was that the practical (as opposed to the theoretical) educational culture at LSE, like many other language-teaching institutions, was not geared to the presence of computers in the classroom, so that their integration into a

general teaching methodology simply did not happen. To properly examine the reasons for this would require a separate thesis; suffice it to say that although considerable efforts were made to get a methodology of collaborative, productive, computer-based vocabulary learning accepted into general practice in the School, actual use of LEXIS tended to be by individual learners pursuing autonomous learning goals outside the classroom.

0.2 Autonomous Learning

The second avenue of potential development, therefore, and the one which was eventually explored and which has led to this present work, was as a tool for individual, autonomous learning, outside the classroom. By autonomous learning I simply mean the learner working alone, having to make their own evaluation of the success of whatever learning strategies they are employing. The practical need for such a mode of working at the LSE arose out of the size of the incidental vocabulary-learning task experienced by students following short, intensive, special purpose courses, such as English for Law, where the major part of class time was taken up with communicatively-based, subject-relevant activities such as discussions, presentations, and role-plays. In the normal course of general English teaching at the LSE, it was assumed that new vocabulary would be learned and practised in the context of this kind of communicative activity, but in the case of special purpose courses where the target vocabulary is more specific, the process is less reliable. It is very difficult to ensure, within a fairly natural communicative activity, that target vocabulary items receive adequate exposure, or that equal opportunity exists for all participants to practice them. In longer courses, these activities could be supplemented by more focused study-type learning, in which teachers could systematically cover any required vocabulary. In the short (one or two week) legal English courses, though learners encountered a lot of new and specialised vocabulary much of which was likely to be required for production at some point, teachers did not have the time to provide formal explanation-practice-testing sessions to ensure that all of it was internalised. It was therefore clear that learners needed to be helped to review this vocabulary individually and in their

spare time. Lexis was a tool available to be adapted to this purpose. The adaptation of the program involved switching the emphasis from those of its functions intended to support lexicon-building, to those aimed at word-retrieval. In other words, a return to a more conventional testing role. However, I felt that the design theory which supported its lexicon-building role, ie: the network aspects of the organisation of the target lexis, would still be relevant to the way it was used for word-retrieval, and I produced a system design aimed at encouraging the development of associative links in the mental lexicon as part of the recycling of a fixed set of target-words for productive practice.

0.3 Background - CALL for Legal Vocabulary

For the English for Law courses, a 'target-word' list of 54 common but specialised legal words and expressions was drawn up (derived mainly from contract law, items such as: privity, surety, covenant, recission, breach, deem, lessee, mandatory, lien, assign etc.), and an introductory class-based procedure devised as follows:

Learners were given a hard-copy list of the target-words (each supplied with a definition and an example sentence in English), and spent an hour browsing through them, without the aid of a dictionary or teacher explanation, trying to group them according to a range of criteria such as "words referring to people", "words referring to money", "nouns which can probably be used as verbs", "words which come from French" etc. The purpose of this was to familiarise them with the forms of the target-words, and with the idea of a range of possible mental association for foreign language words. The explanation that the teacher gave at this stage was kept deliberately slight, because it was felt that learners might otherwise be distracted from viewing the whole target set in terms of common general features. They were encouraged to review the whole list of 54 words, however incompletely, within the allotted hour. They were then introduced to Lexis (a half-hour session), and given an objective for the rest of the course, of achieving a score of 100% with the program, which would correspond to having correctly entered every one of the 54 items. The program functioned, initially, as a fairly straightforward 'type in the

word' exercise, where the learner could choose whether to be prompted by a definition, or by an example sentence (with the target-word gapped out). The prompts were presented in random order, and the learner could browse through them before selecting one to answer. There were, however, two additional features related to the associative network approach, and one inherited from the program's genesis in lexicon-building: a) the user could search for a target-word by entering a syllable, affix or group of letters - the program then displayed all the target-words which contained those letters; b) if the user typed in a guess which was not correct for the current prompt (ie: definition or example sentence), but which was nevertheless the correct form of one of the other target-words, then the program would indicate this and allow the user to cycle through a group of definitions, one of which was guaranteed to be the correct definition for the target-word they had typed in; c) when a target-word had been correctly retrieved, the option became available to edit its definition, to include any information (eg: an L1 translation) that the learner judged might ensure its retrieval in subsequent attempts. The learners' attention was drawn to these extra features in the induction session, and it was explained how they could use them to speed up the guessing process, as opposed to relying on the hard-copy 'crib' (which could nonetheless be used to identify a target-word in the last resort).

The principle underlying these aspects of the design of the interaction was that learners should be encouraged to view the separate things that they, and the program, did during an interaction, as part of an iterated process aimed at creating an active region in their lexical memory, which they could subsequently access for communicative purposes. This region, although centred on the target-word set, would also include words used in the definitions and examples, and any L1 translations or other information they provided for themselves. The strategy both for creating the region and for retrieving words from it, was to make guesses based on whatever associations came to mind (eg: a few letters, a word known to be connected etc.). By encouraging productive guessing (as opposed to laborious searching through the crib for the correct answer) it was hoped to activate associations at the level of relations between target-words, as well as between

individual words and their specific meanings. This higher level of association was expected to assist the search for words in any subsequent communicative practice. In addition, the interaction with the program would be more interesting, and therefore motivating, than a simple 'wrong-try-again' test, and allow the learner to vary the type of activity engaged in (browsing definitions and examples, responding to definitions, responding to examples, searching the target-list, matching definitions to a target-word, editing a definition etc.), and thus suit their particular learning style.

0.4 Research Questions

Lexis, with its accompanying procedure, was used for a number of legal English and other courses at LSE, over a period of a year. No formal evaluation was carried out, but from informal observation and questioning of learners, the following general assessments were made:

i) Most learners were happy to use the program in their spare time, and would spend 1 or 2 hours a week working with it. They appreciated the need for supplementary vocabulary study, and they viewed the program as a logical and interesting way of tackling the problem of learning large numbers of words in a short time. This assessment confirmed the potential of a LEXIS-type design for vocabulary-learning and was taken as sufficient justification for further research into its design principles. It was noted, however, that although target vocabulary word sets were produced for general English as well as the legal English courses, there was a much lower level of take-up and use of the program by teachers or students in the general courses. This appeared to point to a relation between the learners' perception of the role of vocabulary in their learning, and their willingness to pursue an activity focused on it. In the case of the legal English courses, as I have indicated, the learners were convinced of the need to concentrate on expanding their store of words related to the subject. This may have been because of a predisposition (as lawyers) to see words as important, or the obvious pressure imposed by the shortness of the courses they were attending. In the longer general

English courses, however, the vocabulary problem was subsumed into the general, situated, communication problem. Although these learners would readily admit that they needed more vocabulary of all kinds, they were less likely to be convinced of the importance of paying attention to the particular word sets chosen for use with Lexis. This raised a question about the capability of the program, with its fixed target-word sets, to support the necessary level of personalisation of learning content.

ii) Although learners were generally positive about their interaction with the program, claiming that it was helping to increase their vocabulary, most did not exploit its association-related features. This assessment raised the question of how learners were actually using the program, contrary to the assumptions made in the design. For example, they tended not to use the facility to browse through the prompts, preferring instead to try and answer the 'next' question (even though they were told it was randomly selected). They tended not to use the search facility, nor the 'find the correct definition' procedure, concentrating instead on making strenuous efforts to 'remember' the 'right answer'. Many of them spent time looking through the hard-copy target list to find the answers they didn't know; these learners said they thought the program should provide the right answers if requested. Such approaches seemed to undermine the expected value of the activity and the subsequent learning outcome. One problem was that the notion of the associative network, expressed in the word-grouping activity, was not an intuitive one for the learners. To them, the creation of associations between target-words was subservient to the need to grasp precise meanings in context. For this reason they did not explore the functions the program offered for reinforcing awareness of word features (the search facility, the 'find the right definition' procedure). Nor did they take much advantage of the facility to edit definitions, seeing it as an unnecessary step once they had correctly retrieved the target-word in question. In addition, the introductory classroom session did not satisfy the learners that they were pursuing a worthwhile activity. The number and difficulty of words on the list meant that the browsing approach simply confronted them with a large number of problems, which the grouping activity did not appear to them to adequately address.

In short, the learning strategy which the program design assumed and was built to support, was not the one adopted by the learners. This raised questions about whether the strategy was itself theoretically sound, and how the program design could make its aims explicit and its key procedures intuitive for the learners.

iii) Few learners managed to achieve the maximum score, and most did not manage to cover more than half the words in the target set. No way was found to demonstrate a link between learners' use of the program and subsequent improvement in their ability to use the target-words in classroom activities. This assessment drew attention to the need for a methodology of evaluation of this, and other, CALL programs in use. Several kinds of data are, in principle, available from the interaction, including comments by the learners themselves and the records of their actual performance. However, the collection of such data is always likely to be time-consuming, and it is important to establish in advance what it is going to be useful for, and how it can be validated. This question was therefore added to the research agenda which was emerging from the Lexis experience, and which was to determine the course of a fundamental redesign of the program. This agenda can be summarised as follows:

i) What is the nature of the content, in foreign-language vocabulary-learning?

ii) What kind of learning strategy is best suited to acquiring this type of content in an individualised, autonomous learning context?

iii) What are the requirements for a computer-based design to support such a strategy?

iv) How might the learner's approach influence the outcome of their interaction with such a design?

v) How can we evaluate such an interaction and its outcome, in practice?

An investigation of these questions was seen as pre-requisite to the development of an optimal computer-based vocabulary learning strategy geared to an individualised, autonomous learning situation. The thesis of the following work is that the principles of learner autonomy, lexicon-building, productive learning and the associative network, are theoretically well-grounded and can support a rational design for a learning system. I also want to show that the apparently unpredictable ways learners approach interacting with such a design, are in fact amenable to systematic description which may be incorporated into the operating principles on which the design is based. Finally I want to suggest a way forward for research into the recording and evaluation of computer-based language-learning interactions.

0.5 Structure of The Thesis

In Chapter One, a strategy for autonomous vocabulary-learning is described, based on a review of the literature in the fields of linguistics, second-language acquisition and second-language learning. The nature of the content in vocabulary-learning is considered, together with the kinds of teaching and learning approaches which this has given rise to, and appropriate quantitative and qualitative objectives for a strategy which has a vocabulary-learning aim. Chapter Two looks at the state-of-the-art in computer-based vocabulary-learning, and proposes design principles for a system which addresses these requirements. It discusses the constraints that technology and design impose on both aims and processes for the vocabulary-learner. Chapter Three considers principles for evaluation of learner interaction with the system. A psycholinguistic model of the processing which is assumed to underlie the building up of an L2 'mental lexicon', is proposed, and a framework for assessing the learning outcome is discussed. Chapter Four proposes functional specifications for an implementation of the design, and describes two pilot tests and their results, concluding with refinements to both design and evaluation methodology. Chapter Five presents quantitative and qualitative results from a programme of testing with learners of English and Spanish. The results are considered from the point of view of what they reveal about the learners' performance and approach, and also what they indicate about the success of the

design in promoting its learning objective. Chapter Six summarises the argument, presents the conclusions of the investigation, and looks at directions for future research.

CHAPTER ONE: PRINCIPLES FOR A VOCABULARY-LEARNING STRATEGY - BASES IN LINGUISTICS AND SECOND LANGUAGE LEARNING AND ACQUISITION

Introduction:

In the previous chapter I established that this research is concerned with learners who are adult non-beginners studying autonomously (without the benefit of human teachers or co-learners), with the aim of increasing their knowledge of vocabulary in their target language. The purpose of this chapter is to determine the nature of the content of such a task, so that objectives for a strategy may be set, and the kind of learning processes that this content implies, so that the strategy can be defined. In Section 1, I will review some of the literature addressing the function of vocabulary as part of the linguistic system, and consider whether there is sufficient structure there, to support a principled approach to learning it. Section 2 looks at some of the different approaches that have been taken to the learning and teaching of vocabulary, and assesses how far they are consistent with the content as described, and how far they support the aims of autonomous learning. In Section 3 I will review some of the research into the psychology of second-language vocabulary acquisition, and discuss the unconscious processes and mechanisms which operate in the learning of words. Finally, in Section 4, I will synthesise these findings into a statement of objectives and procedures for a strategy for autonomous vocabulary-learning.

Section 1: Vocabulary and Linguistic Description

1.1 The Neglect of Meaning - The Open Choice Principle

If we take the usual description of language (eg: Wilkins 1974, ch.1, Miller 1986, p.171) as having four main component systems: a grammatical system, a socio-linguistic system, a lexicon, and a phonological system, then it is clear that when we are talking about vocabulary we are referring to the lexicon, or to the system which organises units of meaning. Now, according to Crystal (1971) - and the view is supported by Levenston (1979) and Meara (1984) - meaning as a field of study is

of relatively late concern to linguists. Crystal traces the development of modern linguistic science (op cit. ch.4) from the study of phonetics in the 1930's (eg: Bloomfield 1933), via phonology, morphology and surface syntax, to deep syntax (Chomsky 1957) and finally to semantics in the 1960's (eg: Fillmore 1968). It was, he says, because of a too-general level of discussion in much of the historical work in the field of semantics, and because of the dominant influence of behaviourist psychology, that linguists of the structuralist persuasion came to insist that meaning is 'extra-linguistic', an internal phenomenon not susceptible to direct investigation (Crystal op.cit, p.209). As he describes it, the post-Bloomfield and pre-Chomsky schools of enquiry, (1930 -1960), provided linguistic descriptions and classifications which were structural almost to the point of "eliminating all mention of meaning". The Chomskyan (post 1957) position, positing meaning-related 'deep structures' underlying surface syntax, was developed partly in reaction to this. However, word-meaning in this formulation is still subordinate to syntax, reinforcing the view that the grammar is the natural home of generalisation in linguistic analysis, and that the lexicon is important mainly for its idiosyncratic content. Owen (1993, pp.168-169) gives the following quote as an example:

> "[the grammar]...formulates the rules to which lexical items may occasionally be exceptions, the generalisations which hold 'unless otherwise stated'. If the lexicon is the home of exceptions and idiosyncracies, the grammar is the home of significant linguistic generalisations..." (Smith & Wilson 1979, p.58)

The effect of this focus of theoretical linguistics on grammar has been to confer a 'slot-filler' role on the lexis, what Sinclair (1991, ch.8) has called the 'open choice principle', which means that, at any moment of word-selection in the construction of a text, there is a very large number of potential lexical choices constrained only by the requirements of grammaticality, independent of any organisational principles within the lexicon itself.

1.2 Structure in the Lexicon - The Idiom Principle

The open choice principle is challenged if the level of description is shifted from facts about grammatical structures to facts about words, as characterised by the

Firthian tradition (Firth 1957), which has consistently paralleled the mainstream since the 1930's. This tradition was given additional impetus in the 1960's, by the work of Halliday (1966) and Sinclair (1966), and more recently by the development of computational techniques of textual analysis. The approach demonstrates that words occur in the combinations they do, for reasons of meaning as much as of syntax. Sinclair, working from computer-analysed corpus data which reveals the forms and combinations in which individual words occur in use, proposes that there is a close correlation between the senses of a word and the structures in which it occurs in discourse.

> "It seems that there is a strong tendency for sense and syntax to be associated. The correspondences are overwhelming...[This study]...supports the contention that adjustment of meaning and structure is a regular feature of a language..."
> (Sinclair 1991 op.cit p.65).

Statistical evidence, he claims, shows that there is an "..underlying rigidity of phraseology.." in which grammatical structure overlaps with semantics. For example, the meaning of the word 'back' as referring to a part of the human body is far less common than its adverbial sense. This means the word is much more likely to be found in combination with verbs of motion, than, for example, with pronouns. Sinclair says that there is tendency for the most frequent words to have less clear and independent meanings (eg: '..take..', '..get..', '..give..' etc.), and therefore to occur in a wider variety of collocations. Most normal text is made up of the occurrence of these frequent (and less independent) words in conjunction with the more frequent senses of less frequent words. He puts forward the view that text is thus constructed mainly on an 'idiom principle', whereby lexical selection is made on the basis of a lexicon organised around pre-constructed phrases operating as a single choice.

1.3 Structure in the Lexicon -Lexical Semantics

Another view of the lexicon as inherently structured by relations of meaning, has been put forward by Fillmore (op.cit), Cruse (op.cit), Hudson (1990). This view has acquired close associations with work being done in cognitive psychology on the

13

structure of conceptual memory and the mental lexicon. Miller & Fellbaum (1991), theorising from observation of how dictionaries are constructed, suggest that we have 'dictionaries in the mind' organised as matrices of lexical concepts and defining phrases, relating to different subjects, otherwise known as semantic fields. A semantic field is a set of words whose decompositions share at least one nuclear concept (Miller 1986, op.cit p.176), eg: the semantic field of motion verbs, 'walk', 'run', 'fly', 'swim' etc. share an atomic concept 'change-of-location-over-time' (Miller & Fellbaum op.cit, p.176). Evidence of how these fields are organised can be found in word association experiments, in which there is considerable agreement among subjects in the words which are produced in response to common stimuli. Strongest of all associations are those existing between words which exemplify certain core relations such as antonymy, hyponomy, synonymy and meronymy (the part-of relation). Miller & Fellbaum extend the principle to include relations pertaining to verbs, as well as the more usually cited nouns and adjectives, eg: troponomy (the 'manner' relation which exists between 'walk' and 'run'), entailment (the necessary relation between 'snoring' and 'sleeping'), and causality (op.cit, pp.215-223). They put forward the idea that different parts of the verb lexicon have different structures, and that the different syntactic behaviour of different verb-classes can be related to their semantic components (op.cit, p.224), a position sympathetic to Sinclair's view (1.2 above), though arrived at via a different route.

1.4 Lexis in Linguistics - Summary

The view of the lexicon as composed of a vast number of essentially unrelated and idiosyncratic entities has been responsible for a comparative neglect of vocabulary in descriptive linguistic research, which favours the more generative kind of structure found in the grammatical and phonological systems. This view has, however, been challenged from a perspective which emphasises the psychologically and statistically salient patterns which can be found in semantic relations between the senses of individual words, and in features of lexico-grammatical collocation among groups of words. The 'vocabulary-content' of a text, from these points of view, is structured in at least two ways: a) in terms of alternative 'choices' for any

14

individual word-slot in the text, as determined by paradigmatic meaning and word-class (eg: the set of adjectives which might occupy the place of 'individual' in this sentence), and b) in terms of the 'forcing' of certain words to follow others, as determined by collocational convention (eg: the restriction on almost any other noun but 'view' following the pattern 'from these points of..' in this sentence). Although lexical relations like these are not generative in the same way that grammatical or phonological rules are, they are pedagogically much more promising than a list of exceptions and idiosyncracies. However, the espousal by language educationists since the 1930's, of a traditional view of vocabulary content, has largely set the agenda for the development of teaching approaches, and it is only within the last 20 years (Carter & McCarthy 1988 - introduction) that the potential of vocabulary as an organising factor in both language and language-learning has been recognised.

Section 2: Approaches to Vocabulary Teaching and Learning

2.1 The Neglect of Vocabulary in Teaching

The task for the learner of a language is to acquire knowledge of all the four component systems (see 1.1 above), and to integrate this knowledge for the purposes of comprehension and production. The corresponding task for the teacher is to structure the intended learning content for optimal acquisition and the practice environment for optimal integration. Since the learning/teaching process is a progressive one, acquisition and integration are usually treated, pedagogically, as separate steps. Similarly, because human cognitive processing capacity is not unlimited, knowledge acquisition is not usually attempted in all areas simultaneously, and some kind of explicit emphasis is given to one or other area on the basis of a theory about which is more important in terms of either its generative power or its value for communication. Decisions about how to structure practice environments are usually based on general psychological or educational principles or on practical (and often subjective) experience. Theories about how to structure content, however, tend to be founded on linguistic descriptions, such as the ones described in the previous section. The audio-lingual (phonology-based) and the

cognitive code (grammar-based) learning approaches are examples, their rationales based respectively on structuralism (eg: Fries 1945) and generative grammar (Chomsky op.cit). The functional approach (socio-linguistic-based) is another example, drawing on the work of Halliday (1978) among others. Vocabulary (the lexicon), on the other hand, has been something of a maverick area. Although, of the four components, it is arguably the most immediately apparent to the naive enquirer, until recently it was regarded by linguists as the least systematic (see previous section). This paradox has been reflected in a comparative neglect at the pedagogical level:

> "...It is the experience of most language teachers that the single biggest component of any language course is vocabulary. No matter how well the student learns grammar, no matter how successfully the sounds of L2 are mastered, without words to express a wide range of meanings, communication in L2 just cannot happen in any meaningful way. And yet vocabulary often seems to be the least systematised and the least well-catered for of all the aspects of learning a foreign language..." (McCarthy 1990, introduction)

The reasons for this neglect are to be found in the traditional, phonological and syntactical, approaches to description that mainstream linguistics has taken, and in the consequent difficulty of finding principles on which to base a theory of lexical acquisition or a methodology for vocabulary teaching. But one consequence of such neglect, for the vocabulary-learner, is that they are faced with a massive potential learning content and no apparent way of organising it. This perception of the learning task has been of considerable influence in the development of teaching approaches. Some have adopted an 'anti-instruction' position, arguing that explicit vocabulary instruction is counter-productive until basic syntactical patterns have been mastered (eg: Rivers 1968). Others (eg: Halliday et al. 1964) have tried to identify exactly which words a learner needs to know at each stage in their development. Yet others (eg: Judd 1978, Twaddell 1973) have simply assumed the need for 'massive' vocabulary instruction as early as possible.

2.2 'Anti-Instruction'

The view that there is little point in trying to teach vocabulary explicitly derives some of its justification from research in first language acquisition, which clearly favours linguistic exposure over instruction:

> "..The number of words to be learned is simply too large to be covered by instruction specifically aimed at teaching the meanings of individual words: only a small fraction of the words that must be learned could be covered. Furthermore, the rapid rate of vocabulary growth experienced by most children shows that explicit vocabulary instruction even at its best could account for only a small proportion of the words learned by the average child in a year. For many children, the speech of parents and peers may be the single most significant source of vocabulary growth.."
> (Nagy & Herman 1987 p.32)

A 'hands-off' approach to L2 vocabulary derives additional theoretical support from the notion of unconscious acquisition as described by Krashen & Terrell (eg: Krashen & Terrell 1983 - although Lewis 1993 p.41 points out that Krashen himself did not advocate a teaching approach which restricts vocabulary learning until after grammatical structures have been acquired). The theory puts the main emphasis on large quantities of comprehensible input as the most influential factor in language learning.

> "..The central hypothesis of the theory is that language acquisition occurs in only one way: by understanding messages. We acquire language when we obtain comprehensible input, when we understand what we hear or read in another language.." (Krashen & Terrell op.cit preface)

However, there are reasons to suppose that leaving vocabulary to be acquired in this way may not have the desired result, especially if that result involves learning for production. Laufer (1991) cites evidence to show that foreign-language learners, in the absence of specific vocabulary instruction, may not significantly increase their productive vocabulary at all, unless they are below average to start with. There is a recognised tendency for learners to favour simple, general and frequent words in production (eg: Blum & Levenston 1978), and this may lead to the establishment of a kind of 'vocabulary threshold' within the language-using community, which there is no communicative need to cross. Given that an L2 learner's opportunities to

17

encounter comprehensible input, in an autonomous learning environment, are considerably more limited than those of the L1 learner, these findings suggest that whatever is the L2 learner's equivalent to the '..speech of parents and peers..' may not be sufficient to support much vocabulary growth at all.

2.3 Lexicometric Selection

Another approach addresses the problem of the size of the learning task by looking to the field of lexicometrics for principles of vocabulary selection which will ensure that the learners' instructional exposure to vocabulary corresponds to what their natural exposure would be in the target language environment. It has been shown that there are distinct distributions of frequency of words in a language, and also that some more frequent words have a wide area of 'coverage' including applicability to the definition and explication of other words (West 1960). These criteria have been put to use in identifying which words should be taught; the General Service List (West 1953), for example, provides 2000 English words selected on the basis of their frequency of occurrence (in a corpus of 2-5m words), their utility in terms of the range of subjects they are relevant to, and their usefulness for the purpose of definition. The Threshold Level courses produced in the 1970's (Van Ek 1977), selected words on the basis of their membership of 200 semantic fields relevant to daily foreign-language functions, such as ordering meals in restaurants etc. Recent work in computer-based vocabulary analysis for secondary school courses in European languages by the Didascalia group (Decoo 1993) has extended the range to 5000 words, ordered on the basis of a comparison of frequency, utility, familiarity and coverage, and intended to be presented to learners over a 4-year period. This approach and the CALL system developed from it is discussed further in Chapter Two.

Although there are logical and practical reasons for requiring learners to acquire the most frequent and general words in a foreign language first, there is little evidence to suggest that this is what actually happens in informal, natural language acquisition (Levenston op. cit. p.151). In any case, frequency as a criterion for

selection quickly becomes irrelevant, as Twaddell shows (Twaddell 1976, pp.69-70), because although the most frequent words (eg: 'the', 'and', in English) occur around once every 15 words of running text, by the time we get to the 100th in the list ('down'), the rate of occurrence is only once every 1133 words, and the 1000th ('reach') occurs only once every 9,568 words. After that the frequencies soon become so low as to be insignificant. Ordering by frequency is therefore only sensible at the elementary stages of learning. Decoo also acknowledges that high frequency words alone do not make for other than "..colourless and artificial texts, impeding logical communication.." (op.cit, p.129) and for this reason advocates the use of 'frequency values' based on the formal, structural and semantic features of the occurrences of items. However, these added criteria involve '..a constant semantic reflection, taking into account the complexity of language functioning.." (Decoo op.cit p.126) which returns us to the problem of selection as originally stated.

2.4 Semantic Fields

The organisation of the mental lexicon postulated by Miller & Fellbaum (op.cit) has been used as a basis for structuring the presentation of vocabulary for learning. Target items are represented by the use of physical matrices in which words or lexical concepts are put along the top axis and senses or definitional phrases along the side axis. This has been seen in teaching approaches such as that typified by Rudzka et al. (1981), in which learners acquire the meanings of words in contrast to each other, via differential combinations of semantic primitives (eg: 'surprise' combines the concepts wonder/unexpected, as opposed to 'astonish' which combines wonder/difficult-to-believe). Channell (1987 p.121) claims that this approach partly supplants the use of synonyms and translations and leads to more precision in production and less avoidance.

The main problems that occur with this approach are connected with the intuition that is used to determine what the semantic components of individual words are. Channell, for example (op.cit, p.119), distinguishes the words 'surprise' and 'amaze' by the presence of the component 'so as to cause confusion' in the meaning of

'amaze' but not in 'surprise'. Similarly, in her illustration of the collocations distinguishing 'pretty' and 'charming' (op.cit, p.120) she claims that 'charming' cannot be used for a man, a picture or a present. It is doubtful whether such analyses would automatically be accepted by all native speakers. Nevertheless the insight into the way that lexis is networked, and the suggestion of psychological salience that accompanies certain relations (synonymy, antonymy etc.) has a strong basis in psychological research and has a clear application to the organisation of a pedagogical strategy.

2.5 Lexical Phrases

An implication of Sinclair's (op.cit) findings regarding the idiom principle, and of the claim that "..lexical patterning is very, very much more repetitive than open-choice models suggest.." (Owen op.cit, p.173), is that the language user has available to them "..a large number of semi-preconstructed phrases that constitute single choices, even though they may appear to be analysable into segments.." (Sinclair op.cit, p.110). Lewis (1993, p.90) and others (Nattinger & DeCarrico 1992, Cowie 1988) have applied this insight to the requirements of language teaching , proposing an approach which views underlying competence as neither grammatical, as Chomsky suggested, nor conceptual, as Miller proposed, but lexical:

> "..rather than relying on generative power of grammar...users rely on a vast store of fixed phrases and pre-patterned locutions by which they routinely manage aspects of interaction.." (Lewis op.cit, p.90 quoting Cowie op.cit)

The characteristics of lexical phrases, according to Nattinger & De Carrico (op.cit) are that they are readily accepted by native speakers as institutionalised sentences, that they reflect instantly identifiable pragmatic meaning, exhibit different degrees of variability, and cover a representative range of possible (grammatical) subjects.

One of the most important pedagogical implication of this analysis is that the learning task is greatly increased by the inclusion of lexical phrases. Pawley and

Syder (1983) have estimated that there are hundreds of thousands of 'lexicalised units' made up from the basic vocabulary, and if, as Cowie (1992) claims, creativity in production depends on these ready-made units, then the question of how they are learned must be answered, and the problem of how to teach them must be overcome. But how to develop a pedagogy which meets this challenge '..is a task which we are only now beginning to address..' (Cowie 1988, p.137). Lewis too (op.cit , ch.13), in his section on methodological implications of a lexical approach, is at pains to avoid recommending techniques, and instead emphasises language-rich materials and receptive skills practice, a position which seems little advance on the input-hypothesis-oriented approaches discussed in the previous section. Nattinger (1988, pp.78-80), proposes pattern drilling on the audio-lingualism model (although with a warning not to allow them to become '..mindless exercises..'), leading to controlled variation and then to analysis of syntactic structure:

> "..correctly identified lexical phrases can be presented to L2 learners in identifiable contexts, mastered as learned wholes, and thus become an important resource to mastering the syntax.." (Lewis op.cit p. 96 summarising Nattinger & DeCarrico op.cit)

He admits, however, that there are uncertainties in this approach, both in method and description (op.cit p80).

2.6 The Discourse Level

As the existence of lexical phrases only really becomes evident when fairly large tracts of discourse are examined, it may be that their importance in learning is related to other factors which operate at a discourse level. Such factors include register and the role of textual features such as connection, information weighting, and marking phrases, which help to organise, structure and evaluate the message. An example given by Carter & McCarthy are the phrases "..one significant difference.." and "..a central question..", phrases indicating serious polemical discussion (Carter & McCarthy p.209). To comprehend these aspects of meaning it is necessary to be able to understand not only the words and their co-text (ie: the words immediately surrounding them), but also their context, - the overall meaning

and intention of the discourse they occur in. This requirement gives the lexical approach a dimension which is lacking in the other approaches, ie: it bestows on vocabulary a role in the structuring of meaning at text-, as well as at word- and sentence-level. Although it is far from obvious how this idea is to be translated into pedagogical practice, it is clearly of importance as an additional level of structure by which to organise the impoverished input of the L2 learner, and it does strongly suggest that context is an indispensable consideration in the design of a vocabulary learning strategy.

2.7 Learner Skills

In contrast to approaches which are based on the attempt to specify which items are to be learned, or how they are to be presented, there is a learner-centred approach which dispenses with the idea that the most appropriate content vocabulary or the most effective means of presentation can be predicted for all learners and argues instead that they should be given the means to tackle the vocabulary problem for themselves. The issues raised by this approach tend to focus less around the nature of lexis itself and more on the learner's comprehension and the skills associated with reading and the guessing of unknown words from context. Twaddell (1973), for example, proposes the use of guessing strategies and tolerance of imprecision in the early stages of learning.

> "..What we can try to do is guide his development of skills to compensate for [the learner's] lack of resources, and let his resources grow as a result of his success in using his skills.." (Twaddell op.cit p.70)

He claims that there are degrees of 'familiarity and precision' (op.cit, p.72) of words even in L1, and that the learners ability to tolerate vagueness whilst considering contextual clues is the key to successful guessing and the eventual expansion of vocabulary. The approach leads him to advocate extensive practice at skim-reading and a 'complete boycott' (op.cit, p.73) on the use of translational equivalence. His assumptions about the lack of regularity in the relation between word forms and meanings (eg: see Twaddell op.cit, pp.65-66) echo the analysis that supported the anti-instruction view discussed in section 2.1., but where Twaddell differs from the

anti-instruction approach is in concluding that vocabulary expansion does need to be addressed directly, and in proposing that passive knowledge based on consideration of formal and pragmatic features of the printed word does eventually translate into learning (op.cit, pp.74-5).

2.8 Learner Knowledge

Description of lexical knowledge from the learner's point of view is taken still further by Richards (eg: Richards 1986). Drawing on lexis-oriented research in linguistics and psychology, he lists the different types of knowledge involved in native lexical competence, ie:

- Knowledge of word frequency and collocability. Native speakers have an intuitive knowledge of how frequent words in their language are, eg: the likelihood of encountering 'book' as opposed to 'directory' or 'thesaurus'. They also recognise collocation, which describes the most likely combinations of particular words within phrases. Such combinations are recognised where they take the form of fixed expressions, such as '..user-friendly..'; where they are selections from a restricted set, as such the choice of '..metal..lava..iron..' to go with '..molten..'; and where they are simply likely to co-occur within the same discourse, as is the case with '..collect..' and '..stamps..'. The validity of this type of knowledge can be demonstrated by the native speaker's instant recognition of unfamiliar collocates such as '..lying astray..', '..light possibility..' etc. (from Carter & McCarthy p.36, see also McCarthy p.14 regarding creative aspects of untypical collocation).

- Knowledge of register. Native speakers take into account aspects of fashion, geographical variation, social class, topic and mode of discourse etc.

- Knowledge of syntactic behaviour. Words have specific structural and grammatical properties, eg: some nouns can be pluralised ('foot/feet'), some cannot ('information'). Verbs can be transitive or intransitive, adjectives form comparatives and superlatives in different ways etc.

- Knowledge of underlying form. Words have root forms to which affixes can often be applied for the construction of other, derived, words (eg: 'solid', 'solidify', 'consolidate').

- Knowledge of typical associations. This includes semantic structural associations such as those Miller & Fellbaum (op.cit) describe, and also merely conventional ones (eg: 'accident/car').

- Knowledge of semantic value. Thesaurus-type information relating and distinguishing words by their shared semantic components (eg: 'damaged' has the component 'inanimate', whereas 'hurt' has components 'animate' and 'living').

- Knowledge of polysemy. Meaning is the product of a system of relations between a word and its surrounding linguistic and non-linguistic environment. Meaning therefore changes according to the way the environment is perceived. This can take the form of extension, eg: metaphorical use such as the comic '..dying on stage..', or polysemy, eg: '..I hurt my foot..' as opposed to '..at the foot of the stairs..'.

The extent and variety of the lexical knowledge he identifies, precludes any suggestion of a 'global' teaching approach (although he does advocate the cloze exercise, op.cit, p.88). But Richard's main contribution to the learner-centred view is to emphasise that the way in which a particular word is known by a learner is just as important as what the word is. Richards claims that a word is not a label but a "..process for dealing cognitively with the environment.." (op cit, p.83) . This has laid the ground for a more process-oriented approach, in which word-learning activities themselves are considered as significant as the state of knowledge to which they are supposed to lead.

2.9 Learner as Researcher

A development of the process-oriented approach is 'data-driven learning' (Johns 1991a & b, 1994 etc), in which the learner's ability to "..puzzle out how the

language works.." (1994) plays a role as important as the specific language features they learn. The starting point of this approach is the evidence provided by collocational data generated by a keyword-in-context (KWIC) concordancer - a computer-generated printout of every occurrence of a particular word, along with its accompanying context, in a large corpus of authentic text (see Fig 1.1). Johns gives the learner the role of "..a research worker whose learning needs to be driven by access to linguistic data..." (1991a, p.2) and argues that inductive learning strategies, such as perceiving similarity, discrimination, hypothesis forming and testing are all developed in the attempt to make significant generalisations on the basis of this data (1991a, pp.30-31). Although the point is made that once one starts to take seriously the ways in which words behave in context, the distinction between lexis and syntax starts to disappear altogether (1994), there is scope for vocabulary-learning, at the collocational level of description, in this approach, and further attention will be given to it in the next chapter.

Fig 1.1: Example of KWIC Output for the Word 'further'

```
 1 /2d.       Imperial Chemicals gained a further is. at 61s. 6d. and Typhoo Tea is. 3
 2 tion, and, as reported last week, a further +2Om. order was awarded to Babcock a
 3 e Chapel of the Shrine was erected. Further A34 rmation of its purpose was seen
 4 o show for it.       McKenna put Alloa further ahead in the second half in a breaka
 5 ely, some defence officials go even further and say that the original Sandys pol
 6 t that this meeting will lead on to further and decisive ones; but without it, w
 7 d the price rose to around 253s. on further buying.      Of the Eisenhower ban - a
 8  to send in more men, not weaken it further by desertion.      It is unjust to pas
 9  their claim for relief by means of further contribution from the Exchequer, tha
10 o full strength it is unlikely that further depletion of overseas garrisons can
11 be no hesitation in mobilising if a further deterioration in the situation warra
12 s and increased fees at colleges of further education outside Cheshire which som
13 1 be +398,500, and a new College of Further Education at Crewe costing +600,000
14 hire (Hartford) Central Colleges of Further Education at acost of +465,000 and +
15  andsecondary modern schools.      In further education, provision is being made f
16 ary, secondary and special schools, further education, the training of teachers,
17 y schools, and students undertaking further education.      The number of pupils o
18 as seven mills. In 1960 the company further extended its interests by becoming t
19 HNNY HAYNES is anxious to avoid any further half-time dressing room wrangles wit
20 cent motor shows abroad foreshadows further increases inactivity in the industry
21      Hopes will now grow brighter of further international co-operation, which is
22 ny's report centres chiefly on what further moves will be made to distribute som
```

Johns does not advance any research evidence for his processing claims, but the issues he raises, of learning strategy and processes of insight and retention, are closely related to questions recently under investigation in the field of second language acquisition research.

2.10 Section Summary

There is now a widespread recognition that individual learners have individual vocabulary needs not necessarily met by approaches which rely on generalisations about either the linguistic environment or the supposed most useful words in the target language (see Carter & McCarthy, 1988 op.cit, ch.3, for a review of developments in the teaching of vocabulary since the 1970's). New ways of describing the lexicon have brought with them new levels of understanding of the nature of lexical knowledge, and these in turn have introduced a pedagogy for vocabulary learning which draws on psychological as well as linguistic assumptions. These assumptions place the learner, and the development of the learner's own cognitive skills in the processing of new vocabulary, at the centre of the learning process.

A number of important implications thus arise for the design of a vocabulary-learning strategy for autonomous learners. The first is that the selection of learning content should be done by the learners themselves. This is because there are not sufficient grounds to pre-specify which words should be learned, or which words they already know, and because the identification of unknown or unfamiliar words is part of the processing which learning involves. Learners do, however, need to make their selection on the basis of encountering words in a meaningful context, such as a text which affords them some degree of comprehension. Secondly, content should not be restricted to individual words and their meanings. Lexical structure exists in paradigmatic and syntagmatic relations within the text, and where it is evident it should become part of the learning content. The development, by the learner, of an appreciation of this structure then becomes one of the overall objectives of the strategy. Third, the development of the learners' skills in dealing with vocabulary is also an important aim. Guessing meaning by inference from context, deconstructing and analysing words into their semantic and formal components, identifying common features among words, applying metalinguistic knowledge, skill in retrieving words from memory, etc. are all part of the learning content. These skills involve both conscious and unconscious levels of processing.

In the next section I will examine some of the relevant research into second language vocabulary-acquisition.

Section 3: Research in Second Language Acquisition

3.1 Background

Research into lexis in the field of Second Language Acquisition also suffered as a consequence of the grammar-centric perspective (Levenston calls it 'discrimination' - 1979, p.147), and much of the work that was done before 1979 focused on the influence of the first language on the development of L2 lexical forms. Although the data provided by such studies, eg: the classification of lexical errors caused by L1 interference, is inherently interesting, it has not, according to Meara (1984, p.226), been able to indicate any coherent developmental pattern or suggest any way of providing remediation of the more fundamental errors. However, Levenston's call in 1979 (op.cit, p.151) for research to take into account features of the learner such as motivation and previous knowledge, and features of the learning situation such as the degree of control and the specificity of the learning aims, succeeded in changing the focus sufficiently for a number of new and potentially applicable areas of research to emerge.

3.2 Mnemonic Strategies

Much of the psychological discussion about vocabulary-learning strategy has concentrated on mnemonic strategies. Although the empirical research is often inconclusive, a central distinction between implicit (unconscious)and explicit (conscious) learning has been clearly demonstrated (Ellis 1995 forthcoming). Implicit learning is mainly related to form (spelling, pronunciation etc), relatively unaffected by elaborateness of processing, and more subject to the effects of repetition and structured review. Associated teaching techniques address, respectively, the means of presentation of items (explicit learning), or the means of testing them (implicit processes). Research into presentation has usually focused on the relative importance of sensory (aural and/or visual) versus semantic elaboration

(eg: Brown & Perry 1991, Crow & Quigley 1985, Ott et al 1976, Atkinson & Raugh 1975). Conclusions are varied, eg: using both strategies simultaneously, enables learners of varying proficiencies to become more versatile in handling words with differing levels of concreteness (Brown & Perry); semantic processing strategies are better for advanced learners than no-strategy conditions (Crow & Quigley); visual elaboration is most effective with concrete words (Ott et al); combined aural and visual elaboration is more effective for recall than repetition and rehearsal (Atkinson & Raugh) etc. Generally there is agreement that processing of the kind which Craik & Lockart (1972) have called 'deep' because it operates at a conceptual as opposed to simply associational level, favours longer-term retention and retrieval. Research into testing has usually aimed at establishing principles for the most effective systems of re-presentation in terms of spacing and recency, eg: spaced tests produce better learning than massed tests, spacing can be more important than repetition, recall facilitates long-term memory etc. (Ellis op.cit section 8). The most significant finding of this research, for vocabulary-learning strategy, is possibly that what is called 'expanded rehearsal', ie: the spacing of re-presentations over progressively longer periods of time following successful recall. is superior to other review techniques.

There is no doubt that research in these areas has important implications for theories of vocabulary-teaching, and some of the techniques which have been advanced on the basis of its findings will be considered in 3.5 (below). It should be noted, however, as Meara points out, (1984, p.227), that as important as this research is, its pedagogical implications tend to address 'peripheral' concerns such as the method of presentation, and not the more central question of what the 'end product' of vocabulary learning ought to consist of.

3.3 Receptive and Productive Knowledge

A distinction relevant to the question of the 'end product' of vocabulary-learning, is the one commonly made between receptive and productive vocabulary knowledge (Carter & McCarthy op.cit, p.94). Palmberg (1987, 1988) gives a clear summary of

research which has addressed this issue, concluding that the end product and transitional stages of word learning can be studied both qualitatively and quantitatively by reference to a continuum which connects potential vocabulary (unfamiliar words with inferable meaning), receptive vocabulary (words recognised and understood but not used), threshold vocabulary (words which are subject to 'tip-of-the-tongue' states), and productive (fully usable) vocabulary. This model enables the specification, as a learning aim, of levels of word-knowledge appropriate to the kind of processing which the learner and the learning medium can support. Schouten-Van Parreren (1992), for example, describes complementary and differential approaches for strong and weak learners based on the distinction between receptive and productive learning objectives. Weak learners were judged to have difficulty with the inferring strategies necessary to entertain potential vocabulary items, (eg: they knew fewer L1 cognates, or they were unable to recognise relations between different forms of the same L2 word), and thus to need help with receptive vocabulary learning and reading strategies. Strong learners were found to have better powers of recognition (eg: they could often remember not only the meaning of a word but also when and where they had learned it), favouring productive practice and grammar acquisition. For a computer-based system, learning objectives at both ends of this scale are feasible, but the aim to support production of 'fully usable' vocabulary is probably unrealistic, because, as we have seen from the previous section, the scope of the knowledge that may be implied in the use of a word, and the extent to which it may be procedural or tacit knowledge, eg: in 'creative' misuses of collocation, is immense. Knowledge at the 'potential', 'receptive' and 'threshold' levels, implying the relation of word and meaning in the mental store, though not necessarily complete, or with guaranteed retrievability, is therefore a more practical objective.

3.4 Contextualisation & Decontextualisation

The issue of learning from context versus explicit memorisation has been taken up as part of the increased interest in features of the learning situation. Although research intended to establish which type of learning is 'better' has been

inconclusive (Carter & McCarthy op.cit, p.15), some related issues with important implications for teaching and learning have been raised. Hulstijn (1992), for example, reviewing the question of inferred versus given meanings, concludes that where learners focus on comprehension of a text, as opposed to learning vocabulary, they are more likely to remember unknown words whose meanings they have inferred rather than been given. This is due to the higher level of mental effort involved in inference. However, the inference procedure is more likely to result in incorrect understanding and this has to be weighed against more stable, if less effective, techniques such as glossing or the giving of synonyms or translations. Hulstijn's view is that when vocabulary learning is intentional, the effect of individual memorising and rehearsal techniques "..completely washes out the difference between the processing of given and inferred meanings.." (op.cit, p.120). This point of view can be contrasted with that of Sternberg (1987), who stresses the superiority of inferring meaning from context (for L1 learners) and advocates the training of learners to be aware of factors such as the number of occurrences of unknown words, the variability of their contexts, their importance to overall comprehension etc. The difference in opinion merely serves to emphasise the importance, for a general strategy, of supporting learners in both approaches, as was concluded in the summary to the previous section.

3.5 General Implications for Teaching and Learning

General approaches to teaching and learning, having reference to the kind of research described above, tend to stress the importance of variety and eclecticism. Oxford & Crookall (1990), for example, distinguish between decontextualising and semi-contextualising techniques based on the research into mnemonic phenomena (3.2 above), and fully-contextualising approaches based on the research into inference strategies (3.3). They conclude that techniques for getting new L2 words into memory (linking new information with existing schemata) and for recalling them when needed, are required before the words become accessible for productive use (op.cit, p.24). This supports the view that 'threshold' level knowledge is the objective most appropriate for a computer-based system (3.3). They especially

advocate word grouping and semantic mapping (based on studies such as that by Cohen & Aphek 1980, which showed that if learners use explicit techniques of mental association of word- features, both L2 and L1, it helps them to retain and retrieve the words they encounter), and structured reviewing techniques. They recommend that learners should be introduced to a range of techniques and encouraged to adopt the ones that best suit their learning style (op.cit, p.25). This last conclusion, also supported by Levine & Reves (1990), can be taken as further justification for the principle of learner autonomy which has emerged in this discussion as central to the definition of a learning strategy for vocabulary.

3.6 Section-Summary

Second language acquisition research has helped to shift the pedagogical emphasis in vocabulary-learning away from the task of organising the target lexis, and towards that of understanding the way learners process new words. Research has focused on describing mnemonic strategies, on defining receptive and productive levels of word-knowledge, and on distinguishing between knowledge which results from inferring word-meaning from context, and that which is the product of explicitly-given definitions or glosses. It has been found that mental semantic elaboration of target words positively affects retention and recall, that there is a continuum of stages, from potential to fully usable, in knowledge of a word, and that the contribution of contextual inference to meaning-retention is related to the learner's intention to either learn words or to understand a message. The main implication of all this for teaching is that the attempt to prescribe the learning content, so as to maximise the number of encounters of a given word, has to give way to an intention to promote cognitive depth in the learner's efforts to make sense of the words they do encounter. As it is clear that learners may differ widely in the way they process different types of lexical item, and consequently in their style of learning and the strategies they adopt, a general strategy for autonomous learning needs to be flexible enough to support a range of individual approaches. Providing a flexible strategy with the learner in control, means supporting processing at both the 'weak' and 'strong' ends of the scale. For lower level learners the learning

31

process is more closely identified with reading and the development of receptive skills to do with the isolation of features of meaning, by inference or by explicit research. For stronger learners, it is more to do with the elaboration of meaning and the exploration of syntactical relations. Treating a learning strategy as an explicit, conscious approach to the achievement of learning objectives enables learners to judge its effectiveness for themselves, and to modify it accordingly.

Section 4: The 'Logistics' of the Task

The need for a learner to achieve a measurable and acceptable increase in the overall size of their L2 mental lexicon means that, despite the considerations of the qualitative complexity of the lexicon discussed above, the strategic learning aims should also have a quantitative dimension. This is normally measured in terms of the number of items in the target lexicon, but the emphasis given so far to the variability of learning situations and processes suggests that the logistics of the task are likely to vary from individual to individual. It would therefore not be appropriate to suggest an absolute target for all learners. The criterion instead needs to reflect, for each learner, the rate at which the processing cycle is successfully completed, ie: the time involved in the promotion of words from 'potential' to 'threshold' levels of knowledge. 'Learning rate' thus defined can then be compared with the kind of rates of acquisition assumed for other types of successful learning, eg: L1 vocabulary-learning, or efficient L2 instructed-learning.

4.1 The Size of the Lexicon

Whilst it is not uncommon for studies based on the analysis of dictionaries to estimate the size of the (English) lexicon as anything between ½ and 1 million words (Zechmeister et al 1993), in practice the learning task is considerably less. Zechmeister (p.203) introduces the concept of 'base words' (word families excluding homographs) and quotes Goulden at al (1990) "..the average educated native speaker has a vocabulary of around 17,000 base words..". This estimate, as a minimum, corresponds with that of several other sources (Lewis op.cit, p.122, Nagy & Herman op.cit, p.21, Miller 1986, p.174). On the question of learning rates,

however, there is less agreement. Miller (1986 op.cit, p.174) suggests that L1 learners between 6 and 8 years old acquire 21 words a day. Nagy & Herman say children learn their native language at 3000 words per year (half as fast). Goulden et al (op.cit) propose an average rate into adulthood of about 2 to 3 words a day. The problem is, of course, that it is very difficult to know what stages L1 word learning goes through, because of the holistic and largely unconscious way that the native language is acquired. Zechmeister's study, however, also reports that most native speakers believe that it is possible to increase vocabulary size by formal study, by as many as 1000 words in 3 months, which suggests that people have an intuitive notion of the L1 mental lexicon as something that can be fairly quickly added to.

4.2 L2 Word-Learning Rates

The same confidence is not manifested by those concerned with the L2 mental lexicon. Meara (1982, p. 224), for example, discusses 6 words a day on the basis of "..the capacity of the brain to acquire new information.." but refers to others' opinions that this would beyond the capacity of many learners to sustain. Gairns & Redman (1986) claim that 1000 productive items in 125 hours (1 hour per day) of study is probably beyond most L2 learners. However, teaching and learning strategies are obviously significant for both acquisition and assessment, and if the learner is required to deal with new vocabulary items in context, the rate at which their reading comprehension improves is relevant to how fast they will be able to learn words. Similarly, if they are required to recall new items from memory, their access to clues and mnemonic aids will also have an effect on how fast they learn. The optimal strategy proposed here, emphasising both contextual and conceptual processing of words, might be expected to achieve an optimum rate. In the interest of providing some kind of yardstick, albeit somewhat arbitrary, Gairns & Redman's 1000 words in 125 hours, or 8 words an hour, is proposed as a target processing rate for this system, in which each word should be taken from potential to threshold level.

Section Five: Chapter Summary and Parameters for a Learning Strategy

The purpose of this chapter has been to determine the requirements for a strategy for autonomous vocabulary learning. The traditional tendency of descriptive linguistics to view lexis as having a secondary role in the construction of text, has mitigated against a principled description of vocabulary in terms of learning content, but more recent theoretical developments, showing that the lexicon is structured at both paradigmatic and syntagmatic levels, and that it plays a major organising role within the language system, have opened the way for pedagogical description and for the application of strategic teaching and learning approaches. In addition, the role the individual learner plays both consciously and unconsciously, in determining the processes and outcome of learning in terms of a range of receptive and productive skills, is now recognised as a critical factor, as is the overall size of the learning task in terms of the number of items to be acquired, and the rate at which this can be done. The general requirements for a strategy for autonomous vocabulary-learning can therefore be summarised as:

a) Centrality of the learner's own definition of the required learning content.

b) Richness and scope of the learner's conscious semantically-related processing.

c) Psychological appropriateness of the means of testing and rehearsal.

d) Optimal rate of promotion of new items from the 'potential' to the 'threshold' level of productive knowledge.

The overall aim of a strategy should be to support the learner in creating a sizeable L2 mental lexicon at threshold level, as a basis for subsequent productive communicative practice. This can be expressed quantitatively, in terms of a learning rate commensurate with the total number of vocabulary items required for an improvement in reading competence, and the amount of time available in which to do this. Figures are relative, but general targets of 1000 items per 125 hours have

been adopted from the literature, giving a target learning-rate of approximately 8 words per hour. Learning content should consist of: firstly, individual learner-defined vocabulary items - single words, multi-word items, lexical phrases etc. - which the learner identifies as being of interest to them, either because they are unfamiliar, or are being used in an unexpected way, or because they contain some feature of meaning or form which merits further exploration. The source of these items would be textual discourse which is assumed to be representative of 'normal' frequency and distribution of lexis in the target language. Secondly, contextual features of the texts, including the way individual items are combined at sentence and paragraph level, the sets of words formed by topic relations, and the contribution of features of register and style. Thirdly, the lexical structure implicit in the comparison of individual items, ie: the features of meaning (sense, semantic primitives etc.) or of form (grammatical or orthographic) which are held in common or in opposition by two or more items. The conscious skills and unconscious processes which the strategy would aim to promote in order to process this content, are those concerned with comprehending text, selecting, analysing and elaborating target vocabulary-items, and using clues to retrieve meanings and word-forms from memory. The general approach to processing would thus involve the selection, elaboration, and retrieval of items in a sequence which is varied at the learner's discretion, allowing for the exploitation of efficiency in the implicit processes where possible (eg: acquiring an item directly from the text and retrieving it without going through the elaboration stage). Over time, and through the iteration of these stages, there would be a progressive build-up of the target lexicon and of the learner's awareness of potential associations and of the lexical-structural basis of the target language.

In Chapter Two I will examine how far CALL is able to mediate the kinds of learning processes outlined here, and try to establish a framework for the general design and functionality of a system which can be implemented within the constraints of current technology.

CHAPTER TWO: PRINCIPLES AND SPECIFICATIONS FOR A DESIGN - THE STATE OF THE ART IN COMPUTER-BASED VOCABULARY LEARNING

Introduction:

The vocabulary learning strategy described in the previous chapter is founded on the idea that the learner's own strategic decision-making, about what and how to learn, is central to both the quality and the outcome of the learning process. The support which a computer-based system could provide for this decision-making is, theoretically, considerable, especially when we take into account the capabilities of current technology in the areas of information storage and access, audio and video, graphic and textual presentation etc. However, there are practical constraints arising from the realities of autonomous-learning environments for many language-learners, and there is also the need for the theoretical basis of any design to be as clear as possible. These two considerations taken together suggest that the technological platform for the system under discussion should be kept at the lowest level which can safely be assumed for most contemporary computer-assisted learning (eg: see Fox et al 1990, p.12 for what is considered 'widespread'). That is to say, we will not be considering the possibilities of multi-media or telecommunications (as interesting and feasible as these are). The design proposal which will be advanced here assumes a human-computer interface such as is standard with the majority of personal computers at the present time (ie: a 'WIMP'[3] interface like PC Windows or Macintosh). The linguistic environment is assumed to be in the form of written text - though this may be graphically manipulated - and the interaction is assumed to be one-to-one between system and learner.

The purpose of this chapter is to establish the basis for a computer-based system, thus constrained, which is capable of supporting a learner's strategic decision-making in vocabulary-learning, as defined in Chapter One, ie: capable of meeting the following requirements:

i) Learners select their own vocabulary-learning content..

---

[3] Windows, Icons, Mouse, Pull-down menus

ii) They are supported in the elaboration of the items they select - in terms of paradigmatic and syntagmatic relations to other vocabulary in the source contexts and in general.

iii) They are supported in productive practice of the items they select.

iv) These processes are integrated and flexible, amenable to individual learning styles, and supporting a rich learning process.

v) The overall interaction is progressive - extending over multiple sessions, and supports vocabulary-learning at a rate which is consistent with the logistics of the vocabulary-learning task as described in Chapter One.

Existing practice in CALL for vocabulary is the starting point, and Section 1 will outline some past and current approaches to the issue. Design principles relating to testing, selection of content, and the use of on-line tools are discussed. Section 2 is devoted to a statement of functional specifications, based on general principles of computer-aided learning design, for a system intended both to support the learning strategy and to collect data on the learning processes.

Section 1: Computer-Aided Vocabulary Learning, Past and Present

1.1 Testing vs Teaching

CALL has always claimed to specialise in vocabulary, perhaps because of the tendency in language-teaching, in the past, to let vocabulary look after itself, thus confining systematic study of it to the learner's spare time. This is a learning situation in which CALL has thrived - in Jung's (Jung 1988, p.4) survey of descriptor frequency in the international bibliography of CALL, vocabulary-learning comes 5th out of 20 for number of occurrences). However, we may want to be a little sceptical about how far principles of design of such programs are specifically oriented to vocabulary-learning as opposed to mere memory-testing. Computerised versions of word games such as anagrams, crosswords, word ladders etc., have been

used by learners for many years as incidental ways of building up vocabulary (see Kenning 1990 for a review), but as Clarke points out (Clarke 1993) they are essentially concerned with testing what has been previously learned, and not with the learning processes themselves. In the most commonly-used testing programs, learners are required to type a target word in, or to select an answer from a range of options, or to create matches between two lists of elements. The usual format for testing vocabulary is a sentence with a gap to be filled. This technique, often referred to as the 'cloze' test, has a long history in CALL in particular and language teaching in general (eg: Chapelle 1993, Laufer & Osimo 1991, Carter 1988 op.cit), so it is worth considering how far it can be said to support vocabulary-learning processes other than memorisation.

1.2 Total Cloze

Perhaps the best-known and most developed of these testing programs, an extension of the basic cloze technique, is the text reconstruction program known as Storyboard (Higgins & Johns 1984). This program, described as an "..acceptable paradigm for good practice in first generation CALL.." (Last 1989 p.47) remains one of the most widely used and is a good illustration of how established paper-based language testing techniques can be given a uniquely computational character. Storyboard has been called a 'total cloze' activity because it takes the cloze principle and extends it to cover all the words in a passage. This works on a computer because the machine is able to give instant feedback on the learner's guesses and thus convert a potentially impossible task (reconstruct an unseen text) into one where deduction and trial-and-error render a solution progressively more achievable. It thus engages the learner in processing of context as well as words. The object of the program is for the learner to reconstruct a text from scratch, by entering individual words. The words the learner inputs are matched against the ones in the text and if correct are revealed at every place that they occur (see Fig 2.1). Word-content information comes initially from the broad semantics of the topic (which words are likely to appear in this sort of text under this title), and then, progressively as the gaps get

filled in, from syntactic and collocational features as well (ie: which blanks are likely to be nouns or verbs, which words are likely to follow or precede others).

Fig 2.1: Example of Total Cloze

```
Four-stroke cycle              STORYBOARD            Words found :  11 / 46
                          1. ████████ stroke:
   ████ the ████ valve open ███ the ███████ valve █████, the
   piston █████ down, ███████ █ ███████ ██ ████ ███ ███ ████ the
   ██████████ ████ the ████████.

                          2. ██████████ stroke:
   ████ ████ the █████ ███ ██████ █████ █████, the piston
   █████ up, ██████████ the ███████, ████ █████ the piston
   ███████ the top ██ ███ stroke, ██ ████████ spark ████ the
   ████████ plug ██████ the ███████.

                          3. █████ stroke:
   The ████████ ████████ drives the piston down the ████████.
   ████ ██████ ██████ ██████.

                          4. ███████ stroke:
   ████ the █████ valve ██████ ███ the ███████ valve open, the
   piston █████ up the ████████, ████████ the ███████ ████.

 f1 : letter   f2 : word  f3 : text  f4 : score   f5 : intro    Escape : end
        Guess a word : _
```
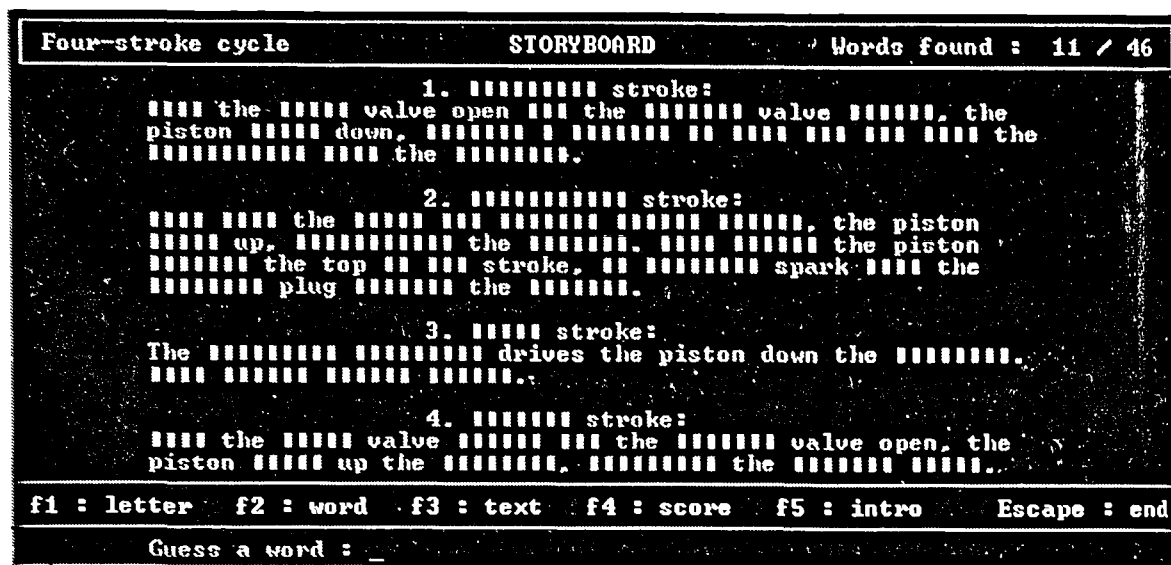
The program has options to 'peek' at the complete text before and during the exercise, and to get a 'free' word (usually the first unrevealed one). There is also information at the word level in the blanks representing the number of letters a word has (and some versions have an option to enter common affixes). The task from the learners' point of view may begin as an open one (try any word) and then gradually become more constrained (try words appropriate to particular gaps), or it may be approached in a linear fashion, using the peek and free-word options to help recreate the text from first word to last.

Tim Johns, (the self-confessed "grandfather" of total cloze), has doubts about the appropriateness of the technique for vocabulary learning, "..for the simple reason that if the student does not know or only half-knows a word, it is highly unlikely that s(he) is going to find that word for one of the gaps.." (Johns 1992). This is true for all cloze-based exercises, but given the range of contextual clues and types of help provided in Storyboard (single letters, affixes, suffixes, whole words, whole text etc.), it could be said to support a certain amount of lexical learning by inference, with the learner progressively more able to guess the meanings represented in unfilled gaps as the overall context unfolds. Single words,

collocations, discourse markers and even grammatical structures are all theoretically learnable in this way. There are also reasonable grounds to believe that a complex range of strategies may underlie the learner's performance (eg: see Goodfellow 1993 op.cit). For this reason it has been claimed that this program is compatible with a view of language learning as information processing (Phillips 1985, p.41). Storyboard, considered from the point of view of vocabulary-learning, illustrates well the complex inter-relation between vocabulary knowledge, reading ability, inference and production skills. Although neither it, nor the various developments from it (eg: Camsoft's Fun With Texts), stand alone as solutions to the design problems posed by the requirements of the autonomous learning strategy under consideration here, because of the lack of decontextualised information, and the relative poverty of the performance elicited from the learner, it may nevertheless serve as a model of a relatively rich learning interaction, in contrast to the memory-testing approach of much conventional CALL-for-vocabulary.

1.3 Semantic Association

Testing techniques tend to be context-free and relate to form (spelling) rather than meaning (Clarke op.cit, p.140). As such, they are probably better described as dealing with word-memorising rather than vocabulary-learning in any of the more complex senses described in the last chapter. Some developers have consequently incorporated psychological principles regarding the most effective uses of repetition and review. Ellis (forthcoming) gives the example of the strategy of 'expanded rehearsal' in which the number of items between successive test presentations is progressively increased, and notes that Siegel & Misselt (1984) incorporated it into a Japanese vocabulary-learning program, with results superior to those of more traditional review techniques. Other psychological insights imported into CALL-for-vocabulary have focused more on the way meaning is associated with word-form. The Linkword series, for example, promotes memorisation of target words via association with visual images, and others have exploited the kind of principles of lexical structure that were discussed in the last section. Meara (1985) describes a program which uses the traditional 'wordsquare' puzzle format to present learners

with blocks of letters concealing German words which they have to identify by typing in. The target words in the square are semantically related and clues are available to this effect. In a further development (Links - see Kenning & Kenning 1990 for a description), words are utilised in an explicit association-forming exercise which requires learners to link words in a chain by selecting appropriate intermediate words which make the association logical. In another semantically-based approach, Catt (1992) describes the use of simple matching and gap-filling routines for exercises in which the learner is required to associate vocabulary items with short contexts exemplifying synonymy and hyponymy relations, eg: (op.cit, pp.134-136)

```
You look tired.................................................I'm absolutely exhausted (synonym)
Have you got a hammer?...........................No I haven't got any tools (hyponym).
```

The strengths of these more lexically-principled approaches are in the emphasis they put on the integration of new items with existing semantic knowledge. Their limitations are in their rigid specification of content, requiring specialised preparation in terms of problem/solution or problem/clue frames, and a consequent lack of individualisation. The performance elicited from the learner is also relatively impoverished, and may not require them to actually produce anything. Additionally, although both the latter approaches involve the learner in using a dictionary (in the case of the wordsquare a small 200-word on-line dictionary was incorporated into the program), in neither case can the information thus acquired be processed (other than internally) outside of the input format required by the test. The same limited performance restricts the operation of feedback to a simple match with the 'correct' solution. Harland (1990, p.92) describes the simple matching of input and target as "..the deadly 'wrong, try again!' variety of system response..", deadly because such feedback does nothing to rectify poor learning approaches. There is a need instead to provide feedback which enhances the learner's subsequent input attempts. A simple example of this is described by Higgins & Johns (1984 op.cit, pp.174-175) who devised a semi-intelligent routine to check student input for reversals, insertions and omissions of letters. The rationale in their approach is that it is unmotivating for a learner to have small errors punished, but

there is potentially a more significant role for such analysis in the identification of elements of a partially correct response which can be used to focus the learner on possible remediation strategies. Harland, for example, proposes the use of a 'simplified lexicon' specific to a given exercise, from which clues to the correct answer might be given (op.cit, p.92).

Some more recent approaches utilising semantic relations have given the learner more to do. A system for teaching German vocabulary to French speakers, developed at the University of Cologne (Widdig & Esser 1988), presents its target words in context (eg: in a business letter), and allows the learner to investigate them for lexical properties such as synonyms, antonyms, transitivity etc. The system even does some error checking on learner input, ie: for spelling mistakes, incorrect endings, confusion between words and their antonyms etc. Another system, for teaching French vocabulary to English speakers, from US Army research in Alexandria USA, (Lexnet-Insitu - described in Swartz et al 1990), attempts to model some of the cognitive processes involved in word-learning, and maintains an explicit network of semantic links between target items (incorporating synonyms and hyponyms). Learners have to reconstruct parts of this network to demonstrate that the relations between the words have been understood. These approaches have much to commend them, but again they necessitate a considerable materials-preparation effort, which means that they cannot, practically, support the learner in specifying their own learning content. In the continuing absence of a natural-language-processing technology which is able to automate the use of semantic cross-reference information within the lexicon (Boguraev & Briscoe 1989), the effort required to give such support to even a limited amount of learner choice in the selection of target items, is considerable. Nevertheless, an important design principle remains, which is that the kind of learner processing represented in these programs is cognitively 'central' to vocabulary-learning, and some kind of representation of semantic association will have to form part of the design in order to exploit this.

1.4 Target Words

Selection of target-words has usually been confronted with either of two broad approaches made possible by recent developments in computer technology. One is to utilise the power of database storage and retrieval facilities to specify a vast amount of decontextualised material from which the learner can select, the other is to use screen text-manipulation functions to allow the learner to customise their learning material directly. A number of systems have tackled the problem of individualising pre-specified lexical content by building large databases from which learners (or teachers) can select blocks determined by frequency or topic. Such a system is Didascalia's Wordchip, referred to in the previous chapter (Decoo 1993a op.cit, 1993b). This is the only system reviewed here which is explicit about the learning rate it is intended to support, approximately 1000 'lexical didactemes' per year learned in 'a few class hours a week' (Decoo 1993a, op.cit). This rate does not have any explicit implications for the learning processes involved, but only for the content, ie: vocabulary in 'frequency blocks' corresponding to the developing communicative language needs of 10-17 year-old school children. It therefore has nothing to say about the kinds of learning strategies which are necessary to achieve the rate. Another system which takes the quantitative road is Leuwen University's Adam & Eve which is able to generate cloze-type exercises from a database of text selected by learner and/or teacher. Gains in flexibility of content and scope of the learning goals, however, are not matched by equivalent gains at the local level or by enhanced integration of knowledge or learner performance. In the absence of any other processing stage the fixed conventional exercise types (cloze tests, multiple choice, matching) confine the representation (and potential analysis) of the learner's performance, to the testing stage alone. These programs are perhaps better viewed as teaching aids, where the teacher can provide extra scope for performance and feedback, than as autonomous-learning systems. Decoo, in fact, has observed that teachers and/or school syllabi are often influential in the decision to buy language learning software for self-access use (Decoo 1993b), which suggests that there is often a close link between the modes of use.

1.5 Learner-Customised Learning Content

The approach in which learners generate their own content from the linguistic world at large was originally exemplified by the Wordstore program (see the introduction), in which the focus is on supporting the learner in compiling their own 'dictionary' of target words. Technical developments in applications software such as word-processors, graphic user interfaces (GUIs), hypertext etc. have made the approach more viable, because of the enhanced means of manipulation of text made available to the user, and the greater scope for the program to interpret the user's intentions. Gillespie and Gray (1992), for example, developed a hypertext system to enable users to select words, expressions etc. directly from text and store them in a network of 'datacards' constructed to reflect the substance and distribution of topics etc. in the text. They consider that using it challenges the learner to deal creatively with new words, phrases and grammatical principles which they have discovered, and to apply them again in different contexts. By doing so it helps them develop the intellectual skills of information gathering and the classification of language (op.cit p.10).

The logical extension of the facility to have the learner select their target vocabulary directly from text, is the provision of text commentaries and glosses (which today can be augmented by pictures and sound) to help them relate selected items to their overall understanding. Several systems have been developed around this principle (eg: Harland op.cit, Lyman-Hager et al 1993, Tiebault 1993), the basic operation of which involves the learner in marking areas of text on the screen, which activates links to various kinds of informational content. The strength of such an approach is in the multiple levels of integration of new knowledge that is thus made accessible to the learner, together with the increased learner autonomy which results from embedding the content in a higher-level reading task. However, the problems associated with materials preparation time remain: Lyman-Hager (op.cit), for example, describes the effort involved as so labour-intensive that it has motivated a number of tests designed to see whether it is worth it in terms of vocabulary acquisition. (This research will be referred to again in the section on research issues

below). However, further developments in the direction of computerisation of dictionaries and other lexical information resources have begun to provide ready-made learning content which can be linked 'on-the-fly' (Harland op.cit, p.92) to the learners' selections. A model of a learning process which might be supported in this way has been suggested by Kukulska-Hulme (op.cit), who proposes that the two most important stages from the point of view of computer support, are finding meaning from reference works, and retrieving a word from the written record for productive use (see Fig 2.2)

Fig 2.2: Kukulska-Hulme's 'Journey of a Vocabulary Item'
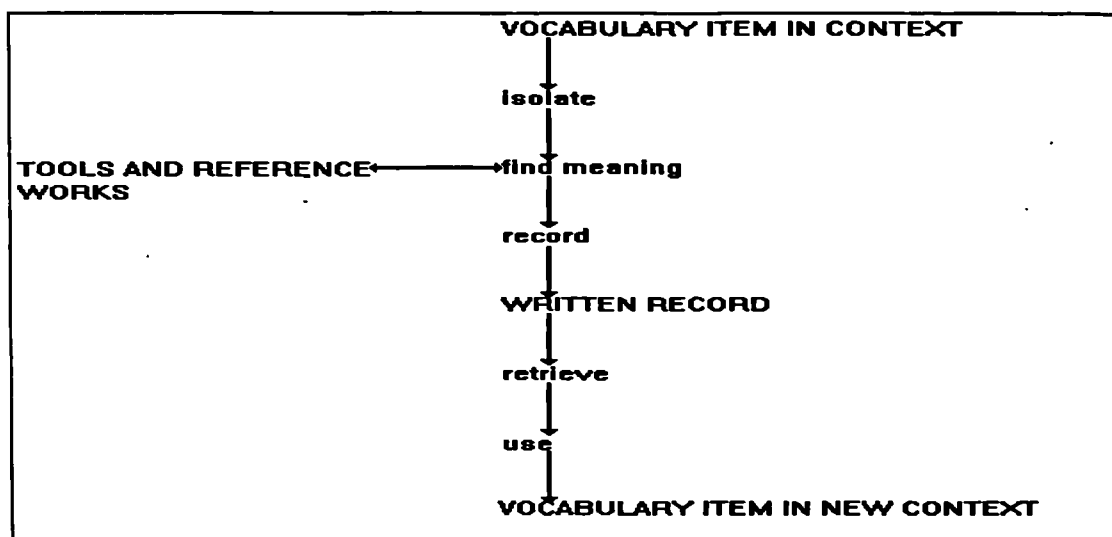
Her model was intended to apply to the sharing of lexical information for collaborative work (see Introduction), which is why she stresses the written record, but in a separate point she acknowledges that there is a parallel process of memorisation in which the written record becomes a record in memory. This process is not necessarily conscious (the point made by Ellis referred to in the previous chapter), but where conscious effort is made it can be enhanced by the application of semantic and orthographic clustering criteria in the mental storage of items. The tools and reference works are therefore required to support search on the basis of these criteria as well as simple definitions or translations.

1.6 Tool-based Approaches - Dictionaries

There are many different kinds of computer-based information resources relevant to the processing of language for one purpose or another - Knowles (1990) lists 40 different types of tool that are theoretically available, including dictionaries, thesauri, lexical databases, concordancers, parsers, indexers, grammar-checkers etc. The most common type of lexical resource used in CALL, are on-line dictionaries, into which there is research going on in a number of areas (eg: Harland op.cit, Guillot & Kenning 1994, McElligot & O'Neill 1993, Zahner et al. 1994). Although the first of the widely-available on-line bilingual dictionaries (eg: Collins On-Line, Harraps) are simply computerised versions of book dictionaries, developments in user interfaces and the design of lexicological tools have created the possibility for dictionaries to free themselves from the conventional types of look-up and to support autonomous learner investigation of a much more complex kind. There are several systems currently under development which aim to harness one or more of this type of tool to the development of language skills (eg: Horton et al 1990, Cumming et al 1993, Hayet 1994, Fox 1993).

One project which is specifically addressing issues of learner use of lexical tools is Queensland University's and La Trobe University's Mayday project (Sussex et al 1994, Cumming et al 1994, Cumming et al 1993). This project, which has produced a system for learning aspects of word composition in English (negative prefixes, verb affixes, noun suffixes), has incorporated an on-line learner's dictionary, a synonym dictionary, and a dictionary of derived words, into an exploratory environment for lexical learning, and is studying the advice that teachers give to learners who are using these tools to help them with various lexically-based activities and games. The Mayday researchers have customised their tools in order to support particular learning processes. For example, the dictionary has been organised so that compound words are not used as headwords but have to be looked up in a separate 'word building' resource (Cumming et al 1994, p.119). This reflects a view that word composition is a distinct process from word comprehension, and also the wish to keep the processes separate so that they can be studied

independently. Most of the reported work to date has concerned students' strategies when words are not found in the main dictionary (because they are compounds), although conclusions are geared to the general hypothesis that there are distinct 'discussion' and 'task' levels of teacher-student interaction, rather than to descriptions of the learners' autonomous use of the tools. The synonym dictionary is organised by word senses, which are related to senses in the main dictionary by numbers, in contrast to the style of standard thesuari, so that learners can "..follow a particular sense...and build up a deeper, more informed understanding of it.." (Sussex at al op.cit, p.142). There are plans to incorporate Miller et al's Wordnet reference system (Miller 1990), which is a lexical database based on a semantic association network. The work in this area is still in its infancy, so it is not yet possible to draw any lessons about the kinds of semantic-clustering processes which these tools support, or about the effect of their use on the learners' lexical knowledge.

Hayet (1994) makes a distinction between lexical classification of the type found in conventional dictionaries, and the semantic hierarchies and thesaurus-type combinations which she utilises in Dicologique, a lexical resource which is integrated into a software platform under development at UMIST[4] intended to support vocabulary learning, translation etc. in technical subjects. This system associates lexical items and their underlying 'notions' (which may be descriptive terms such as 'green' or 'metal', or classifying terms such as 'Christian symbol' or any of several other qualifying categories) in a polyhierarchical structure which can be searched for groups of lexical relatives such as polysemes, homographs, and synonyms. Of particular interest are what she describes as 'usage' or 'documentary' synonyms. These are associations of the kind "..term A can be used for concept B in a particular environment.." where there is no direct synonymy. Hayet claims this kind of semantic relation is useful in CALL to support the study of lexical semantics, "..with the description of words in terms of their relations with other words within a subject field, while developing the awareness of the groupings dictated by usage in a foreign language.." (op.cit. p.100). Learner investigation

---

[4]University of Manchester, Institute of Science and Technology

using Dicologique therefore engages them in processing semantic information which reveals directly the structure underlying the target language lexicon. Current research with the courseware into which this system is integrated includes looking at the ways in which optimal use of the integrated tools constitutes a methodology for learning which can itself become a teaching aim. As with the Mayday project, however, this research is not sufficiently advanced to have any immediate pedagogical implications.

The support that these innovative tools can provide for autonomous vocabulary learning is clearly dependent on the learners' ability to use them. As Sussex et al (op.cit, p.142) point out, there is always the possibility of cognitive overload, especially for lower-level learners, resulting from unfamiliar analytical tools and methods. Given the lack of research, to date, into how learners interact with such tools, and taking into account, also, the practical difficulties of integrating innovative software into a standard development environment, it is probably not feasible to propose the incorporation of anything more complex than existing on-line dictionaries (bi-lingual and/or mono-lingual) and thesauri where available, into the system under consideration. If the learners' attention were adequately focused on the idea of semantic clustering, it would anyway be possible for them to use a conventional bilingual dictionary to find synonyms, antonyms etc. by utilising their knowledge of L1.

1.7 Tool-based Approaches - Concordancers

Learner investigation, using reference sources, can be motivated by a 'data-driven' approach. Johns (1986,1988,1991a & b) has demonstrated a computer-based methodology for language learning which puts the learner into the role of "..linguistic researcher.." (1991a op.cit, p.2), identifying, classifying and generalising facts about the target language from the evidence provided by large corpora of text. The instrument which makes this possible is Microconcord,[5] a KWIC (keyword-in-context) concordancer which searches text files very rapidly

---

[5] By Tim Johns, Published by Longman

and locates and displays every occurrence of a specified word together with its immediate co-text (see Fig 1.1 in Chapter One). The information a learner can get about a word, from this data, includes its different meanings, collocational features, main forms, the kinds of syntactic roles it can play, its register and discourse functions etc. Because the information is implicit in the data and has to be extracted by the learner via processes of observation and inductive reasoning, the approach is called "..data-driven..". One example of the kind of linguistic generalisation about English that has been arrived at by learners using this method, concerns the difference between the words 'convince' and 'persuade' - the former generally being used in relation to truths whilst the latter tends to refer to actions; John's learners (1991a op.cit, p.5) concluded that these two words exemplify the main difference in use between the 'that-clause' (eg: 'to <u>convince</u> someone <u>that something is true</u>') and the 'to+infinitive' construction (eg: 'to <u>persuade</u> someone <u>to do something</u>'). This particular example has a lexico-grammatical character which is highly relevant to current ideas about the lexicon in the field of descriptive linguistics (see Chapter One). Other examples focus on lexical meaning, eg: hypernyms revealed by the use of 'such as' ('..<u>industries</u> such as steelmaking..', '..<u>crops</u> such as coffee..'), or discourse function, eg: the difference in use between 'however' and 'nevertheless', or grammatical forms, eg: the varieties of 'have to' ('..the institute <u>has had to</u> rely on..', '..councils <u>are having to</u> give..'). A further development of the approach suggests that the fragments of discourse typically generated by the concordancer, if derived from a text database which is familiar to learners, may be more useful in reinforcing vocabulary than the more commonly used cloze tests (Stevens 1991).

Johns has shown that learners do acquire linguistic knowledge from working with concordance data, and has gone on to develop a teaching package (Contexts) which puts his approach into a more rigid pedagogical framework (fixed numbers of pre-selected citations, demonstration procedures, simple research and practice tasks). Like the lexical-tools researchers mentioned in the previous section, he is interested in the way learners use the procedures, and the program logs student sessions as 'histories' which can be later analysed for patterns of learner behaviour. Like the

Mayday and Dicologique research, also, it is too soon for any more than very general conclusions (Johns 1994), but he does claim to have identified certain tendencies including one in which learners avoid the program's inspection-of-data functions (the 'Show' option) in favour of its testing functions ('Quiz' option). Johns puts this down to what he calls the "..beat the machine effect..", whereby "..we are blowed if we are going to be bested by a mere machine, and to show who is the master we make things as difficult for ourselves as possible..". In considering possible remedies for this, he opts for learner training rather than giving the program more control over the procedures. This is clearly an important consideration with learning processes, such as doing KWIC searches, or using a lexical database, which users may not be familiar with, and which may not, at first, be particularly intuitive. It raises the issue of explicit tutorial support in the development of procedural linguistic skills - an area which has been neglected in CALL, outside of 'process-prompting' programs for written composition (eg: see Phinney 1989). In the absence of clear principles for providing this support for the vocabulary-learner, it would seem that the onus falls on the design of a computer-based autonomous-learning strategy to ensure that the kind of knowledge represented in the output of a particular tool or resource (such as a concordancer) is as far as possible integrated with the more conventional procedures (such as reading a text or doing a cloze test) in an apparent a way as possible. This can be done by: a) giving the learner an immediate and obvious means of relating a KWIC citation 'back' to its originating text, so that any comprehension problems arising from its truncated form can be resolved, and b) using KWIC citations as part of the testing, as well as the elaboration, stage of the interaction (eg: in the manner suggested by Stevens, above), so that they appear as part of a familiar cloze-test etc.

1.8 Section Summary

The development of CALL-for-vocabulary roughly corresponds to a progression from simple machine-controlled word-testing, to complex learner-controlled lexical information processing. Some conventional systems appear to support relatively rich, semantically-based, learning processes and meet some of the requirements of

the strategy for productive practice of target items. They fall down, however, in the rigid pre-selection of content, restricted scope of learner performance, and general lack of system adaptability. GUI-based developments have made possible an approach in which learners manipulate text on the screen rather than select items from decontextualised lists. This approach has been claimed to support a more autonomous and richer level of learner processing, by creating a detailed linguistic environment which is able to support remediation of errors. The extensive glossing of text, however, which is needed in order to make all potential learning content explicit, requires a level of preparation of material which is incompatible with maximising the numbers of items which can be processed. Some emerging systems have demonstrated the feasibility of incorporating general language-information resources such as on-line dictionaries, thesauri and concordancers. This approach meets the requirements for elaboration of learning content and the self-regulation of processing, and is compatible with a large set of potential target items. It provides for an extensive base of potential learning content at two levels - meanings and behaviour of particular words, and types of generalisations that can be made about classes of words. It aims to support the learner's development of information processing procedures which will allow the integration of new with existing knowledge. It presents the learner with linguistic information 'en masse', and expects them to apply their own strategies of selection and analysis. Implicit assumptions about the learner are that they are well-motivated, sophisticated (even with experience in research methods in their own subject areas), have particular needs specifiable in terms of target texts, and are in a learning situation in which a great deal of emphasis is placed on the development of their own learning strategies and on their responsibility for their own learning (Johns 1986 p.161).

The general shape of a design that meets the requirements outlined at the beginning of this chapter, can thus be envisaged in terms of the system features discussed in this section. It would be based on a GUI interface, giving the learner self-regulated access to: a corpus of text, a bi-lingual dictionary and keyword-in-context concordancer, and procedures for context-based productive testing of target items, organised according to principles of structured rehearsal and review (see Chapter

One). Emphasising the learner-controlled aspects of the interaction in this way, involves a trade-off against opportunities for structuring the system for teaching purposes, but this is the consequence of prioritising individualisation and scope in the selection of learning content. This would be, therefore, a system which does not set out to do any explicit teaching, but intends, instead, to support the learner in the development of strategic awareness and the ability to profit from the interaction. As such it should be based on an understanding of how such a development can come about. We do not yet, however, have such an understanding, because we are not yet in a position to describe clearly the processes involved. Contributing towards such a description is one of the aims of this design, in common with much contemporary CALL (eg: the Mayday project and Marie Hayet's and Tim Johns' work referred to above). A brief review of some of the issues involved will be given in the next chapter. Meanwhile, in the, hopefully not prolonged, absence of such an understanding of learner-related aspects of the interaction, it is important to have a firm basis for the system design in principle, so that its potential performance can be optimised in key areas such as learning goals and feedback. General principles for computer-based learning systems are therefore adduced, in the following section, as the basis for a specification of system functionality, covering the requirements which have thus been identified.

Section 2: Requirements for a Design

In this section I will specify the requirements for a system which is intended to support autonomous vocabulary-learning of the kind described in the previous chapter, following the model provided by systems which focus on supporting the organisation of learner processing rather than the provision of precise information regarding the knowledge which is to be acquired. It is an exploratory system, for maximising the rate of vocabulary learning over a period of time, and the pedagogical features it incorporates are to support the development of effective vocabulary-learning strategies. User interaction with this system is intended to provide information about how learners go about the vocabulary-learning task, and the stages they go through in the development of strategies for optimising the

learning outcome. Where the purpose of a CAL program is to support a self-contained learning process, adaptable to the needs of individual learners, there are certain general principles which apply. These relate to: the information contained in the program, the learning goals, integration of new content with existing knowledge, learner performance, and program feedback (Laurillard 1990, 1991). The following requirements will therefore be classified under these headings.

2.1 Information

The principle is that a system should provide the learner with access to the 'new' information they need in order to achieve the learning objectives. For a vocabulary program, new information should be contained in suitable text and in the lexical tools. Vocabulary items for learning should be selected by the learner and should be locatable in their context, as part of the process of reading a text. The items may be single words, multi-word items, and phrases. The requirement for a large number of potential items overall, and the necessity to avoid an overload of unfamiliar words in any single text, suggests that a number of texts, a corpus, is needed. Selection of the texts themselves is therefore also an important feature of the design. Access to information about unknown words, eg: definitions, translations, synonyms and collocations could be available via on-line tools.

2.2 Learning Goals

Learning goals are what define the aims of learner performance, and what motivate reflection on that performance. They can be global in their scope, relating to the system's overall aims, or local, relating to the learner's current objectives. The global goals of this design are: to acquire new vocabulary at threshold level at a rate consistent with the logistics of the task (see Chapter One), to develop an understanding of the structural relations which underlie the lexical system in the target language, and to develop strategic skills such as inferring meaning from context, using lexical reference tools and resources, generalising about lexical similarity, and constructing and using clues to aid productive retrieval. These goals can be made explicit by dividing the interaction into 3 separate stages, selection,

elaboration and testing, with appropriate functions available in each. The local goals thus begin with the selection and general comprehension of a text, and the identification of target items (the new information), and go on to the elaboration of the items and productive testing of them. The learner works with clues to individual items, with the goal of reproducing each item from memory. Success in retrieving items is an thus explicit measure of success at achieving the learning goals.

2.3 Integration

The principle of integration refers to the relation of new information to what the learner already knows. The design should support, or allow for, integration in a number of ways. Firstly there are the processes of inferring meaning from context, which go on at the selection stage. The learner should be supported in their goal of comprehending the source text, by the ease of identification and selection of target items, which can be done without interrupting the flow of meaning. This would allow the learner to utilise existing vocabulary knowledge and reading ability in the guessing of meaning from context. Secondly there is the selection and recording of information from the lexical tools. Here again ease of reference is an important factor for integration, as is the design of the tools themselves. The latter consideration is beyond the scope of this design, which is committed to using existing programs as tools, but the incorporation of a standard dictionary at least allows learners a basis of familiarity with the format of the information provided. Integration could also then be supported by the facility for free-text entry of the learner's own description of the item. Finally, the design supports the learner in utilising, and therefore reinforcing, associations they have created between L2 items and their meanings, related items, definitions, L1 translations etc. as clues to assist in the productive retrieval of these items from memory.

2.4 Performance

The main principle of performance is that the learner should be engaged in activities which afford learning. For this strategy, these have been determined as: selection of vocabulary-items from text, recording of information on items researched via the

use of lexical tools, grouping of items according to structural similarity, and cued productive retrieval of items. Performance should also demonstrate that learning has taken, or is taking, place. Learner performance can be assessed qualitatively, at the level of the development of effective learning strategies, eg: the global goals referred to in 3.2 above, or it can be assessed quantitatively, at the level of the learner's speed and effectiveness in achieving local goals. The ideal roles for both kinds of assessment would be in the provision of feedback for the learner, and the automatic adaptation of the interaction to their individual requirements, but in the current state of our understanding and representation of qualitative factors (see the next chapter for a discussion), this role is only feasible at the quantitative level. Qualitative aspects of performance supported by the design can, however, be used in the building of a description of learner processing at the theoretical level, for example: the notes and item-groupings produced by the learner at the elaboration stage may be evaluated for their contribution to success in the testing stage. In the design, the clues used for testing retrieval draw on the processing which has preceded this stage. If a qualitative judgement is made about the processing involved in each of these methods of stimulating retrieval of target-knowledge, then the learner's performance in respect of their strategic use of the clues also becomes a determining factor in the overall learning outcome. Quantitative measures of performance can be used to adapt the interaction via the creation of a learner model which guides the organisation of testing according to psychologically -based principles of rehearsal and review (ie: of spacing and recency). In a conventional form of this interpretation, performance would be considered mainly as the incidence of 'correct' retrievals of target -items, but within the framework of the kind of 'process' definition of vocabulary-learning which has emerged in this discussion, other aspects of quantitative performance are equally relevant. These include the number of items selected, the numbers and distribution of items grouped,. and the total amounts of processing of various kinds involved at each of the three stages of the interaction. The structure of this 'process-based' learner model will be considered in the next chapter.

2.5 Feedback

Feedback is the means by which the learner judges the quality of their performance. It is usually either intrinsic, wherein they can see the result of an action (eg: the system's 'change-of-state' itself confirms that the action is correct), or extrinsic, where there is a commentary on the performance, such as a 'correct' signal, or an updated score, or an explanation of why something is incorrect. Intrinsic feedback in a language-learning system implies some kind of language understanding on the part of the system, but this, like adaptation on the basis of qualitative assessment of performance (see 3.4 above), can be considered beyond the capability of current CALL technology to provide for any complexity of meaning. For this design, therefore, feedback should be extrinsic and operate on learner performance at the testing stage. Basic feedback, intended to maintain the learner's awareness of the quantitative objectives of the interaction, could be provided by a score representing the percentage of selected items which have been successfully retrieved. Further extrinsic feedback should, however, be given on the nature of retrieval attempts, meeting the requirement to enhance learner input at the testing stage.

Section 3: Chapter Summary and Design Considerations

The aspects of CALL-for-vocabulary design which are the most relevant to the system under consideration, are those which address: the complexity of processes involved in vocabulary-learning, the nature of embedded pedagogical principles, the degree of adaptation to the individual learner, and the optimal use of the power and flexibility of the technology. These considerations can be translated into design principles for an approach to CALL-for-vocabulary design which capitalises on theories of the mental lexicon, utilises on-line lexical tools, and supports individualised processing of target items, adapting to the learner in aspects of the interaction concerned with testing and feedback.

In general, the proposed design whould embody a strategy for vocabulary learning which is based on reading and item-selection, exploration of lexical relations and item-elaboration, and the use of self-constructed clues for item-retrieval. The

strategy should be supported by the design, though not explicitly taught, because the aim is for the learner to develop an awareness of its effectiveness, through exploration of the functions of the design in the pursuit of improvement in their performance. The three processing stages in sequence: selection, elaboration and testing, represent a single cycle in the overall learning process. Over a period of time these stages are iterated according to the learner's developing appreciation of the global strategy. To support this iteration, processing information should be recorded in a learner-database, such as: selected items, their locations in originating texts, item-definitions and translations, item-groupings, indexes representing correct and incorrect retrieval attempts. The same information, in an abstracted form, can be used to build a learner model, to guide the sequence of presentation of items in the testing stage. The learner database should be constantly updated, so that the results of processing accumulate and are available to be inspected in a relevant form at each stage. At the stage of selection they can access any text that has been used as a source, see a list of target-items and locate each item in its respective text. At the lexicon-building stage they can display and amend the definition and grouping information for any item they have already processed. At the testing stage they can see a representation of their success in retrieval to date. The learning rate thus reflects the accumulation of all this information, as well as the incidence of successful retrieval. Other performance data, classified according to the processing stage in which it occurred, may also be recorded and used in the theoretical analysis of learning processes. This is further discussed in the next chapter.

CHAPTER THREE: OUTLINE SPECIFICATIONS FOR A DESIGN AND PRINCIPLES FOR EVALUATION

Introduction:

The aims of this chapter are to specify the main design features of a vocabulary-learning program that will satisfy the requirements outlined in Chapters One and Two, and to establish a basis for evaluating the vocabulary-learning it supports. Because the learning strategy is not going to be 'taught' to the learner, (in the sense of giving them an explicit sequence of procedures to follow), because this would contradict the principle of learner-determination of content and process, it has instead to be 'afforded' by the interaction, and implicitly promoted in the way the design supports the realisation of local learning goals. These specifications therefore focus on means of making these goals explicit to the learner and of promoting the kinds of skill that pursuing the goals entails, and on the adaptation of the interaction to the individual learner. The central issues for evaluating the design are how to determine: a) the extent to which aims and processes implied by the strategy are evident in the way the learner uses the program, and b) the degree to which quality in the learning outcome is associated with those features of the strategy which are present. The interaction will therefore need to be recorded, and an approach to the interpretation of this data devised. As the learner's use of the program has both conscious and unconscious aspects (eg: selecting target words is largely conscious, recalling them unconscious), their actions need to be interpreted in terms of a model which takes into account both kinds of processing. Also, as learning is realised in both process and product, assessing the quality of the learning outcome has to proceed within a theoretical framework which relates one to the other.

Section 1 will outline the main features and functionality of a proposed design, drawing on a psycholinguistic model of word-production for an account of how the interaction can be made adaptable. Section 2 will consider the question of the interpretation of data recorded by the program during interaction with a learner, and propose the additional need for a subjective source of information about what they

are actually doing. In Section 3 issues of quality in the learning outcome are dealt with, in the context of a framework for evaluating the approaches which the learner brings to the various aspects of the task which the design represents for them. This leads on, in Section 4, to a general statement of evaluation criteria for the interaction, which will subsequently be put to the test in pilot studies (Chapter Four) and a main test programme (Chapter Five).

Section 1: Features of a Design

1.1 Learner Selection of Target Items

The requirements derived in Chapter One from the consideration of the nature of vocabulary learning, emphasise the learner's own definition of the required learning content. This principle was operationalised in CALL terms, in Chapter Two, as the selection by the learner of target items which are locatable in their context, as part of the process of reading a text, eg: Fig.3.1 below.

Fig 3.1: Selection of Target Items


```
texttexttexttexttext          word1
texttexttexttext              word2
texttexttexttexttext          word3
texttexttexttexttext          ...
texttexttexttexttext
.....etc


SOURCE TEXT                   TARGET
WITH A TARGET                 ITEMS
ITEM HIGHLIGHTED
```
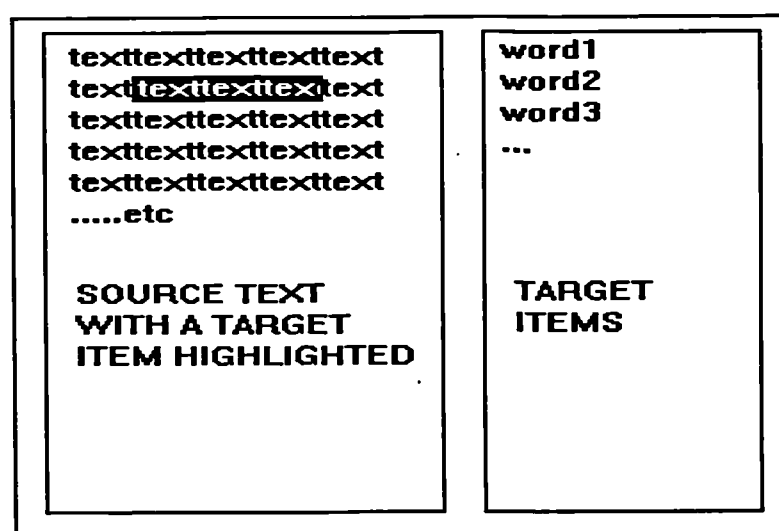
The skills involved in this process of highlighting target words in the text and transferring them to the target-word list, include the ability to infer the meaning of an unfamiliar item from the context, and to judge the relevance of an item to the lexical-structure-related aims of the learning strategy, eg: on the basis of its orthographic, morphological or syntactic patterning. This screen could be part of an

initial system module, corresponding to the first of the 3 stages proposed in Chapter Two (ie: selection, elaboration and testing).

1.2 Elaboration of Items

The second stage would address Chapter One's requirement for richness and scope in the learner's conscious semantically-related processing, together with the specification in Chapter Two for access to definitions, translations, synonyms, collocations etc. via on-line lexical tools, and support for grouping of items according to structural similarity. The learner's recording of descriptions of their target-words, and their assignment of them to groups, might be realised as shown in Fig 3.2:

Fig 3.2: Elaboration of Target-Items
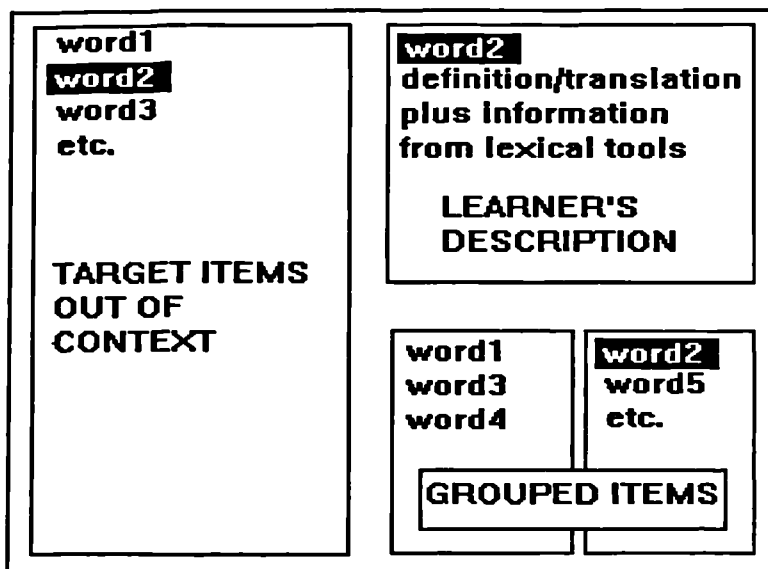
The skills involved in these processes include the ability to extract relevant information from a dictionary definition and to recognise relevance in a translation, to employ knowledge about the composition of words to identify root forms, to spot significant syntactic or collocational regularity in a list of KWIC citations, to identify useful conceptual relations between elements of the meanings of different words, to determine whether grammatical or spelling features are significant as structural relations etc. and to devise group titles encapsulating the similarities in the items contained. KWIC concordancing would allow learners to widen the range of

their reference to existing knowledge, and to include familiar collocations and knowledge of affixation (using 'wild-cards' would enable them to experiment with searching for root forms of target items). The facility to record their own notes on this information would also encourage learners to relate what they have found out to what they already know, possibly by experimenting with the writing of notes in the target language itself. The design should provide an easy and intuitive means of indicating items which have features in common, and of collecting them together, and entering a heading for the group. This would promote reflection on existing knowledge (of semantic concepts in L1 as well as L2) and the application of existing schema to new items.

1.3 Testing of Items

The third stage addresses the requirements in Chapter One for psychological appropriateness in the means of testing and rehearsal, and the promotion of new items from the 'potential' to the 'threshold' level of productive knowledge. Chapter Two operationalised this as cued productive retrieval of items, utilising associations created between items and their meanings, related grouped items, definitions, L1 translations etc. as clues to assist in productive retrieval. Fig 3.3 illustrates this. Skills at this stage mainly relate to the ability to 'reconstruct' a forgotten form from available clues, ie: the item-definition which the learner has written themselves, the KWIC contexts in which the item occurs (obviously with the item itself omitted), and relevant groups which the learner has created (again with the item itself omitted). Other more conventional types of clue may also be made available, eg: a 'hangman-type' clue showing the number and arrangement of letters in the item to be retrieved. Following Chapter Two's specification of feedback requirements, the system should provide feedback in the form of a score recording successful retrievals, and also give information about 'near-misses', so that the learner is able to subsequently reconstruct an answer. A further requirement, arising from the wish to promote associations within the target lexicon, is to give feedback on learner inputs which match items other than the current target-item. Such feedback could

be expected to reinforce general areas of the learner's target-lexicon even if it did not lead to the immediate retrieval of the required item.

Fig 3.3: Testing

```
┌─────────────────────────────────────────────┐
│  ┌───────────────────────────────────────┐  │
│  │ texttexttexttext..........texttexttexttext │  │
│  │                                        │  │
│  │ KWIC CITATION FROM SOURCE              │  │
│  │ TEXT WITH TARGET ITEM GAPPED           │  │
│  │ OUT                                    │  │
│  │                                        │  │
│  └───────────────────────────────────────┘  │
│  ┌───────────────────────────────────────┐  │
│  │                                        │  │
│  │ GROUP CLUE: word1 word3 ........       │  │
│  │ DESCRIPTION CLUE: definition etc.      │  │
│  │ ETC.                                   │  │
│  │                                        │  │
│  └───────────────────────────────────────┘  │
│  ┌───────────────────┐ ┌─────────────────┐  │
│  │ LEARNER INPUT     │ │   FEEDBACK      │  │
│  └───────────────────┘ └─────────────────┘  │
└─────────────────────────────────────────────┘
```
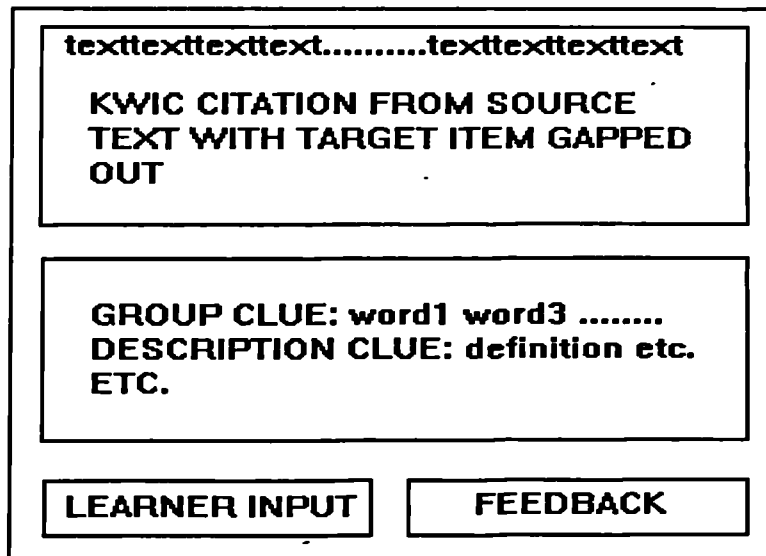
1.4 Adapting the Interaction

The idea that the focus of cognitive effort can be an area of the mental lexicon, as well as a specific item in it, is the basis of the requirement for the interaction to be adaptable to the individual learner, as it is clear that whilst two learners may select the same item, the associations it has for each (including the other items it is selected with, and those it is grouped with) may be very different. Given the assumption, in Chapter One, of retrieval practice effects resulting from the sequencing of target items being tested; if we can establish how items might affect each other's retrievability, then it may be possible to determine sequences of presentation which are optimal for a particular individual, based on a model of the state of their lexicon as a whole. What is needed, for this, is a psycholinguistic model of how words are stored and retrieved in the mental lexicon, compatible with the assumed effects of priming and recency, which can be used to inform the sequencing of items for testing. The model should be consistent with observable facts of lexical performance, such as slips of the tongue and other common errors, and should allow for the representation of its main elements in computational form.

62

Formal models of production are, however, not common in second language learning theory (Matthews 1992, p.25, De Bot 1992, p.1), but first language investigation does provide some insight. De Bot (op.cit) suggests that Levelt's (1989) account of speech production is sufficiently global, and well enough grounded both theoretically and empirically, to be adaptable to the bilingual case. Although it is a theory of spoken and not written production there do not seem to be any grounds for assuming that it cannot account for the latter, given that the main point of interest here is the structure of the lexicon rather than mechanisms of articulation. Channell (1988), for instance, summarising Fay & Cutler (1977) on the mental lexicon, does not distinguish between the 2 modes of production. The main relevance to L2 of Levelt's model, according to De Bot (op.cit, p.11), is that it is consistent with a single lexical storage system in which lexical items from specific languages and registers form closely linked sub-sets which "..fit one context of discourse better than another..", but which may nevertheless be interconnected, accounting for cross-language performance phenomena such as code-switching in balanced bilinguals. In Levelt's formulation, a 'pre-verbal' message, equipped with information derived from discourse and situational knowledge about which language it is to be encoded in, is encoded into a lexical item in two overlapping stages: first the semantic content of the message is matched, in the lexicon, with an appropriate set of semantic and syntactic features (a lemma), then the lemma activates (also within the lexicon) an associated set of morpho-phonological features (a lexeme). The resulting lexeme is then passed on to another set of processors to be incorporated into a phonological plan.

For the purposes of modelling the learner, the most important aspects of this theory are the 2-level structure of the lexicon, and the account it provides of the operation of the selection of items for production. According to Dell (op.cit), production of items is achieved via a 'spreading activation' mechanism (Quillian 1967, Collins & Loftus 1975), which works as follows: when an item (either at lemma or lexeme stage) is being matched it is temporarily active. Some of this activation (the level of which ultimately determines its likelihood of being selected for output) is immediately passed to its neighbours via the network links. Activation travels in

both directions along a link so there is a 'rebounding' effect (Dell 1986, p.288), resulting in the most highly connected items in the active part of the network accumulating the highest temporary levels of activation. This gives these items a high likelihood of being selected for output. When several forms have high activation there is a kind of competition between them to be selected. The overlapping of the 2 encoding stages (meaning and form), is what accounts for the occasional unpredictability of the output, as highly-connected but inappropriate lexemes can sometimes compete for selection with the intended ones. This is manifested in L1 in slips-of-the-tongue (eg: word substitutions, spoonerisms etc.) and hiatuses (eg: word loss, tip-of-the-tongue phenomenon), and in L2 by nonce forms and composites.

This is a very brief account of a complex model, which does not address many of the important questions concerning the relation between meaning and form, the nature of the links, the role of factors related to more general memory functions, such as repetition, and the actual processes involved in activation. (A clear and thorough explication of the workings of the model can be found in Dell op.cit). It is a 'steady-state' model (De Bot op.cit, p.3) which has nothing to say about learning, but instead describes the ordinary operation of spontaneous L1 speech production. Nevertheless, it has been generally accepted as a plausible description of processing in the mental lexicon (eg: Aitchison 1987), and has been used as evidence that second language learners have a "..formidable task in emulating the complexities of L1 storage in an L2.." (McCarthy 1990 op.cit, p.42). Its usefulness for the present argument derives from the fact that it is directly relevant to the local performance goals of the system in question, ie: to create links amongst target vocabulary items, and to recall the items productively. It is therefore a plausible theoretical basis on which to construct a record of the status of individual items within the learner's personal network, as follows:

According to the theory, two factors determine whether an item is produced. One is the item's 'base' level of activation (eg: the state it is in when the matching process begins), and the other is the number and types of links it has to other items in the

network. An item which has few links is less likely to be selected, unless it has a high base level. An item with many links is more likely to increase its activation via its neighbours, even if the base level is low. It follows from this that, in the attempt to produce an item from memory in response to a given prompt, the current state of the network as a whole may be just as significant as the nature of the prompt, and that an item which has failed to be recalled via one path of activation may succeed via another.

The learning strategy therefore aims to maximise connectivity in the network, and to maintain the base activation levels of items as high as possible. In order to do this the program must keep a record of the associations that every item has, as well as an index of its current state of activation, and use this record to optimise the paths to recall. We have seen (in Chapter One) that certain links between words are considered to be more 'salient' than others, eg: synonymy, antonymy. These are the explicit links that the learner will be expected to create in the context of the local performance goals, on the assumption that the effect of making similarities in meaning and form explicit, is to render the words involved more accessible for production. Work on lexical decision and naming tasks has also shown that other types of association, such as general semantic context, sentence context, and episodically established linkages (connections made through personal experience) are similarly effective in priming responses (Forster 1990). The existence of these priming effects suggests that items which have such implicit associations can also be regarded as having higher base activation levels. On the basis of this kind of theoretical approach, we can specify some aspects of an implementable model of the learner's lexicon, wherein the amount and type of processing a particular item has received reflects its levels of activation and association. The system represents the structure of the learner's target-lexicon as follows:

a) As activation of items is increased in each of the selection, processing and practice stages of the interaction, the system maintains, for each item, an index representing the amount of processing it has received at each of these stages.

b) As the explicit linking of items in groups increases the activation levels of the items concerned, the system records how many links each item has.

c) As successful recall increases the activation level of the items concerned, the system maintains an index representing the amount of successful retrieval each item has had.

The likelihood of an item's being recalled is determined partly by its activation level and partly by the links it has. Given that an attempt to retrieve an item will also tend to activate others with which it is associated, items identified by their levels of processing as highly connected and/or highly activated, can be used as aids to the recall of weaker ones, by being interspersed with them in the test sequence. This is the basis on which the system is able to adapt to the individual learner, by generating a test sequence based on the assumed structure of their particular target lexicon. Details of how this may be done are discussed in the next chapter.

1.5 Section Summary - The Strategy and the Data it Produces

Throughout the interaction, procedures are supported which focus on the development of an overall vocabulary-learning strategy, and on the skill of manipulating linguistic information resources, as well as on the linguistic information itself. The learner makes conscious decisions to move from one stage to the next, gaining some program functions and losing others with each move, and is thus encouraged to reflect on performance at each stage. As their awareness of the learning content develops, the strategies they apply to the exploration and elaboration of target-items should come to explicitly influence decisions about which vocabulary items to select. Clue-using strategies at the retrieval practice stage should reinforce memory for the item-contexts and enhance subsequent comprehension of the texts themselves. Improved comprehension should assist the application of inferential strategies and enhance perception of lexical relations within the text.

The learner's mental lexicon - the connections between items, and assumed item activation levels - is modelled in the form of data recording the iterative processing they do in selecting, elaborating, linking and recalling each target-item. This data is used to determine the sequencing of items for testing, so that assumed effects of recency of exposure, priming, local areas of activation etc. can be exploited to maximise the probability of recall. The data is also available to be analysed in terms of quantitative measures characterising the learning outcome, such as total numbers of items selected and retrieved, numbers of links, spread of processing across the whole target-item list, amounts and rates of processing of individual items, and qualitatively in terms of the respective importance to the outcome of different kinds of processes.

Section 2: Interpretation of Interaction Data

The idea that data generated in the interaction between learners and computers can provide us with information about language learning processes, has created a lot of interest (eg: Garret 1991, Sussex 1991, Chapelle & Jamieson 1989, Jung 1987). The attraction is that the computer's ability to record complex processes accurately and unobtrusively suggests that we can use it to tell us exactly what learners do when they learn. It is proposed, for example, that the computer can 'pile up' language acquisition data relating to the frequency and quality of student errors (Jung op.cit). Or that keystroke data can be used in the description of students' monitoring strategies (Chapelle & Jamieson op.cit). CALL environments could provide a 'cognitive workbench' for research (Sussex op.cit). They could provide 'detailed and sophisticated assessment instruments' to be used in the evaluation of our theories of language learning (Garrett 1993). There are, however, relatively few on-going attempts to do any of this for computer-based vocabulary learning processes (I have already mentioned, in Chapter One, the embryonic Mayday, Dicologique and Contexts work). Pennsylvania State University (Lyman-Hager op.cit), have conducted studies comparing the retention of vocabulary in learners who use a CALL program for reading comprehension with those who use conventional paper dictionaries. Their conclusion (op.cit, p.96) that the computer-

users are better able to retain the vocabulary they have learned is, as yet, only the starting point for further investigation into whether this is due to qualitative differences in the learning processes. McKee (1993) reports on research done at the University of Ulster into learners' use of text-analysis tools. Retrospective reports by the learners revealed positive responses on each of the key questions: whether the tools made them look at texts more closely, whether they increased their range of vocabulary, and whether the tools had helped in the development of vocabulary acquisition techniques. However, the authors are reluctant to draw any conclusions about learning processes as such (op cit, p.15).

One of the problems with the attempt to interpret the contents of 'log-files' or 'session-histories', as some writers have pointed out (eg: Chapelle 1990, Laurillard & Manning 1993), is that however thorough and detailed a computer's record may be, it does not describe what the learners think about what they are doing. This can lead to misconceptions, and inaccurate interpretation of results. CALL programs can collect data, but what is also needed, if we wish to interpret the data in terms of learning processes, is information relating to the subjective dimension of the learner's approach to the task. The learner's approach describes the relation between the student and the learning task (Ramsden 1992). In CAL it is based on their perception of the task as represented by the interaction design, and is likely to involve alternative strategies and learning theories to those the program design presupposes; it may even come into conflict with them. Thus, even where we have a principled cognitive model underlying the design of the interaction, a further understanding of what the learner thinks they are trying to do is essential if we are to describe their learning. Evaluation of CALL interactions has to address the issue of bringing together performance data and the learner's introspections. These have to be considered with regard to the pedagogic intentions of the program, and the quality of the learning outcome.

For the evaluation of the vocabulary-learning system in question, a detailed record of the learners' performance is required, but because the learning strategy is essentially autonomous, subjective data from the learners themselves is also

necessary. Whilst performance at the local level (ie: in selection, elaboration and retrieval) can be monitored by the system and evaluated on a quantitative basis, related to criteria such as the 'optimal' learning-rate discussed in the last chapter, performance at the global level, (ie: strategic iteration and integration of learning processes), will need to be monitored qualitatively, using the introspective data, and evaluated in terms of the overall learning outcome. In order to do this a theoretical framework for the qualitative evaluation of learning is required.

Section 3: Learning Approach

3.1 Deep and Surface Approaches

The concept of approach to learning describes how people experience and organise the subject matter of a learning task (Ramsden op.cit, p.42). The approach determines the quality of the learning outcome, in the sense that it defines how the learner sees the learning content, and thus constrains what they can do with it. Within the interaction in question, for example, the learner is required to select unknown vocabulary items from a text. Clearly this could be done without much reference to the meaning of the text, if the learner approached the task in that way. They could simply make a list of words they did not know. However, even if they subsequently managed to retrieve every one of the words on their list, we would probably regard this outcome as being of lower quality compared to that of another learner, who had managed to understand the text itself, as well as learning the new words. One of the local goals of the interaction is to understand the source text, and we assume that learners will attempt, to some extent, to do this. But they could equally view this aspect of the task as a matter of translating the words in the text into L1, adding a second list, of L1 equivalents, to the first. The quality of the outcome here would probably be inferior to what might result from an attempt to paraphrase the text in L2. One of the global goals of the interaction relates to the learner's understanding of the systems of lexical structure in the target language, and we expect them to develop ways of investigating this. L1 translation is only of limited use in the perception of L2 structure. The optimal approach involves them in looking for abstract relations between the L2 target-items, some of which will be

realised through the roles the items play in the originating text. So in order to achieve the best quality learning outcome, ie: to meet the goal of structural awareness, along with the other quantitative and qualitative objectives of the strategy, the learner has to approach the words in the text as linguistic phenomena, as well as carriers of an encoded message.

3.2 Approaches to Vocabulary-Learning

Different approaches to the task of vocabulary-learning, may be described in terms of the distinction between 'deep' and 'surface' learning, that has been used in investigations into the learning processes of students in higher education (eg: Laurillard 1993, ch.3, Ramsden op.cit, ch.4, Marton 1986, Marton & Säljö 1984). This work defines a deep approach as one in which the learner attends to the significance of the task and is able to see the internal relations in the way that the knowledge is organised and/or described, such that the eventual learning outcome preserves the original meaning and structure of the content. A surface approach is one which focuses on the component parts of the task and on procedures which lead to outcomes at that level only (Ramsden op.cit, p.43). Marton (1986 op.cit) gives the example of subjects reviewing a text in which the case of a problem family with an alcoholic father is described in the course of an argument for a more holistic view of the Swedish welfare system. Subjects taking a surface approach tended to understand the text as being about the problem family, those taking a deep approach understood that the example was subordinate to the message about the welfare system. The second group perceived the structure of the text in terms of its component parts. The first group saw only the parts. Thus there is a direct relationship between the quality of the learning process and the level of learning outcome.

This analysis applied to vocabulary-learning may help us to evaluate learning outcome by looking at the learning process, but it is not without complications. It is clear, for example, that comprehending the message-structure of a text, as described in Marton's example above, is only possible if you can understand the text

70

linguistically. We are looking, therefore, to apply the deep/surface analysis to the learning approach at this linguistic level. The quality of the approach to understanding the language, is what makes a deep approach to understanding the message, possible for the foreign language-learner. As I have observed, a learner may adopt different approaches to the linguistic task of understanding unknown words, eg: by inferring their meanings from context, or by looking up each word in a bi-lingual dictionary. The former would equate to a deep, the latter to a surface, approach to the task of understanding the words. The learning outcome of the deep approach should be qualitatively superior to the surface one, because the learner's reading ability is enhanced by the use of inference strategies (eg: Twaddell op.cit, Nation & Coady 1988). But, as Hulstijn has pointed out (op.cit, p.123), inference strategies also run considerably more risk of incorrect understanding, than dictionary lookup or other types of 'given meaning'. A deep learning approach, linguistically defined, may therefore not necessarily result in a correct understanding of the text, any more than will a surface one.

To illustrate this, we can look at the Spanish text which is used for Pilot Study 2 (described in the next chapter). This text (reproduced in appendix IA) is a report about a World Health Organisation study entitled "Los Años Cruciales de la Adolesencia" (The Crucial Years of Adolescence). Whilst the text refers to several of the well-known problems associated with adolescence, eg: drugs, driving behaviour etc., its central message is that these are due to the rapid and radical changes in the composition of society brought about by the industrial revolution. The opening sentence begins as follows:

> Los adolescentes pagan un alto tributo por las ventajas que aporta la moderna tecnología en diversas partes del mundo y la mayoría de los esfuerzos realizados para cambiar esta situación y favorecer a los jóvenes que han fracasado en la presente sociedad postindustrial...etc.

For an English-speaking Spanish-learner of intermediate level, the comprehension task is likely to include the decoding of the vocabulary items 'tributo' in the first line, and 'fracasado' in line 3. A surface approach to this task, as I have described it above, might be to simply isolate the word and look it up in a bilingual dictionary.

Microtak's Spanish Assistant on-line dictionary gives 'tribute' for 'tributo', and two senses for 'fracasado', the first a noun meaning 'failure' and the second an adjective meaning 'unsuccessful'. The outcome of a learning approach which stopped there might be the assumption that 'tributo' is cognate with English, and the formation of a mental association between 'fracasado' and the concept 'fail(ed)(ure)' with the actual syntax of the translation left in some doubt. A deeper approach might be to look at the immediate co-text and/or the internal structure of the unknown words. In the case of 'tributo' the 2 preceding words, if they were known, might effect the inference of its meaning - 'pagan' (they pay) appears to reflect the English collocation 'pay tribute', 'alto' (high) on the other hand, suggests the English collocation 'high price'. The deep approach might be to reserve judgement on whether the adolescents were paying tribute to, or a price for, modern technology, until more evidence became available from the context. Considering 'fracasado' it might then be observed: that it is similar to the English (French-derived) word 'fracas' meaning a noisy quarrel, that it occurs with an auxiliary verb ('han') and is therefore probably a verb itself, that the preceding co-text, '..los jóvenes que..', contains the word 'jóvenes' (young people - probably known because it is relatively frequent), and that 'los jóvenes' is the subject of 'fracasado' and that the latter refers to some negative/disruptive aspect of young peoples' social behaviour etc. This interpretation would be in line with what appears to be the overall theme of the text - the problems Society has with adolescents - and might reinforce the interpretation of 'tributo' as meaning 'tribute' in the sense that it is young people's embracing of modern technology which has caused them to become uncontrollable. In fact the literal meaning of the part of the text quoted above is as follows:

> Adolescents pay a high price for the advantages that modern technology brings in different parts of the world and the majority of efforts made to change this situation and favour young people who have failed in the present post-industrial society...etc.

The outcome of such an approach, despite its failure to decode the actual message, is clearly a deeper level of processing, and the associations generated are likely to be more complex, involving intra- as well as inter-language relations, and formal (syntactical, orthographic) elements as well as semantic ones. However, if we were

looking, as Marton did, at the ability of the learners to understand the message, we should have to say that neither of the approaches in this example would have guaranteed this outcome. Thus, because of the added difficulty in producing accurate meaning, in the context of L2 comprehension, the deep approach should be interpreted here as being <u>more likely</u> to produce accurate meaning at the level of the whole text, by virtue of <u>actually</u> producing an accurate structural analysis of it. Because, as we are also interested in their developing appreciation of the way words are related linguistically, we would say that the second approach works better, and if extended could create more potentially useful associations between the target words in context, and the rest of the text, eg: the thematic associations between the sympathetic connotations of 'fracasado' in this context (have gone wrong) and other words such as 'perturbado' (disturbed), 'padecen' (they suffer), 'la seguridad personal' (personal security), 'la salud' (health) etc. It also works better in terms of the learner's potential to put right initial misconceptions, because the more these apparent relations to the wider text are considered, the more the structure of the text itself becomes evident. For an intermediate-level learner, the cognitive effort required to sift out the message contained in the complex clause structure of the opening (70-word) sentence in this text is considerable, but the deeper approach gives them a chance to eventually succeed, whereas the surface approach gets them no further than a mental (and possibly unstable) list of L1-L2 associates.

Finally, the deeper approach is not restricted to the originating text itself. Skill in spotting lexical and grammatical relations can be applied in different kinds of context, such as those provided by the entries in a dictionary, or by the output of a KWIC concordance. If, as (Marton op.cit, p.2) has suggested, the point of learning is "..to improve one's understanding of some phenomenon in the world...by means of taking part in a symbolic representation of that phenomenon..", then where the phenomenon to be understood is the structure of lexis in a foreign language, the symbolic representation is not only the text(s) in which the words are embedded, but also their representation in dictionary, thesaurus, and KWIC citation. Taking part in the representation is not only the comprehending of the structure of the text

but also the comprehending of the structure of the information contained in these resources.

A deep approach to vocabulary-learning can, therefore, be realised at a level below comprehension of the text as a whole. The deep-surface dichotomy can be applied to the approach of the vocabulary learner with regard to their intention to discover and make explicit some underlying structure among the lexical items they have selected, regardless of whether this leads immediately to a deep understanding of the message of the text. The general characteristics of a deep approach can be interpreted for L2 as: the use of strategies of inference from context, the exploration of features of lexical structure, the association of L2 words with each other. A surface approach is interpreted as: focus at the word-level only, the intention to find a precise meaning-equivalent, and the association of L2 with L1. Qualitative assessment of the presence of these characteristics in the approach of an individual learner, can be made on the basis of data provided by the learner's own introspections into how they view the particular aspect of the task which is current. Whilst it is likely that the same learner will exhibit different approaches to different aspects of the task, depending on the nature of the task as much as on the individual's learning style, their overall tendency to adopt deep or surface perspectives should be a clear pointer to the general quality of the learning that is going on. Part of the evaluation of the system, therefore, will be to investigate the presence and pattern of occurrence of these characteristics.

Section 4: Chapter Summary - Preliminary Evaluation Questions

In this chapter I have proposed general features for a system design which is intended to support and promote a strategy for vocabulary-learning. Following the strategy, the learner adopts a deep approach to the selection of target-items, making judgements about the contribution the items make to comprehension of the source text, and the lexical features apparent in their internal structure and that of the surrounding co-text. This process introduces new nodes into the mental network, and generates weak associations between them and existing areas of the network

containing similar semantic, formal, or contextual links. The learner builds their lexicon, paying attention to semantic and formal structure and using information researched from the lexical resources to incorporate the new items into familiar linguistic categories, and to propose new categories whenever new features are uncovered. This generates activation in the parts of the network connected to the target-items, strengthening associations and increasing the range of sources from which priming or prompting stimuli may be received. In this way the learner consciously generates clues for themselves to use in subsequent productive practice of their target vocabulary. The system sequences the target-items so that the effects of association, activation and recent retrieval are exploited to maximise the priming effect that one item and its associated clues may have on another following it. The learner adopts a deep approach to the use of the clues, focusing on L2-L2 associations to prompt recall, and to interpret system feedback in terms of the structure of individual items and the network as a whole. The outcome is increased comprehension of the source texts, and a steadily-growing collection of L2 items, reference notes, L2 groupings displaying an increasing amount of cross-referencing, and a high rate of successful retrieval.

It is proposed that a model of the mental lexicon as a network of lexical items connected by features of structural similarity, and activated according to the processing that has been done on them (including the links they have to other items), may serve to represent the learner's knowledge. A model based on this theory will be used to record quantitative performance data and to determine the sequence of presentation of items at the testing stage, according to principles of recency, expanded rehearsal, and the assumed operation of spreading activation. It is proposed to adopt the notion of deep and surface approaches to learning as a framework for evaluation of the quality of the learning processes and learning outcome. Qualitative data relating to the learner's apprehension of different aspects of the learning task will be recorded and interpreted as an indication of the extent of their perception of structural features underlying the learning content. These issues will be addressed in the pilot and main studies described in the following chapters.

It is clear that, to ensure the generation of useful quantitative and qualitative data of the kind described above, the design must provide a rich learning experience, involving the learner in the deployment and development of their linguistic knowledge. The learning environment can be characterised in many ways (eg: O'Malley & Charnot op.cit, pp.187-188), but there are two essentials if it is to be successful in the promotion of a deep, autonomous, learning approach: a) it should satisfactorily support management of the learning process, related to the learner's goals of understanding the learning content and regulation of their own learning, and b) it should support successful planning, processing, consolidation and extension of learning, related to the integration of new information with prior knowledge. These requirements can be summarised as a set of preliminary 'evaluation questions', to be applied to an initial implementation of the design, with the intention of establishing, in pilot tests, whether it provides such a learning experience, and whether learners are able to manage the processes which it aims to support. The questions are as follows:

i) Are the learning objectives and procedures clear to the learners?

ii) Do they engage with the goal of understanding a) the source texts, b) the individual items, and c) L2 lexical relations?

iii) Do they engage with the goal of regulating their own learning?

iv) Does the learning process involve the integration of new with prior knowledge?

The next chapter reports the results of pilot tests addressing these questions, and finalises the design of the system.

CHAPTER FOUR: IMPLEMENTATION OF THE DESIGN AND PILOT-TESTING

Introduction

The design for a computer-based vocabulary-learning strategy for foreign languages, outlined in the preceding chapters, has been implemented as a program in two versions. The first version was a straightforward realisation of the three stages of the basic strategy, ie: selection of items, elaboration (lexicon-building), and testing. It is described in terms of its general architecture in Section 1. This implementation was subjected to a series of pilot tests, to establish whether the interaction it supports is capable of providing a rich learning experience for users. The specific objectives, procedure and results of these tests are discussed in Sections 2 and 3. The second version of the program incorporated improved and more complex design features suggested by the findings of the pilot tests. The functionality of this version is described in detail in Section 4. The improved design was then tested to see if the interaction would support an in-depth analysis in terms of the 'deep/surface' learning model discussed in Chapter Three. The results of this test are reported and discussed in Section 5. The last section sets out objectives for the main test programme which will be reported on in Chapter Five.

Section 1: First Implementation - General Architecture

The program has three learner-input modules: selection, lexicon-building , and testing. These are supported by a corpus of text, a lexical tool-set consisting of a bilingual dictionary and keyword-in-context (KWIC) concordancer, and a learner database incorporating a learner model. The modules can be accessed independently via a graphic user interface. The general architecture is shown in Fig 4.1.

Fig 4.1: General Architecture of the System
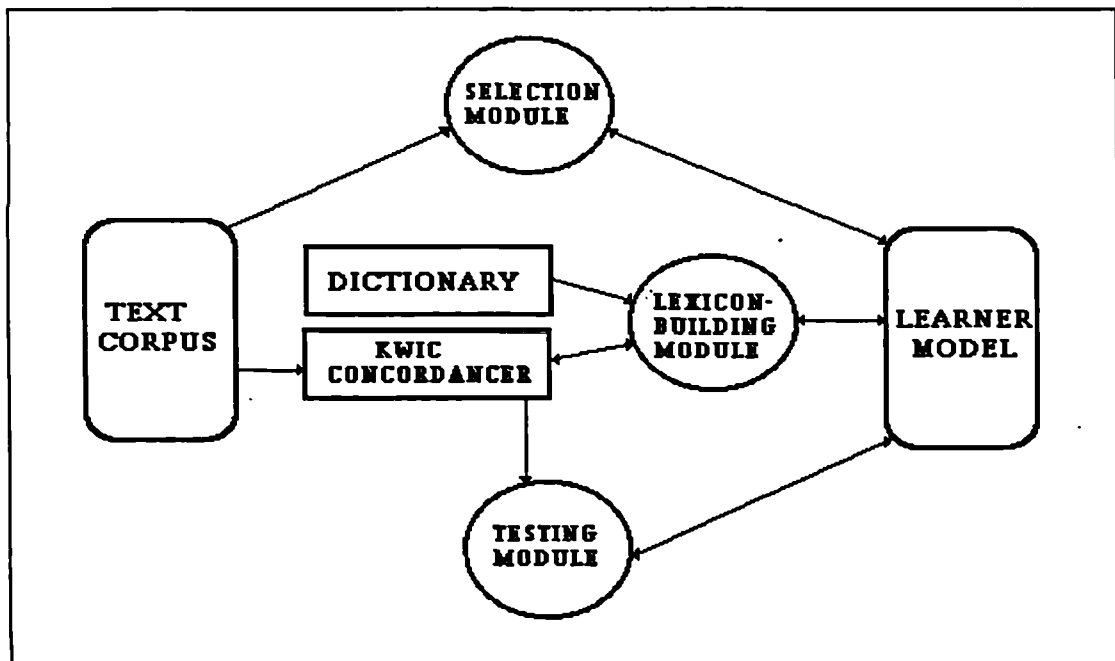
1.2 Modules

This description covers only what is necessary to appreciate the general nature of the interaction which was the subject of the pilot study, (it is described in detail in Goodfellow 1993). The final design is discussed in detail in Section 4 below.

a) Text Corpus

The corpus consists of approximately 50,000 words divided into texts of between 300 and 2000 words. The texts are stored as separate files which can be accessed individually by the learner, displayed in a scrollable window, and selected for the purpose of extracting vocabulary.

b) Selection Module

One of the texts is selected by the learner and displayed. Sections of it (words, phrases, whole lines) can be highlighted and put into a list of vocabulary items for study (target-items). The target-item list itself can be scrolled and individual items

highlighted. The source text for a target-item thus highlighted is automatically displayed and the item located in it.

c) Lexicon-Building Module

A list of texts which have been used for vocabulary-extraction is displayed. Choosing one causes all the target-items originating from that text to be displayed. Individual target-items can then be highlighted and sent as keywords to the concordancer or referred to the dictionary. Notes (eg: translations or L2 synonyms etc) written by the learner can be attached to the items. Target-items can also be grouped under titles invented by the user. Grouped items are linked to each other in the learner model (see f below).

d) Testing Module

The target-items are sequenced for testing by the learner model. The learner is then prompted for each item in turn with the following clues:

i) A KWIC list citing all occurrences of the target-item in the corpus, each citation with the item itself gapped out.

ii) The titles of groups to which the item belongs (if any were created), together with other items which are in the same groups.

iii) The notes which the learner has written and attached to the item.

iv) A 'hangman-type' clue, giving information about the orthographic 'shape' of the item.

The learner types in the item they have recalled. Their input is matched firstly against the current target-item, then against any items linked to it in the learner model, and finally against the whole target-item list. The best match found is

indicated as feedback. The matching procedure gives a higher weighting to letters at the beginning and end of the word than those in the middle. Feedback is displayed as a percentage, with an explanatory comment which indicates which type of item has been successfully matched.

e) Lexical Tool Set

The KWIC concordancer takes as input a string of text (the keyword string) and finds all occurrences of it in the whole corpus (including occurrences of it as part of a longer word). The context of each occurrence (the citation) is displayed to a depth of 30 characters on either side of the keyword string, together with the name of the text in which it was found. Individual citations can be selected and expanded, the entire text containing the citation then being displayed in a scrollable window. The L2-L1 dictionary displays a window which shows the headwords which are nearest alphabetically to the keyword that has been entered. One of these can be selected and the dictionary then gives an L1 translation of the main senses of that word.

f) Learner Database and Learner Model

Each target-item is stored in the learner database together with the title of its originating text, the notes the learner has attached to it, and the titles and composition of any groups which the learner has created. The learner model contains a representation, for each item, of which other items it has been grouped with, and its 'processing level' (representing its level of activation in the network) which is a value representing the amount of processing which has been done on it, ie: the total number of links it has to other items, plus the number of times it has been annotated, plus the number of times it has been successfully recalled. The model sequences the items for testing, first by selecting the item with the lowest processing level, then by preceding it with an item which has the greatest number of links to it. In this way the assumed operation of spreading activation is brought into effect.

1.3 Summary

The system thus supports users in selecting texts for study, extracting vocabulary items from the texts, looking up and recording information about the items' meanings and contexts of occurrence, associating them in groups according to features of structural similarity, and retrieving them from memory in response to prompts which utilise the information which the users themselves have recorded. It maintains a model of the mental lexicon that the user is thought to be creating, in the form of a record of the associations the user has created between target words, and the amount of processing each target word has received. This version of the system was implemented for the learning of Spanish, with the text corpus derived from authentic sources including the newspaper El Pais. The dictionary is Collins On-Line Spanish-English dictionary. The software platform is Microsoft Visual Basic version 1.

Section Two: Pilot Study 1

As indicated in the previous chapter, and in the introduction to this one, the purpose of the first set of pilot studies was to confirm that the program described above supports a rich enough learning experience for quantitative and qualitative data generated in the interaction to be evaluated at local and global levels of the learning strategy. A further objective was to identify any refinements to the first version, necessary to promote the learning strategy and assist the development of deep learning approaches.

These objectives, and the data which was examined, can be summarised as follows:

i)   Are the learning objectives and procedures clear to the learners?
       Quantitative data: time spent in each of the 3 stages of the interaction.
       Qualitative data: learners' comments on their objectives and procedures.

ii)   Do learners engage with the goal of understanding the learning content?

Quantitative data: numbers of notes and groups created.

Qualitative data: learners' comments on their understanding of texts, words and structures, types of groups created.

iii) Do learners engage with the goal of regulating their own learning?

Quantitative data: numbers of items selected and retrieved.

Qualitative data: learners' comments on their planning.

iv) Is new knowledge integrated with what the learners already know?

Qualitative data: learners' references to frameworks of linguistic knowledge, linguistic considerations in their notes and groups.

In addition to addressing these four evaluation questions, an assessment was made of the evidence, in the learners' retrieval performance, supporting the idea that their knowledge of new items is in the form of a network, and a preliminary evaluation was also made regarding the support the program gives in the achievement of an optimum learning rate for the acquisition of new vocabulary (see Chapter One).

2.1 Procedure

8 subjects (S1-S8) were observed in interaction with the program. All were university-age learners of Spanish in Britain, all but one (S1) currently studying at University. Their L2 reading proficiency levels varied from elementary to upper intermediate (assessed by themselves). 4 had previous experience using Windows-type computer interfaces. The subjects spent an average of 3 sessions of approximately 45 minutes each, (Table 4.1), and the results are thus based on a total of just over 17 hours of program use. The program recorded their input in each of the 3 modules and maintained a learner-model recording their selection, grouping and retrieval of target words.

Table 4.1: Number of Sessions/Times

Subject:	S1	S2	S3	S4	S5	S6	S7	S8
Sessions	8	3	3	2.5	2	2	2	1.5
Total time (mins)	385	151	144	144	114	98	86	70
Average time per session	48.1	50.3	48	57.6	57	49	43	46.6

2.2 Evaluation Questions

i) Are the design objectives and procedures clear to the learners?

Most subjects distributed their time reasonably evenly across the three program modules (Table 4.2) and can therefore be considered to have taken on board the underlying principles of selection, lexicon-building, and testing.

Table 4.2: Distribution of Time (minutes) per Module

Subject/ Module	S1	S2	S3	S4	S5	S6	S7	S8
selection	84 (21.8%)	44 (29%)	33 (22.9%)	48 (33.3%)	22 (19.2%)	28 (28.5%)	19 (22%)	25 (35.7%)
lexicon-building	149 (38.7%)	63 (41.7%)	47 (32.6%)	48 (33.3%)	47 (41.2%)	47 (47.9%)	46 (53.4%)	29 (41.4%)
testing	152 (39.5%)	44 (29.1%)	64 (44.4%)	48 (33.3%)	35 (30.7%)	23 (23.4%)	21 (24.4%)	16 (22.8%)

Successful support for differing individual priorities is suggested by the varying proportions of time spent on each of the three modules by each subject. In general, the more time the subjects worked with the program the more time they tended to spend in the testing module, which is due partly to the logical sequence (you have to have selected words before you can test them), and partly to the emphasis given in the introductory session to initial mastery of the tools (dictionary and concordancer) in the Grouping module. The fact that the lexicon-building activities occupied considerably more time than selection, for most learners, is consistent with the fact that note-taking and group-making involve more complex cognitive processes than simple selection, but it also suggests that, within the time of this study, most learners did not progress to a stage where they were applying these more complex processes at the selection stage, as intended by the strategy.

Nevertheless, most subjects assessed the overall interaction favourably, which suggests that they found the objectives and procedures clear. The comments of 5 of them (Table 4.3) confirm their awareness of the linguistic objectives (S2,S3,S5), or their satisfaction with the procedure (S4,S8).

Table 4.3: Assessment of Objectives

Subject	Comment
S2	"..it makes sense when you know what you are doing...it's not a glorified dictionary, it actually links words...it's a better way of learning..."
S3	"..it's different...I don't normally try to learn vocabulary but the computer is more motivating...it's useful...especially testing in context..."
S4	"..it's logical when you think about it.."
S5	"..I'd probably use it if I had it at home...it's a good way to learn a language..."
S8	"..you don't have to go away to learn things...it's reasonable easy to use..."

Three of the non-computerate subjects, however, made comments whilst learning to use the program which indicate that there is a potential for cognitive overload arising from the complexity of the interaction (Table 4.4).

Table 4.4: Comments on Complexity

Subject	Comment
S4	"..there's a lot to learn and remember.."
S6	"..so many things you can do...it would take a long time to master..."
	"..you don't remember the different things...but you need to refer to them.."
S7	"..there are so many things to remember.."

The main implication for the design, arising from these comments, is that procedures need to be clearly defined and labelled, so that learners do not feel they have to remember how to do things (eg: where to find lists of words, how to call up the dictionary, how to get the next test item) and can focus on the learning strategies they are using.

ii) Do they engage with goals of understanding?

This was interpreted as referring to their comprehension of the individual target words, or to their understanding of the texts from which the words were extracted, or to their awareness of features of lexical structure among the target-items. The

comments cited here (Table 4.5) are taken from work in the selection module, where both text and words are displayed together. Comments show that attention can vary from a focus on specific words in isolation (S2,S8), to an awareness of the role they play in the meaning of the text (S1,S4), to a judgement about the comprehensibility of the text itself, and its role in giving meaning to the words (S4,S5,S7).

Table 4.5: Comments on Understanding

Subject	Comment
S1	"..I'm happy with my understanding, as well as getting new words.." "..my understanding is a lot more....I think my level's higher..." "..understanding the words makes the sentences follow on.."
S2	"..I just want to know what they (the words selected) mean.."
S4	"..I don't know the text...I need to read the whole thing.."
S5	"..now we see it (the meaning of a word) again we can put the text together more.."
S7	"..I understand most of the text.."
S8	"..this text is a lot easier.." "..I don't know what they (the words) mean.."

This can be taken as evidence that the pursuit of understanding at text and word level is a fundamental process in the interaction, and that this contributes to the richness of it as a learning experience. Awareness of lexical structure, however, as far as it is represented in the amount of grouping that subjects did, is not so well served. Table 4.6 compares total numbers of items selected, notes and groups, and shows that only S4 focused on grouping in preference to writing notes.

Table 4.6: Comparison of numbers of Items, Notes and Groups

Subject	S1	S2	S3	S4	S5	S6	S7	S8
Total Items	37	46	19	28	18	17	13	22
Notes	37	13	15	2	6	6	7	5
Groups	7	7	4	6	3	3	3	3

This suggests that understanding, at the lexicon-building stage of the interaction, tends to emphasise the connection between an individual word and its meaning, rather than the features it shares with other items. (Composition of the groups is examined in more detail in Table 4.15 below).

iii) Are they able to regulate their learning?

Regulation involves planning and performance. Planning mainly applies at the selection stage, where decisions are made about how words will later be processed. Three of the subjects (S2,S4,S8) referred explicitly to this (Table 4.7). This suggests that the potential for strategic regulation of learning processes exists in the design, even if not all learners make use of it .

Table 4.7: Regulation of Learning - Planning

Subject	Comment
S1	"..I won't look in the dictionary now..it's quicker to select and then go on to the next module to look up and make notes.."
S2	"..I want to make a group of preterites.." "..I want to test myself on endings.."
S4	"..I'm going to make groups.."
S8	"..I'm getting some meanings to bump up the groups.." "..this word I already know...it will fit with a group.." "..I want to see how the groups fit together.."

As far as other types of processing are concerned, some of the subjects made comments suggesting that there are aspects of the interaction which are unsatisfactory as far as their successful control of learning processes is concerned (Table 4.8).

Table 4.8: Regulation of Learning - Dissatisfaction

Subject	Comment
S1	"..the text is possibly too long...there's no sense of achievement...it's a big task ahead.."
S2	"..it would be useful if I'd been more organised yesterday.."
S3	"..it's easy to think you've learned something...but when you test yourself you can't remember it.."
S4	"..the trouble is I haven't studied them.."
S5	"..texts should be chosen to provide at least 2 examples of each word.."
S6	"..I 'm put off by not being given the answer...if you've tried 3 or 4 times it's useful to find out...then you're likely to remember"
S7	"..I could easily fit the word into the sentence if I could see the list..."

Most of these comments occurred at the testing stage when they were confronted with their inability to remember some of the words they had studied. S1 and S5 felt that the originating texts should be constructed to help them in some way

(length and lexical content); S2,S3 and S4 were critical of the ineffectiveness of their own processing; S6 and S7 felt that the program failed to give them certain important assistance (the correct answer after a failed attempt, and the list of target words to select the answer from). These comments are probably informed by the subjects' experience of conventional methods and materials for practice and testing, (eg: graded texts, glossaries, question-and-answer testing, discrete item testing, multiple choice testing), but they also relate to dissatisfaction with their actual performance in retrieving target-items (Table 4.9).

Table 4.9: Regulation of Learning - Selection and Retrieval

Subject	S1	S2	S3	S4	S5	S6	S7	S8
Total items Selected	37	46	19	28	18	18	13	22
Items correctly retrieved	24	9	13	8	6	3	2	3

The discrepancy between the number of items selected and those retrieved indicates that processes of regulation regarding the goal of retrieving new items were not properly under control. It is interesting to note that all the problems about text choice, organisation and testing of words, prompting and feedback could have been dealt with by the learner iterating round the stages of the interaction as designed. What is apparently required, in order to do this, is a more explicit understanding of the means and purposes, ie: the strategy, that the program embodies. This was taken into account in the design of the second version of the program (see section 4 below).

iv) Does Processing involve Knowledge Integration?

Issues of knowledge integration are found in a variety of processes relating to the way that individual target words are characterised, the kinds of conceptual frameworks that they are related to, and the extension of focus to other information present in the environment. Some comments on how the assessment of a new target word is done by the subject, are shown in Table 4.10, demonstrating the use of an understanding of context to guess at the meanings or role of an unfamiliar word

(S1,S3,S8), the relation of newly discovered meaning to that previously-known (S4,S5), the use of knowledge about potential variation in distribution (S6) or word form (S7,S8) to guide research into new meaning.

Table 4.10: Comments on Knowledge Integration

Subject	Comment
S1	"..I'm not sure of this...it's a key word"
S3	"..they're words I don't know..though I can guess them from context.."
	"..I may be able to get it in context..but I know I wouldn't remember it in another context."
S4	"..it has a different meaning..which I want to study.."
S5	"..we've just had this in class..I'm interested in its real meaning.."
S6	"they're quite difficult..not words that are used frequently.."
	"..it's so unusual...it will remind me..."
S7	"..I need to look it up..how can I tell what tense it is..?"
S8	"..I'd like more contexts..to see what it means.."
	"..is it capable of putting in different word forms...?"

Frameworks of organisation of knowledge are central to the integration of a word. L2-L1 equivalence was the dominant framework because subjects needed definitions for the unknown words, and the only dictionary available was bilingual. This meant that nearly all the notes were written in L1 (there were occasional exceptions where L2 canonical forms were used for reference - see Table 4.13 below). Groups, on the other hand, are L2 by definition, although the titles given to them may be in either language (in fact only S5 gave L2 titles to any of her groups, and they were grammatical groupings - see Table 4.15 below).

Other knowledge frameworks, such as semantic fields etc. were referred to during the processes of selection and lexicon-building. Subjects related the words they were working with to a variety of linguistic frameworks (Table 4.11), such as semantic fields (S1,S7,S8), word categories,(S2,S3), sense relations (S5) cognates (S1,S5), orthography (S4), frequency (S6), register (S6) etc.

Table 4.11: References to Linguistic Knowledge Frameworks

Subject	Reference	Knowledge Framework
S1	"..be nice to have a thesaurus..see how many other words mean 'feria'.."	semantic fields
	"..if I can put a similar (L1 translation) word, it's easier.."	cognates

S2	"..these are little nouns..I need a better word..something more specific.."	categorial relations
S3	"..I don't know what to call them..they're participles, but not all used like that.."	categorial relations
S4	"..will it show the words in alphabetical order? ..That would be good for learning.."	orthographic relations
S5	"..synonyms are useful if the word is a cognate.."	sense relations/cognates
S6	"they're quite difficult..not words that are used frequently.."	frequency relations
	"..I assume this is informal, not literal.."	register
S7	"..I want to make a semantic group.."	semantic fields
S8	"..I like to group under meanings.."	semantic fields

This evidence shows that all subjects were engaged, at one time or another, in relating what they were doing to what they already knew, either about the target words or about the way that language can be organised. This suggests a link between the subject's ability to organise their knowledge productively, and their awareness of linguistic structure. This point is taken up with greater emphasis in the discussion of the 2nd pilot study, in the next section, and in the design of the main test programme which follows.

Extension of focus, from the target word to other information that surrounds it, can be seen in the way learners assess the value to them of information provided by the KWIC concordancer (Table 4.12).

Table 4.12: Comments on the Concordancer

Subject	Comment
S1	"..I'd use it when my level was higher.."
S2	"..there's loads of them..!"
	"..can I mark the words in the KWIC..?"
	"..I wanted to do something with 'volver' - with 'en.. de.. re.. etc..."
S3	"..the contexts help if they're from texts I know.. there's not enough of the sentence otherwise.."
S4	"..I'm looking at each context to reapply 'puesta'..what does suppuestamente mean?..'dispuesta' is 'disposed to'..etc"
S7	"..it's only half a sentence..can I move it along..?"

These comments relate to the multiple, truncated contexts that the concordancer provides for the keywords that the learners specify. For most of them it was a novel way of looking at the use of words, and the information contained within was not immediately apparent. The first reaction was often surprise (S2,S3), followed by a wish to see more of the context (S3,S7). However, some of the learners did look

for ways to integrate the KWIC information, eg: S2 noticed that his keyword 'vuelto' occurred with a variety of prefixes ('en..' 'de..' etc.) and wanted an easy way to transfer the relevant contexts to his notes; S4 had the same experience with 'puesta'. Even S1 (the elementary learner) observed that the information would be useful to him if his level was higher, ie: if he understood more of the language in the contexts themselves. This evidence suggests that knowledge integration processes are supported to the extent that the learners are familiar with the formats in which new information is presented. They are able to relate to the texts, for example, and to the dictionary, but the concordancer is more opaque. Displaying the keyword-in-context occurrences of target-items and requiring them to be retrieved in response to the clues provided in the testing module is, however, the nearest that the interaction comes to generating new situations for the target words, and thus to providing opportunities for extension of the knowledge gained. It is clear that the design should be more explicit about what the KWIC contexts can be used for and the way that the information can be integrated into the general approach.

Whether linguistic considerations are evident in their notes and groups can be judged by looking at the kind of notes that subjects recorded (based on their use of the bilingual dictionary and concordancer), at the composition of the groups they created, and at the comments they made which refer to linguistic concepts. Table 4.13 shows a selection of some of their notes, together with an assessment which demonstrates that subjects vary considerably in their linguistic awareness.

Table 4.13: Notes on Selected Items

Subject	Target Word	Subject's Note	Linguistic Assessment
S1	'teme'	'afraid/to fear'	Correct sense, incorrect form and syntax
S2	'cierto'	'use lo cierto: certain/sure'	Word related to expression, correct meaning
S3	'descansaba'	'from descansar'	Imperfect form related to infinitive
S4	'puesta'	'pp of poner/dressed'	Participle related to infinitive form, appropriate meaning for context
S5	'apisonamiento'	'steam roller'	L1 association, incorrect but related meaning

S6	'comportamiento al volante'	'flying compartment'	Incorrect word-for-word translation

S1, for example, has recorded the sense of the Spanish word 'teme' without being accurate in the translation ('teme' means 'he/she/it fears'; 'afraid' is 'tener miedo', literally 'to have fear') and without making any reference to its syntax (is it a verb or an adjective?). S2, in contrast, has combined an awareness of a possible collocational role for his target word, with an understanding of its meaning. S5 has created an English association for her target word (literally 'flattening') which is relevant only by association ('steam roller' is 'apisonadora'). S6 has attempted a word-by-word rendering of 'comportamiento al volante ('driving behaviour'), which does not exist as an expression in the dictionary, and produced an entirely incorrect translation ('flying compartment') which makes little sense in the context.

All the subjects, however, regardless of level of linguistic sophistication, made comments,. recorded during the observation, which demonstrate a continuing exercise of linguistic judgement. Table 4.14 shows a range of these comments. S1, S3, S4 and S7 refer to endings etc; S2 and S4 to affixation and root forms; S4, S7 and S8 refer in a general sense to what is Spanish and what isn't, S8 assesses the level of difficulty of a text.

Table 4.14: Linguistic Judgements

Subject	Topic Word	Comment
S1	'enriquecer' 'conquistadores' 'interrumpe'	"..it only means enriched when it's reflexive.." "..this is plural.." "..it's the 3rd person.."
S2	'se divisera	"..it's the only reflexive one.."
S3	'atestada/padecido' 'cometido'	"..I know these are participles...I'll have to look up the infinitives.." "..it has a verb ending, but it's a noun.."
S4	'campesino' *'empanares'	"..it has an adjective ending.." "..it's a bit dubious, it doesn't look Spanish.."
S5	'padecido'	"...except that we might get the wrong synonym, eg: 'suffrir' or 'supportar'..."
S6	'soit-distant'	"..that's French.."
S7	'esfuerzo' '*esfuerzar'	"..it can be pluralised.." "..it doesn't seem to make sense.."
S8	Text	"..there's an amazing difference in the level of difficulty.."

This reflects well on the capacity of the interaction to generate meta-level processing, engaging the learners in a variety of forms of assessment of features and characteristics of lexical items and their context.

The groups defined by the learners also demonstrate a variety of types and levels of linguistic knowledge, as well as something about the learner's awareness of the way this knowledge can be deployed to help in the goal of learning and retrieving the target words. Each group created can be classified and ranked in terms of the amount of linguistic information it embodies, and its consequent (according to the psycholinguistic model) function as an aid to, or reflection of, the learning/ retrieving process. The most functional groups are those which specify small numbers of words with a specific lexical relation. The least functional are the largest, or the most general, such as top-level grammatical categories.

A. Most Functional

1. (Semantic): sense relations, ie: synonyms, antonyms etc.
2. (Semantic): semantic fields, ie: words related by core meaning or semantic composition.
3. (Semantic): lexical sets, ie: words which are 'topic-related'.
4. (Grammatical): sub-categories, eg: 'ER verbs', 'ado/ido endings' etc.
5. (Spelling/Pronunciation): cognate words, ie: words which have the same form in the L1 or other language.

B. Least Functional

6. (Semantic): discourse-related words, ie: words which are related only by their discourse function in the particular text from which they were selected.
7. (Semantic/functional): 'case' words, eg: words which are related by the role they play, as agents, attributes etc.
8. (Grammatical): words classed by general category, eg: 'nouns', 'verbs', 'plurals'.
9. (Grammatical): underspecified sub-categories (eg: S2' s group 'Little Nouns').
10. (Register): eg: idioms, slang, high/low frequency words.
11. (Reflective): words which, on reflection, have some special significance for the learner.

On the basis of this classification, we can compare the relative effectiveness of each subject's grouping, listing both the type and title of each group created:

Table 4.15: Subjects' Group Types

A. Most Functional

Subject/ GpType	S1	S2	S3	S4	S5	S6	S7	S8
Sense Groups	Start Country							
Semantic fields							Negative	
Lexical Sets	Sea Value Invaders		Smoking	Medical Nouns Country vocab	Emotion	Casual		Emotion Fire
Grammatical Sub-Cat.		ER verbs past. parts Preterites Imperfectos		ado/ido verbs	Reflexivos		ER verbs AR verbs	
Cognate	Cognate French							
Items in groups	20	17	2	9	5	3	7	9

B. Least Functional

Dis-course Specific						Religion Journalism		
Case Groups		Personal words						Personal characteristics
Grammatical General.			Nouns Verbs	Nouns	Plurales			
General unspec.		Little nouns						
Register		Idiom/ unusual expression						
Reflective			Words-i-can't remember	Known				
Items in groups	0	11	11	11	6	7	0	1

93

Clear differences can be seen here between subjects whose grouping was predominantly semantic (S1,S8) and those whose groupings were mainly formal (S2, S7). Also between those whose groups are functional in terms of the psycholinguistic model (S1,S4,S7), and those whose are not (S3,S6). Other interesting individual features which occur in this data can be found in the groupings of S2,S4, and S6. S2, for example, said that many of the preterite words he selected and grouped, he in fact already knew and selected them precisely because he could put them into a group. This is an example of the kind of strategic thinking regarding knowledge frameworks which the design is attempting to encourage. S4's semantic grouping shows the influence of the source text - she put 'brazo' (arm) in the 'medical' group together with 'infermera' (nurse) and 'las heridas' (injuries), because, in the text, it was an arm that had been injured. Her grammatical grouping was also complicated by orthographic form - 'las heridas' being put with 'ado-ido ending' words (past participles which might also include 'ada' as an ending). These confusions show that the rationale behind particular groupings may not be very consistent, and although a framework has been employed it is by no means certain what effect it will have, in terms of the associations created. A similar problem is illustrated by S6 - a more advanced reader of Spanish whose groupings tend to relate to the particular discourse in which words are encountered. Her group 'Religion' contains words derived from a text about the Catholic church in Cuba, but includes 'apisonamiento' (the crushing - used to report the suppression of the clergy), and 'baldosas' (flagstones - used metaphorically), neither of which would have much association with religion outside of this text. Similarly, her group 'Journalism', taken from a text describing a room full of reporters waiting for an important announcement, includes 'emparedados' (sandwiches), and 'aludido' (the person referred to). Although these associations may arise from an awareness of the role of the words in the structure of the discourse (and thus indicate a good understanding of the text itself), it is hard to see them as evidence of an understanding of lexical structure, and therefore as contributing usefully to the overall learning strategy.

Overall, the investigation into knowledge integration suggests that the interaction supports the deployment of a considerable range of knowledge frameworks and thus facilitates the integration of new knowledge with what the learner already knows. The design as it stands, however, does not assist learners in finding the best way to represent what they have discovered, either in terms of the notes that they write or the groups they invent. It only offers them the basic idea, and the facility to record whatever they come up with. As the grouping activity is central to both theory and strategy it is obvious that more direction is needed in the form of encouragement to reflect on the type of group which is being considered, and on the purpose of this group vis-à-vis the learning aims. This issue is taken up in the design of the second version of the program.

2.3 Evidence for Spreading Activation

Performance and comments in the testing module provide the most significant data suggesting the operation of a mental network, as Subjects tried to recall their target words. Table 4.16 details some aspect of each Subject's retrieval performance. Each example has been classified using Dell's (op.cit, p.285) categories of English speech error, showing that these errors conform to some of the types predicted by the assumption that multiple sources of activation are at work in the attempt to recall word forms, as the underlying theory proposes.

Table 4.16: Retrieval Performance

Subject	Output	Target	Error Types	Subject's Comment
S1	*'descrimetu' *'descrimuto' *'descrimutos' *'descrimenutos'	'descubrimiento' ('discovery')	phoneme errors /u/ shift /en/ shift /c/ & /r/ blend etc.	"..I'm confused between 'discovery' and 'described' ..." "descritos' is blanking out the other.."
S2	*'llamuras' *'llamudas' 'llamaras' *'llamuros'	'llanuras' ('plains')	phoneme errors /m/ /n/ substitution /d/ /r/ substitution	"..I'm guessing now.."
S3	*'arregran' 'arreglan' *'arresgieron' *'arr..'	'arriesgan' ('they risk')	phoneme errors /i/ deletion /i/ shift	"..it's similar to gar or something.."

S4	*'incideros' *'ratitos'	'colillas' ('cigarette butts')	word errors non-contextual substitution	"..cigarette ends .. it's something like this .." "..that means complete .."
S5	*'pertubador'	'perturbado'	morpheme errors /dor/ /ado/ substitution	"..there's no 2nd 'r' .. " "..I remember writing it wrong before.."
S6	*'apistonamente' *'apistonamiento'	'apisonamiento'	phoneme errors /t/ addition	(Comment at Notes writing time) "..steam roller is not right but it is so unusual it will remind me.."
S7	'elaborar'	'desprende'	word error non-contextual substitution	"..I just don't know.."
S8	*'alludidor' *'aludor' *'alludor' *'aluidor' *'alud'	'aludido'	phoneme errors /i/ shift morpheme errors /dor/ /ido/ substitution	

S1's attempts to produce 'descubrimiento' (discovery) reveal activation of another target word 'descritos' (described). This is confirmed by the subject's comment. S2's attempts to produce 'llanuras' (plains) are interfered with by stronger activation possibly emanating from 'llamar' (to call), which is a frequent word learned early and much reinforced. His experiments with different gender endings (..as, ..os), and with substituting ..d.. for ..r.. are well-informed guessing strategies, as his comment shows, arising from a knowledge of Spanish morphology and pronunciation. S3's confusion between the target word 'arriesgar' (to risk) and the common word 'arreglar' (to arrange) is compounded by activation of the preterite form of ..IR/..ER verbs in the 3rd person plural: '..ieron'. This inappropriate guess, (her comments shows that she is aware that it is an ..AR.. verb), may be the result of the salience of the letters '..ie..' in the original target word. S4's knowledge of the English meaning of the word she seeks, probably provokes activation, in her first guess, of associated common L2 words like 'incendio' (fire) and 'cenicero' (ashtray). The second guess is inexplicable as an attempt to retrieve 'colillas', but her comment shows she associates it with the meaning 'complete' (in fact there is no connection), and that concept may have been a response to associations generated by the word 'restos' (the remains of something) which also occurred in the context. Such productions are consistent with the theory if it is assumed that L1 and L2 lexicons overlap (see chapter 2). S5 has attempted to produce a common

adjectival form ('...dor'), and this has perhaps combined with activation caused by the role of the 'r' in the spelling of the word (see her comments). It is perhaps surprising that the form she produced should have won out over 'perturb...' because of strong activation presumably coming from the English cognate. However, as her comment shows, this production was subject to a high level of conscious monitoring, which may well have discriminated against the too-familiar form. S6's notes to this word used the L1 word 'steam roller' which, although it was not a translation, was supposed to act as a clue (see her comment). What she produced at retrieval time was perhaps affected by activation from the cognate word 'piston' with its connotations and phonological association with 'steam'. S7 produced a Spanish-English cognate word which is semantically possible in this context - but her comment shows that she is aware that it is not the target word. Strong activation from the concept and from the English lexicon are operating to fill the gap left here by the absence of a Spanish synonym. All S8's attempts are interfered with by the conviction that it is a '..dor' noun form that he is looking for. His target list also contains 'amenazador' (threatening) - selected just before 'aludido' - and 'abrumador' (overwhelming) - both adjectival/adverbial forms of the same suffix. It is possible that activation of this morphological feature is greater for him than the more common '..ido' participle form, because it has been more recently learned.

All the subjects thus exhibit performance in the testing module which is explicable in terms of activation from competing lexical items. In addition, there were 5 cases in which some form of general activation of the L2 lexicon seemed to be at work, resulting in the correct retrieval of a target-item following unsuccessful attempts and with no apparent additional clue. For example, S2's target word 'dedos' (fingers), was passed on, but 3 words later, after successfully retrieving 'emparedados' (sandwiches) he remembered 'dedos' and went back and correctly retrieved it. The closeness of many of the attempts and the apparently random variation of letters in subsequent guesses (eg: S1,S2,S3,S5, S6, S8 above) suggests that the design could incorporate some means of helping the learners to make more principled amendments to their guesses, perhaps based on some indication of which

parts of the response are correct. This point was also addressed in the design of the second version.

2.4 Learning Rate

A target learning rate was estimated, on the basis of the literature (see Chapter Two) at 8 words per hour. Because one cycle of processing, with this system, involves the selection, elaboration and retrieval of a target-item, the learning rate can be taken as the time required to retrieve an item for the first time. It can therefore be calculated by dividing the number of items correctly retrieved at least once, into the total time spent in interaction with the program. Although a target rate was not made explicit to the Subjects in these tests, it is interesting to see how their learning rates compare (Table 4.17).

Table 4.17: Learning Rate

Subject	S1	S2	S3	S4	S5	S6	S7	S8
Items correctly retrieved	24	9	13	8	6	3	2	3
Total Time (minutes)	385	151	144	144	114	98	86	70
Words per hour	3.7	3.6	5.5	3.3	3.1	1.8	1.4	2.6

None of the subjects in fact came near to 8 words per hour, although probably for different reasons. S1 perhaps due to his lower level. S5,S6,S7,S8 perhaps because they had 2 or fewer sessions and an overall lack of time in which to process the selected words. These results are not taken as evidence, at this stage, for the design's capability to support processing which meets the target rate, because the design is not finalised. But they do suggest that the complexity of the interaction and iterative nature of the strategy require subjects to work for longer before meaningful results can be obtained. Bearing in mind the qualitative evidence of useful processing which we have seen in the foregoing analysis, it is also clear that the precise relation between the performance-defined learning outcome (ie: items selected and retrieved and the learning rate) and the approach-defined outcome (ie: the richness of the interaction as a learning experience) needs to be investigated

more closely. This is the objective of the second pilot study, following the re-design.

2.5 Summary of Pilot Study 1 and Design Improvements

The evidence from this study can be summarised to give the following assessment of the success of the design in meeting the requirement to provide a rich learning environment:

- The general objectives and procedures of the interaction in the selection, lexicon-building, and testing of target vocabulary are clear, but its underlying complexity means there is a requirement for a more transparent interface with procedures more clearly defined and labelled.

- The interaction represents a rich learning environment as far as the pursuit of understanding is concerned, but learners may not make full use of the potential for strategic regulation of learning processes that exists in the design. They need support, therefore, in progressing to a point where they address the more complex strategic processes, involving awareness of lexical structure, as well as local performance goals.

- The design supports the learner in organising their knowledge productively, and in the continuing exercise of linguistic judgement, but the emphasis in the interaction falls too easily onto processing associated with the finding of translation equivalents. It does not embody a sufficiently explicit description of its purposes and means regarding the building of a network of L2 associations.

The following improvements to the design are therefore proposed:

i) Directions of transpositions of target-items: from text to word-list and vice versa, and from word-list to group etc. should be clearly visually marked. Learner actions

which are related to the assignation of items to groups should be clearly distinguished from actions which record features.

ii) The selection of a target word should involve an initial categorisation of it in terms of broad lexical structural categories. This initial categorisation should be flexible and able to be changed, but it should serve as a 'first level' of grouping, having an organisational role in the grouping module and an informational one at the testing stage. Learners will thus be required to consider the structural features of the items they select, from the beginning.

iii) The use of translation notes as clues for retrieval should be restricted once a target word has been judged by the learner as known. Learners will thus be encouraged to generate group-type associations in order to be sure of getting adequate clues when revising these words. A judgement of 'known-ness' can be forced by limiting the total number of 'not-known' words that can be in the target list at any one time.

iv) The clues and feedback at the testing stage should be appropriate to activating a variety of different forms of partial knowledge of a target word, and to helping the learner make more principled amendments to their guesses.

In addition it has been seen that the interaction provides some evidence for the psycholinguistic model (ie: the operation of spreading activation within a mental network) in learners' responses at the testing stage. However, this performance does not discriminate between the different types of processing from which activation can result, and does not take into account the effect of the learner's approach on the building of their mental lexicon. Thus there are significant qualitative outcomes to the interaction which remain unrepresented in the quantitative model maintained by the system. Two refinements are required if evaluation on the basis of this model is to be productive, firstly the model of the learner's network must represent all the relevant processing that contributes to performance, and secondly it has to be related to the learning outcome assessed

independently of the quantitative interaction data from which it is built. As was concluded in 2.4 (above), evaluation needs to take into account the learner's approach to the task, and to relate this to the learning outcome as represented by performance measures. The learning model outlined in Chapter Three provides, in the 'deep/surface' distinction, a way to evaluate the approach independently of performance, allowing the two outcomes to be compared. The learner model, however, is itself a representation of performance data and so cannot be similarly compared. What is needed, therefore, is a third means of assessing outcome, which is also qualitatively based, and against which the learner model can be validated. It is proposed that this other means of evaluation should be based on the learner's own assessment of the state of their knowledge.

The final system resulting from the implementation of these improvements is described in detail in the next section. This is followed by a second pilot study, focusing on a single subject, which will consider the scope that the resulting interaction offers for qualitative evaluation of the learning outcome on the basis of a 'deep/surface' analysis of the learner's approach, and also on the feasibility of using the learner's self-assessment to validate the learner model generated from the performance data.

Section Three: Improved Design

3.1 Modules

The general architecture remains as shown in Fig 4.1 Enhancements to the component modules are as follows:

a) Selection module

The source text is displayed in a scrollable window. Learners select target vocabulary, ie: those words or phrases they are uncertain of, by highlighting the items they want.

Fig 4.2: Selection Module
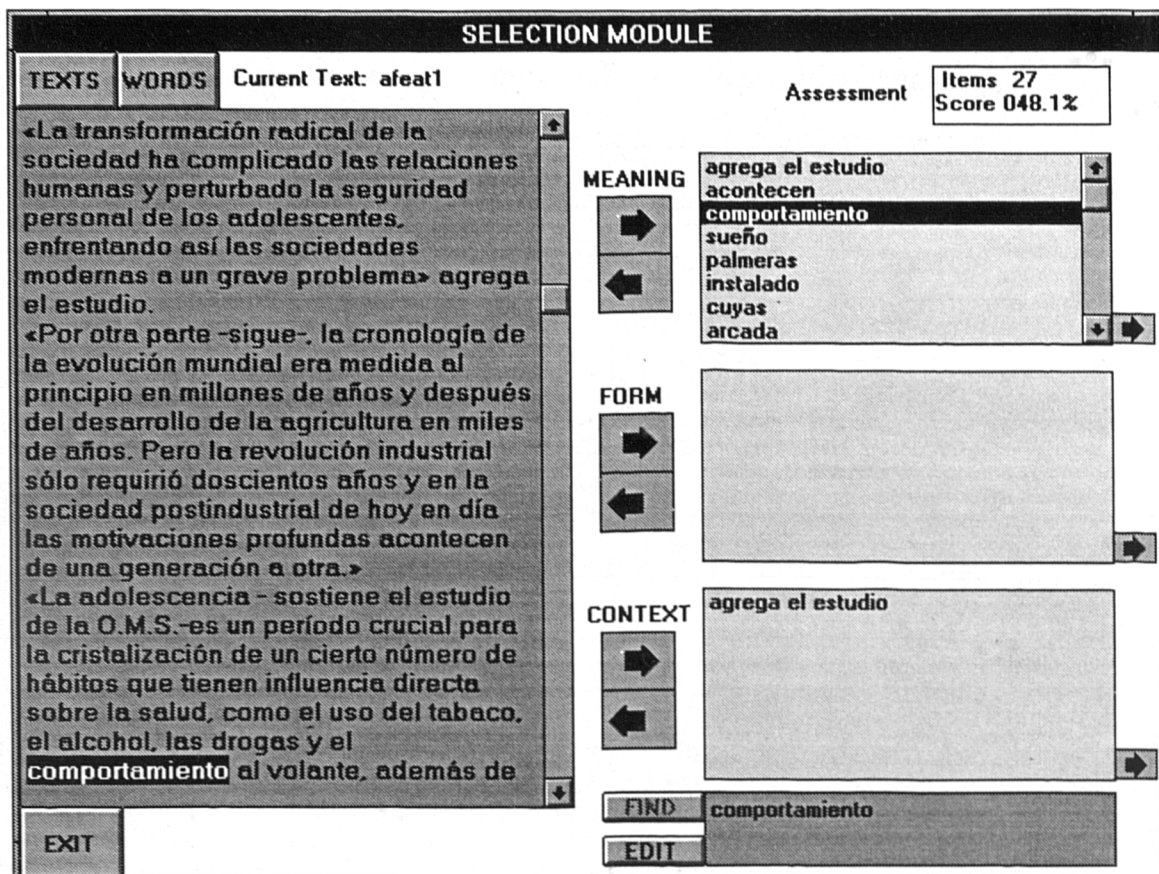
They select items and enter them, on the basis of a subjective judgement, into one or all of 3 category-boxes entitled: MEANING (for items where it is the meaning which is of primary interest), FORM (for items with some interesting grammatical, pronunciation or spelling feature), and CONTEXT (for items with some specifiable relation to surrounding words, eg: collocations). An element of structuring of the target-word list is thus present from the start, provided that the learner is consistent in their criteria for matching items with categories.

In the example (Fig 4.2), the learner has chosen the text discussed in Chapter Three. He has highlighted 'comportamiento' and assigned it (using the large rightward arrow) to the category MEANING, in this case because he is simply interested in the meaning of the word. If he thought further, he might also assign it to the FORM category on the basis of its ending ('..iento') which is a feature of some Spanish nominalised verb-forms. He has earlier selected the expression 'agrega el estudio' (the study adds) and assigned it to the CONTEXT category because it plays a discourse organising role in this particular text (a report on a study of the problems

of adolescents). Clicking the leftward arrows re-locates an item in the text so that it may be seen again in its context and perhaps re-assigned to another category. Clicking one of the small rightward arrows removes the highlighted item from the corresponding category box. The category boxes contain a maximum of 25 items, so that in order for the target list to grow larger than that it is necessary to remove some. When an item is removed the learner is asked if they want to 'finish with' it. This is assumed to mean that they know it. A known item can only be prompted, at the later testing stage, by group clues and not by notes, translations etc. This is to encourage learners to prioritise grouping above other types of elaboration.

c) Lexicon-building module

Once items have been selected, they are expected to be processed in their categories. This is represented by the visual arrangement in the lexicon-building module (Fig 4.3).

Fig 4.3: Lexicon-Building Module

Items within a box can be processed downwards, ie: into groups or notes, but not horizontally, ie: across the categories. If a learner wishes to group two items from different categories they must go back to the selection module and reassign the items to the same category box, which may, if the total target-item set is more than 25 items, involve 'finishing with' some items in that category. The categorisation process is thus integrated into the overall strategy. The learner can create new groups within each category, assign target-items to the groups by clicking on the large downward arrow, and also assign non-target-items (eg: items from outside the source texts, such as well-known words which may provide additional strong associations).

In the example, the learner has selected an item, ie: 'agrega el estudio'. He has placed it in a group called 'phrases', which he earlier created for the item 'de hoy en día' (this name does not in fact reflect a grouping according to meaning, but one according to context - this learner's performance is discussed later in section 4). He has also used the dictionary to get a translation of 'agrega el estudio' (..the study adds..) and written this into the 'notes' box. He may also have highlighted part or all of the current item and referred it to the concordancer to find out what other constructions it might appear in (Fig 4.4).

Fig 4.4: Concordance and Expansion

In this example, the learner has referred the word 'estudio' to the concordancer. The citations show all its occurrences in the entire corpus, together with the originating text title. Selecting a citation and clicking the 'expand' button shows the whole of that originating text.

c) Testing module

The system sequences all the target-items according to the information in the learner model (see d below). The learner can browse this sequence, forward and backwards, using the 'Current item' arrows. A prompt is given for the current test item, which consists of the truncated (60-character) context for that item in the originating text, with the item itself gapped out. The learner has to recall the missing item and enter it. Other available clues are demonstrated in the example (Fig 4.5):

Fig 4.5: Testing Module

105
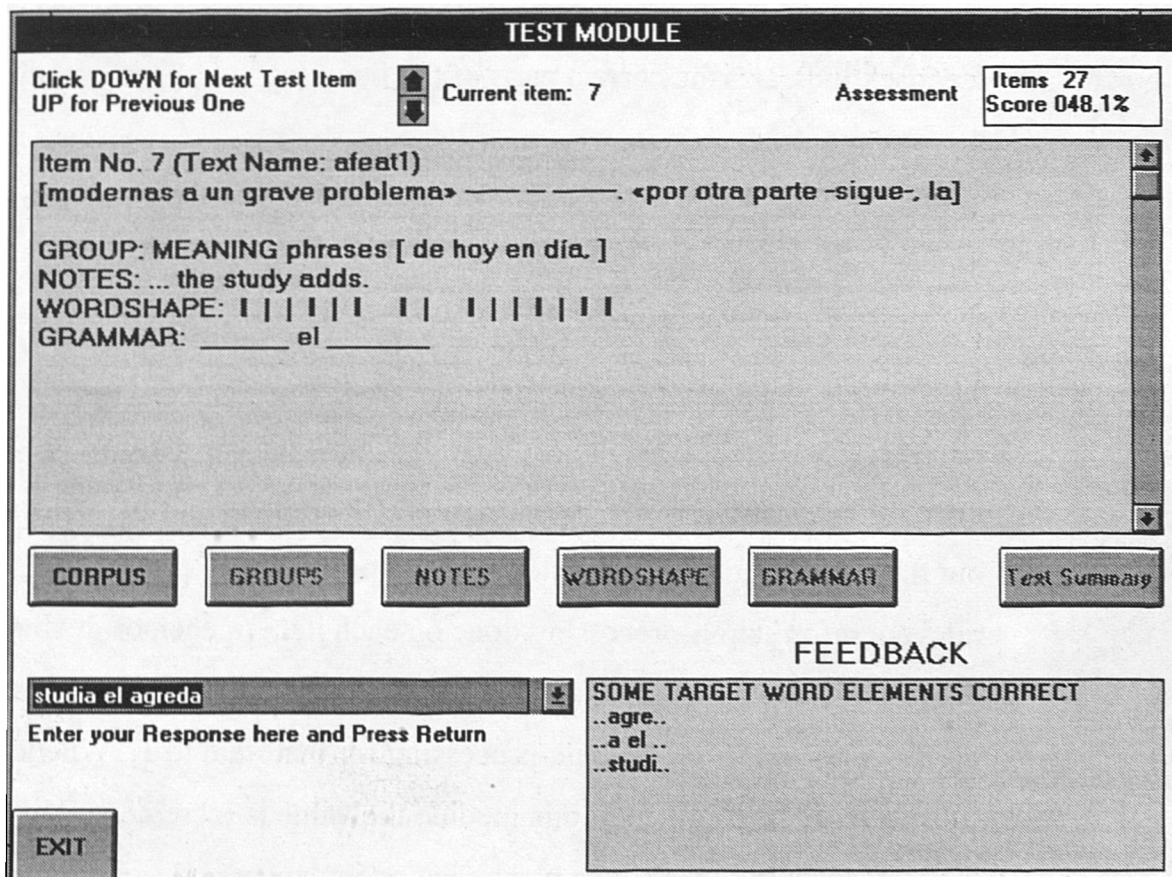
i) Corpus clue: gapped KWIC citations for all the occurrences of the target-item in the rest of the corpus (in this case there were none).

ii) Groups clue: groups that the target-item has been included in (the program displays the category and the group title plus all the other items which belong to the group).

iii) Notes clue: notes that the learner has written for the target-item.

iv) Wordshape clue: vowels and consonants in the target-item.

v) Grammar clue: affixes, endings etc. in the target-item where appropriate.

vi) Summary: any summary of the originating text that the learner has written.

Only the first two clues are available for items that are 'finished with'.
The program gives feedback on the learner's response, identifying correct strings of letters in the answer. In the example the learner has entered *studio el agreda. The feedback indicates the correct parts of the input (agre.., ..a, el, ..studio) in the order they are required.

d) Learner Model

The learner model is abstracted from the general performance trace, which records all the learners' contributions (actions executed via keyboard or mouse) to the interaction. There are 39 possible actions (including entering and leaving modules), but the learner model is restricted to a 'snapshot' of the types and amounts of relevant information-processing done on each item in each of the three modules. When the learner first selects an item in the selection module the learner model sets the value for selection-module-processing for that item to 1. Whenever that item is subsequently highlighted in this module the value is incremented by 1. Similarly in the lexicon-building module, using the concordancer and saving notes each result in

the incrementing of the item's score for that type of processing. In addition, when an item is assigned to a group, the code for that group is added to the item's record. The learner model thus records how many and which groups each item belongs to. In the testing module, the learner's attempts to recall items also result in changes to a value for testing-module-processing for the item, incrementing it if the attempt is successful or decrementing it if the attempt is an incorrect one on a previously successful item. A representation is thus maintained, measured by the number and type of events involving a particular item in each module, of the presumed level of activation of that item in each of the general categories of processing.

This representation enables items to be compared in terms of their relative levels of assumed activation, according to the cognitive model which underlies the testing strategy (see Chapter Three, 1.4). This comparison is used to sequence the presentation of items for testing, so as to maximise the effects of activation spreading from highly-processed items downwards. The order is determined according to the rationale that extensive processing in the first or second stages of the interaction results in high levels of both association and activation for the items concerned. This renders those items more likely a) to prime subsequent items, and b) to be retrieved themselves. Items with a high level of processing at lexicon-building stage only, are likely to have associations which can be triggered. Items with high levels of retrieval are least in need of the retrieval-practice effect. This rationale is implemented in the following testing sequence: items combining a high level of processing at selection or lexicon-building stages with a low level of retrieval are practised first, items with high levels of processing at lexicon-building stage alone are practised next, and items with high levels of retrieval are practised last. This strategy is further amplified in the next chapter.

3.2 Section Summary

The system thus described has been enhanced in several important ways, reflecting an increased emphasis on the functions of word-grouping and the capacity of the program to record the events which make up the interaction with a learner, and to

compile them into a representation of the mental lexicon which the learner is assumed to be creating. The features of the system which were assessed in the first pilot study as able to provide a rich learning environment and engage learners in the exercise of their linguistic knowledge have been retained, but there is added emphasis on the promotion by the program of the overall strategy, and the use of the learner model to support the assumed processes of association and activation. The new features of the design have created new evaluation requirements, focusing on the learner's approach, the validity of the learner model, and the qualitative assessment of learning outcome. The capability of the interaction to provide data for this level of assessment is examined in the next section.

Section Four: Pilot Study 2

This study concentrates on a single subject - S1 from the first study. This subject was chosen because he was the lowest-level learner to work with the program, and should thus provide a benchmark for establishing a base level at which the interaction can reasonably be expected to be fruitful. Whereas in the first study he worked from a text which was in deliberately simplified Spanish, in this study the text is an unmodified one (again the 'adolescents' text from Chapter Three) which could be expected to pose considerably more reading problems for him. The consequent interaction should therefore test the support that the program can give him to its limits. The objectives of the study are to assess the extent to which the learner's approach, defined in terms of deep and surface learning, is evident in the way he interacts with the program, and also to determine whether the quantitatively-based learner model can be meaningfully compared with the learner's own assessment of the state of his knowledge. The data examined is of the same kind as that used in the first study, with the exception that more detailed information is available from the program's trace of the interaction (the performance trace), and from the subject's introspective comments on what he was doing. Also, an additional source of data is available in the procedure for self-assessment, which is described below.

4.1. Procedure

The program used was as described in the section above. The Subject was a 25-year-old English learner of Spanish in London (subject S1 from Pilot study 1). His level was just above elementary, as he had not studied Spanish formally and had just been 'picking it up' for about 12 months in short visits to Spain and in occasional classes and conversations with Spanish-speaking friends in London. For this study he set himself the task of using the program to learn 50 new Spanish words in 2 weeks. There were 5 observed sessions during this period, consisting of an introductory session, 3 observed practice sessions, and a recorded interview. He was permitted to use any other on- or off-line resources as he saw fit. The program recorded all his actions with the mouse and keyboard (the performance trace) and built a quantitative model of his lexical knowledge in the form of a set of integer values attached to each selected item, representing the amount of processing of different kinds that he had done on that item (the learner model). His comments and replies to questions while he was working were noted by the observer. Questions usually took the form of "why did you do that?" referring to some action or other, or "what are you thinking about now?" when he lapsed into silence. In a final interview he was asked to assess his success at learning his target words. The assessment took the form of a series of judgements about each of the L2 words in his target list. Three criteria for assessment were proposed to him, corresponding to the main kinds of processing assumed by the theory of the interaction: i) whether he could remember any of the other words in the context from which he originally selected the item, ii) whether he understood the meaning of the item, iii) whether he thought he could produce the item in a new context of his own. He was asked to consider each item in the list and mark it yes, no, or maybe, for each criterion. He was asked to think aloud whilst doing so and his comments were recorded.

4.2 Evaluation of Results - Performance Trace

During the five hours eight minutes of the Subject's interaction with the program (spread over two weeks), the performance trace recorded 552 mouse or keyboard events distributed across the three program modules (Table 4.18).

Table 4.18: Overall Use of the Program

Module	No. of Events	Time (mins)
selection	65 (11.5%)	78 (25%)
lexicon-building	330 (60%)	180 (59%)
testing	157 (28.5%)	50 (16%)

The distribution of time across the three modules is considerably more skewed (in favour of lexicon-building) than any of the results from study 1. In addition, the totals for events in each module, when compared with the times, show that selection processes took significantly longer, which is consistent with the increased difficulty of the text. There are two possible explanations for the difference between the two studies, in the Subject's performance at the lexicon-building stage: one is that the way he managed the interaction was affected by the new design features or by the increased difficulty of the language content, the other is that it was affected by a change in his approach. The results considered below should suggest which is the case. Whether his strategy was different from that intended by the design, and whatever kind of approach he adopted, the performance outcome represented by the total number of words selected, the number successfully retrieved, and the number of groups created (Table 4.19), indicates that it had far-from-optimal consequences:

Table 4.19: Performance Outcome

No. of Items selected	21
No. of Items correctly retrieved	13
No. of Groups	1 (2 items)

This falls well short of his 50-word target. The learning rate (23.7 minutes of study for every 1 successful retrieval, or 2.5 words per hour) is also considerably slower than either his own figure (3.7 wph), or the average (3.4 wph) for study 1. The small number of items selected (21) and the number of those retrieved (13) reflect the shorter time he spent in the selection and testing modules (25% and 16% respectively), but the single group he created in no way accords with the time he

spent (59%) or the extent of his activity (60% of events) in the lexicon-building module. The following analysis of the performance record reveals the extent to which his strategy departed from the optimal one:

i) Processing in the selection module.

The new program design supports the assignment of selected items to one or all of the categories: MEANING, FORM, CONTEXT, and subsequent re-assignment at any point in the interaction. Re-assignment is necessary if the user wants to group together words which have initially been put into different categories. Use of it indicates reflection on the lexical categories and a refinement in the approach to grouping. Of the 65 selection module events in this Subject's performance trace, only 7 were related to re-assignment. The trace also shows that his use of categories consisted of single assignments to the MEANING and CONTEXT boxes, mostly to the former, with only 4 assignments made to FORM. There does not, therefore, appear to be much consideration of structural aspects of the selected items. The ratio of selections to time in this module (1 every 3.7 minutes), on the other hand, suggests that there was fairly extensive mental processing of some kind going on, but we can see that it was not particularly related to the establishing of categories for the selected items. The conclusion is that his selection of items was slowed down by his attempts to understand the text itself. The Subject's comments during selection processing confirm this:

"..I was trying to hang on to the whole context...to focus on the whole thing..."

As we saw in the discussion of deep and surface approaches in Chapter Three, this might be evidence of a deep approach to understanding the text, but it does not guarantee a deep approach to understanding the linguistic structure.

ii) Processing in the lexicon-building module.

The program design intends the creation of groups to be the principal processing activity in this module. Dictionary and concordance lookup, and the making of notes on meaning or translation of individual items are supposed to assist in the discovery of lexical features which can then be recorded as group names and have relevant items assigned to them. Of the 330 events in the performance record, only 11 related to the creation of groups and assignment of items to them (1 group created, containing 2 items, put together on the comparatively non-functional basis that they were expressions and not single words). The trace also reflects the fact that all his notes were L1 translations. But again, the trace shows that some other kind of processing was going on, which was not related either to grouping or information lookup. In session 3, for example, in 29 out of the 38 events where he clicked on an item, the highlighting of the item was not followed by either a grouping event, dictionary or concordance event, or a notes-saving event. In session 5, none of the 29 clicks on items had any follow-up. In session 6, only 1 of 27 events had any follow-up. An explanation of what he was doing, is provided by his comments:

> "..Once I've selected the words I want, I go into the lexicon-building module. I click on each of the words and write it down on a piece of paper. Then I go into the dictionary and look up each of the words in turn. I write down next to it the main translations. Then I go back and go through clicking them again. I do a concordance on each one and also look it up in the big (off-line) dictionary. I work out which of the meanings is right and enter that into the notes box.."

His strategy was thus to identify and memorise 'correct' English translations for the target-items. Where the program did not support this, eg: dictionary and concordancer were inadequate or could not be used simultaneously, he resorted to off-line solutions such as noting words to look up in a book dictionary, printing out the text in hard copy etc. He was operating, in effect, with a mental list of L1-L2 pairs and not with a network of L2-L2 associations, a strategy tending to isolate each item from the others and, in effect, replace it with an English translation. If we relate this approach back to his activity in the selection module, we may reasonably infer that he was trying to translate the text word-by-word, rather than investigate

the contextual lexical properties of his selected items, effectively a surface approach.

iii) Processing in the testing module.

The program design intends the original contexts of target-items, plus associations rising from the way they have been grouped, to be the primary means of prompting production in the practice module. The Subject's overall reluctance to practice his target words, which is also consistent with his objective of translating the source text rather than increasing his vocabulary, means that the performance trace data does not give an adequate picture of his processing in this module. Of the 157 events, for example, only 4 were related to the use of the groups clue, an obvious consequence of his lack of attention to grouping at the lexicon-building stage. With the total number of target-items barely exceeding 25, and his restricted use of testing, the consequence was that the embedded strategy for encouraging grouping (the procedure for 'finishing with' items - see previous section) was irrelevant.

It is clear, therefore, that this subject's strategy does not conform to that assumed by the design. The evidence is also that the translation strategy he did adopt did not produce the kind of performance outcome which is consistent with the vocabulary-learning aims of the interaction. Qualitative data from the interaction is nonetheless required to support an interpretation of this approach in terms of deep and surface learning, and the extent to which it does so will be examined below. In addition to the interpretation, we will also look for some correspondence in the qualitative description of learning processes and outcomes, to the performance outcome, even though the strategy from which this arose was not in accordance with the program's pedagogical intentions.

4.3 Evaluation of Results - Qualitative Analysis

The surface-deep dichotomy can be used to examine the record of learner actions for those which are not related to a structure-perceiving approach, and those which

are. Surface-level actions are those which are intended to create links between target-items and their L1 translations, deep-level actions are those which pay attention to associations (contextual or decontextual) among the target-items themselves. Not all learner actions can be thus classified by sight, some require the Subject's introspective comments to elucidate his intentions, others are outside the pedagogical framework of the interaction. For the examples shown in Table 4.20, the comments are derived from the Subject's remarks during the self-assessment session and from protocols recorded during the observed sessions. These examples are illustrative, a full classification of the entire performance record would cover many more action types, including the grouping and note-writing performance, in which it would take into account qualitative criteria such as that discussed in the first pilot study. Although there is a certain amount of ambiguity in some of the interpretations, it is clear that such a classification can indicate the extent to which the learner's performance conforms to a surface or deep profile, thus relating it to quality in the learning process. In this Subject's case, the examples of deep processing identified are almost the only ones that occur in the whole interaction, so the quality of the overall learning process, as represented by the data, can be characterised as surface level and therefore likely to yield a poor outcome. This, as we have seen, would be in accordance with the quantitative evaluation of the learning outcome. Nevertheless, the existence of even a slight indication of deep approaches to the task is encouraging in the case of this Subject, whose level, as was pointed out in the introduction to this section, represents a kind of base-line for the system.

Table 4.20: Classification of Learner Actions by Approach Type

Action Type	Surface Approach	Deep Approach
Select target-item from text (selection module)	Choice of item is arbitrary, eg: Subject's comment "..I chose 'esfuerzos' because it looked nice.."	Choice determined by role in the text or by meaning or form related to other items selected. eg: "..I chose 'perturbado' because ....I wanted to know what else the text had to say about human relations.."
Assign or re-assign target-item to category boxes (Selection module)	Assignment of item is arbitrary, eg: Subject 's comment "..I just put everything into the CONTEXT box.."	Assignment of item is considered, eg: he re-assigned 'aporta' from MEANING to FORM because he noticed that it was a near-cognate: "..I wanted to change the box because it means 'bring' as in 'port'.."

| Create group (Lexicon-building module) | Group title does not reflect significant lexical feature, eg: his creation of group 'expressions' including 'agrega el estudio' and 'de hoy en dia' | Group title reflects significant lexical feature

(no examples) |
|---|---|---|
| Write notes on item (Lexicon-building module) | L1 translation copied verbatim from dictionary: eg: Subject's notes for 'aporta'..show up, arrive, port..' were irrelevant to the contextual meaning 'bring' | Select appropriate L1 sense or modify grammar to fit form of target-item or write L2 paraphrase for selected sense and form of target-item. eg: his notes for 'fracasado:'..'have been unsuccessful..' reflected its form and contextual meaning rather than a literal translation |
| Refer to KWIC (Lexicon-building module) | Do the concordance on the current form of the target-item. eg: he did concordances on word forms as they appeared in the target word list and not on their root forms, eg: 'acontecen' (3rd pers. pl.) produces 1 citation - 'acontec' would have produced 3 | Do the concordance on a root form for the target-item

(no examples) |
| Select clue type (Testing module) | Use translation clue eg: he used translation to prompt retrieval of 'favorecer', 'enfrentando', 'esfuerzos' | Use groups clue, eg: Subject used groups clue (despite inappropriateness of group title) to prompt retrieval of 'agrega el estudio' and 'de hoy en dia' |

A surface/deep analysis can also be applied to the data which is generated in the self-assessment exercise which was introduced for the purpose of comparison with the learner model (see 4.4 below). It has already been suggested that this Subject's approach was based on the perception of links between target-items and their L1 translations, resulting in the creation of an L2-L1 'list' rather than an L2 network. His comments on his knowledge of the meanings of some of his words, elicited in the self-assessment interview at the end of the test programme, support an interpretation which confirms the surface nature of this approach. In the absence of in-context processing or retrieval practice, some of the L1-L2 pairs appear to have become decoupled (Table 4.21).

Table 4.21: Comments during self-assessment

Word	Subject's comments as he tried to assess his knowledge of it
'acontecen' (they happen)	"...I've got a few meanings in my head...it could be another one.."
"...it could have meant 'confront' but I think 'enfrentando' means that..."
"...I was looking at 'comportamiento'.."
comportamiento ('behaviour')	"...I think that meant 'the behaviour..."
rodeado ('surrounded)	"...what's come to me is the English translation of a word I used which was 'surround'..."
"...I don't know if 'rodeado' means 'surround' or if that is another word...
...or 'acontecen'.."
enfrentando (confronting)	"...I think that means 'to confront'

115

(re: Eng. word 'surround')	"...'surround' is either 'acontecen' or 'comportamiento'..one means 'behaviour'..." "...if I was to guess I'd say 'acontecen' means 'surround'..."

This surface approach to the knowledge which is involved in knowing the meaning of a word could be contrasted with one which emphasised either the contextual role of the word, eg: "..comportamiento was used in the description of how young people drive...", or its co-text, eg: "..comportamiento al volante...", or its structure, eg: "...the verb comportar-to behave...in noun form..." etc. Such deeper approaches would indicate that the learning processes which contributed to this recall were of a qualitatively higher order than those which simply serve to produce an L1 translation, and could also be expected to result in a better performance outcome. These examples demonstrate the capability of the interaction to generate data which supports qualitative analysis in terms of deep and surface learning processes. This is true even where the learner is at a baseline level of target-language competence, and where their overall strategy tends to confound the pedagogical intentions of the program design. In the final sub-section below I will look at the status of the learner model in comparison to the same qualitative assessment.

4.4 Evaluation of the Learner Model

The learner model represents the interaction viewed in terms of the theoretical processes of activation in the mental lexicon which are assumed to underlie the learner's performance. It is, correspondingly, used to influence interaction in the testing module, which is where implicit cognitive processes (eg: retrieval) are thought to be at their most active. The model is based on the general performance trace and thus represents the outcome of whatever learning strategy the learner has adopted. How good this representation is, in terms of its ability to stand for what the learner has learned and thus to be effective in the application of a testing strategy, is dependent on the actual quality of the learning processes which have given rise to it. Where the learner has adopted the strategy-as-designed, together with a deep approach to the task, it could be assumed that the record in the learner model, at least as far as it accurately embodies the theory, would quite closely

represent the learning outcome. Where the strategy is an alternative one, however, or where the learning approach has significant surface characteristics, the model's record could not be expected to reflect the outcome. The extent of the reflection, therefore, might be taken as an indication of the closeness of the learner's strategy to the optimal one, or of the general quality of their learning processes. In this study, the learner model was compared with the outcome as represented by the learner's self-assessment. Self-assessment was carried out, at the end of the test programme, by getting the learner to judge each item (yes/no) according to the criteria: a) Can you remember any of the other words in the context from which you originally selected this item? b) Can you now understand the meaning of this item? c) Could you produce this item in a new context of your own? The learner model was made to generate an equivalent set of assessments on the basis that high processing levels in the selection module are equivalent to a 'yes' judgement for criterion (a), (and low ones to a 'no' judgement), that high levels in the lexicon-building module are equivalent to 'yes' for criterion (b), and that high levels in the testing module equate to 'yes' for criterion (c). A comparison was then made by overlaying the grid generated by the model onto that generated by the learner. Table 4.22 shows the result. The Subject shows that he has considerably more confidence in his knowledge than his performance data would appear to warrant, especially in his assessment of his ability to use the words in another context of his own. This may be a consequence of his predominantly surface approach, failing to appreciate the complexity of the requirement to use a word. In terms of the comparison with the assessments generated by the model, we can see that there are 33 cells in the grid (50%) in which there is agreement. Most of the positive correspondence is in the meaning category because it was in the lexicon-building module that the Subject did the most (and the deepest) processing - involving looking up meanings, using the concordancer etc. Where the model overestimates most is in the context category, because whilst the Subject did a certain amount of processing in the selection module, most of it was of a superficial kind (ie: the attempt to translate the text into L1). Where the Subject overestimates most is in the use category, because a) his surface approach encourages him to identify this judgement with the 'meaning' one and not to see the added complexity involved in production for use,

and b) the lack of processing in the testing module means the model is unable to assess the majority of items in this category.

Table 4.22: Comparison of Learner Model with Subject's assessment

S=Words known according to the Subject
M=Words known according to the model

Item	Selection/ Context	Lexicon- building/ Meaning	Testing/ Use
esfuerzos	S M	S M	S M
profundas	M	S M	S
palmeras	S M	S M	S
favorecer		S M	S
desprende	S M	S	
perturbado	S	S M	S M
fracasado	S M	S M	S M
de hoy en día	S M	S M	S M
acontecen	M	M	
comportamiento	S M	S M	S
motivaciones	S M	S M	S
sueño	S	S M	S
ventanas	S	S M	S
rodeado	M	M	
arcada	S M	S	
estrecha	M		
enlozado	S	S M	S
húmedo		S M	S
juzgar		S	S
agrega el estudio	S M	S M	S M
aporta	S M	S M	S M
enfrentando	M	S M	S M

Thus, there is a discrepancy in the overall pattern of correspondence, which reflects the fact that the learning processes diverged from the pedagogical intentions of the program, although it does not indicate whether these divergent processes are more superficial than the expected ones. However, there appear to be patterns of 'overestimation' within the 3 categories, which may constitute further evidence of depth or superficiality of the learning approach. These patterns may be made more evident by a re-design of the structure of the model so that it reflects a gradation in the assumed strength of activation of items, rather than a simple 'yes' or 'no' for each category of processing. This issue is taken up in Section 4.5 below.

The interaction data recorded by the learner model, therefore, does in some way represent the Subject's knowledge state. This is despite the fact that it does not record optimal learning processes as defined by the theory and intended by the

pedagogical design. In terms of the explicit objective of maximising successful retrieval of selected words, the learning outcome is poor, and the same is true for the implicit objective of learning about lexical structure. The subjective data also confirms that the optimal network-creating strategy supported by the pedagogical design was not taken up. Judged by the overall aims of the design, therefore, the quality of the outcome is low, demonstrating Pask's principle that if the style of the teaching does not match the approach of the learner then the outcome is poor (Pask 1976). But the learner has not been idle, and his approach has produced some learning and a self-assessment which is confident, at least as far as the 'meaning' category is concerned. This mismatch may be represented in the nature of the relation between the learner model and the self-assessment data, and further explicated, via the 'deep/surface' conceptual framework, in a comparison of what the learner thinks about the task (evidenced by introspective data), with what they do (performance data), and with the learning outcome (qualitative evaluation on the basis of the pedagogical strategy).

4.5 Re-design of the Learner Model

The model can be restructured to provide a more tractable representation of the learner's network of items, presenting it as the result of 3 main types of processing:

i) General processing events (initial selection of items, subsequent location of those items in their source texts, KWIC lookups, notes-saving events).

ii) Grouping events (assignment of items to groups).

iii) Retrieval events (correct retrievals, incorrect retrievals).

Items can be related to each other according to these 3 measures in the following way. First the highest score of any item, for each of the 3 measures, is found. This is used to establish 3 ranges, low, medium and high, for each measure. Every other item is then allocated to one of these ranges, according to its score. If, for example,

item 10 had the highest score for general processing at 6, then that would fix the top range for general processing at 5-6. Items with a score of 1-2 would go into the low range, and items with a score of 3-4 would go into the medium range. When each item has been allocated to the appropriate range for each of general processing, grouping and retrieval, a sequence of items is derived as follows:

i) Items in the low range for retrieval, and the high range for grouping.

ii) Items in the low range for retrieval, and the high range for general processing.

iii) Items in the low range for retrieval, and the medium range for grouping and/or general processing.

iv) Items in the low range for retrieval, and the low range for grouping and/or general processing.

v) Items in the medium range for retrieval and the medium-high ranges for grouping and general processing.

vi) Items in the medium range for retrieval and the low range for grouping and general processing.

vii) Items in the high range for retrieval.

This is then the sequence in which items are presented for testing. The first items to be presented come first because they are assumed to need productive practice (their retrieval values are low), but they are also expected to help activate the learner's network, by virtue of the grouping or the extent of the general processing which has been done on them. These are followed by the items which need both practice and possible further processing (all their values are low). Items which the learner has had most success with in the past (medium and high retrieval values) are left till last. As a result of the learner's success (or otherwise) in retrieving items during the test, the values for retrieval are incremented (or decremented). On subsequent occasions when the learner enters the testing module, these new values (and any others generated by interim processing in the selection and grouping modules) are used to calculate the sequence anew.

The patterns of assumed activation (low/medium/high) in the final network can thus be directly compared with the learner's self-assessment (no/maybe/yes). If patterns of over- and under-estimation, such as those described above, are found to be consistent with the adoption by learners of deep or surface approaches, eg: the model tends to over-estimate where the learner's approach has been a surface one, then this may suggest that the model is indeed a useful representation of the learner's knowledge, and that its structure can serve as a basis for a pedagogical strategy, and also contribute to an overall assessment of the value of the interaction.

Section Five: Conclusions and Further Testing

In the vocabulary-learning program described in this chapter, the main pedagogical objective is explicit in the interaction design, in the sense that the facilities for separating words from context, for testing them etc. clearly promote the learning of vocabulary. But there is a second objective, implicit in the design, which is to help learners develop an awareness of lexical structure, in the sense that those who make sensible decisions about lexical categories, word-groupings etc. get more help from the program. The optimal strategy for vocabulary learning encouraged by the program is based on the principle that mnemonic associations generated between lexical items in the target language will serve a) to make those items easier to retrieve, and b) to build an understanding of the kinds of general and language-specific relations that can exist between words in the L2. The program encourages lexical categorisation from the moment that new words are first selected, supports the creation and expansion of common-feature groups, and promotes the use of these groups as cues in recall practice of the constituent words. A learner who consciously adopts this strategy, and who deliberately organises his/her selections and processing around the need to find new classifications and groupings for existing items, is thought to be approaching the learning task in a 'better' way than one who concentrates on decoding the selected items in their contexts and expressing the resulting understanding in L1.

The two pilot studies demonstrate that the program design is capable of supporting a rich learning environment for the user and can assist them in the exercise and development of their linguistic knowledge. Furthermore, they suggest that the mental processing in which the system engages learners may be in accordance with that which is postulated in theories of lexical processing. Learners, however, can impose their own pedagogic intentions on any interaction, and the second study demonstrates that where these intentions confound those of the program, then processing is likely not to be optimal. Data shows that a poor learning outcome, described both qualitatively and quantitatively, can be related to the learner's failure to adopt (or the program's failure to promote) the learning strategy which is appropriate to the implicit pedagogical objective. This conclusion, deduced from the perspective of a conceptual framework which describes the learner's approach in terms of deep and surface learning processes, is suggested by both performance and introspective data. Several questions can now be identified preparatory to the design of a test programme intended to validate the final design:

i) Is the quantitative learning aim realistic for this system? None of the subjects in the pilot studies achieved it, but they were either working with the earlier version of the design (PS1) or starting from a low level of linguistic competence (PS2). The test programme should be able to establish whether, and under what conditions, 8 words per hour of study is feasible.

ii) What are the parameters of the qualitative learning aim as defined by the system? Developing an awareness of lexical structure, without explicit pedagogical support, (ie: instruction in aspects of lexical structure), is dependent on word-grouping using information from dictionaries etc. PS1 showed that learners create groups of many types and of varying degrees of linguistic specificity. The test programme should be able to demonstrate the results of 'good' and 'bad' grouping strategies.

iii) Does the 'deep/surface' dichotomy for approaches to vocabulary learning stand up? This fundamental question could only be answered conclusively if it were

possible to investigate all conceivable approaches. However, it may be answered in principle if the following can be established: are approaches identified as 'deep' associated with 'good' learning outcomes? Are 'surface' approaches associated with 'poor' learning outcomes?

These questions involve a number of subsidiary considerations relating to the reliability of the objective data and the interpretation of subjective responses. These will be discussed further in the next chapter.

CHAPTER FIVE: EVALUATION - MAIN PROGRAMME OF TESTS, RESULTS AND DISCUSSION

Introduction:

The strategy for computer-based autonomous vocabulary-learning which has been developed in this thesis defines vocabulary-learning as a process involving the selection of target items from context, followed by meaning- and form-related association of these items out of context, and cued productive retrieval items from memory. The pilot tests described in Chapter Four demonstrated that interaction with a system in which this strategy is implemented provides a rich learning experience, and results in a learning outcome which can be evaluated for target content and learner approach. It was also shown that this interaction generates performance features which reflect assumed underlying psycholinguistic processes. The following questions were identified as the focus of a programme of tests intended to evaluate the interaction for a range of learners:

i) Is a quantitative target learning-rate of 8 words per hour feasible?

ii) Does the interaction support the development of deep learning approaches?

iii) Does the quality of the learners' approach relate to their performance outcome?

iv) Does the system's learner model reflect the learner's approach and the learning outcome?

In this chapter a programme of tests is described, focusing on learners of English and Spanish. Results are analysed with regard to these questions and to the implications they have for optimising the design of the system. Section 1 outlines the research methodology and describes the subject groups and the different conditions under which they were studied. Section 2 addresses the feasibility of the target learning rate. Quantitative performance data is analysed and interpreted to indicate the effect of different learning strategies on the learning rate for individual

subjects. Section 3 describes subjects' use of the grouping strategy and evaluates its significance for performance and the quality of the learning approach. Tendencies in subjects' performance are interpreted as indications of their potential to exploit the grouping strategy in the way intended by the design. Section 4 considers further evidence for deep/surface characterisation of the subjects' approach, and the relation between this and the learning outcome as indicated by the performance measures used in the preceding two sections. Section 5 discusses the extent to which the system's learner model reflects the learner's approach and the learning outcome. Section 6 conducts a general discussion of these results, focusing on the differences and similarities between the subjects, and drawing some conclusions about the effectiveness of the strategy and the design of the system in which it is implemented.

Section 1: Procedure and Subject Group

A broadly similar procedure to that used in the pilot tests was adopted (see Chapter Four). Subjects were given an introductory lesson in order to learn how to use the program, then observed over a series of self-directed sessions. Their comments during these sessions were written down, and all keyboard events were recorded by the system. Finally, they were interviewed away from the computer and asked to reflect on their knowledge of the target words they had selected. There were 12 subjects divided into 2 groups. The first group (S9-S14) consisted of 6 learners of Spanish as a Foreign Language (native language English), in the first year of an undergraduate modern languages course at a London university. The second group (S15-S20) were 6 learners of English as a Second Language (native languages Spanish and Portuguese), who had all lived and worked in London for a minimum of 2 years, and who were following a training course in basic computer skills and ESL. All the subjects spent 4 sessions with the program (after the introductory lesson), amounting to between 2 and 5 hours study each, spread over a period of 2 to 5 weeks. The results discussed below are therefore based on approximately 50 hours of program use. The final interviews were carried out between 1 and 2 weeks after the last on-line session. There were some differences in the procedure for the

two groups, due to the different circumstances under which the tests had to be carried out. For the English-learning subjects (hereafter referred to as the ESL group), the test programme was integrated into classwork. For the Spanish-learning subjects (hereafter referred to as the SFL group) it was a voluntary extra-curricular activity. The major difference in the nature of the learning experience for the two groups, was that the SFL subjects were observed individually and thus had a higher level of interaction with the supervisor than the ESL subjects, who were observed in groups of 3. It is possible that this may have resulted in the SFL group becoming more familiar with the program overall (because they were able to ask more questions), and that this might have affected their learning outcome. However, as the language of interaction between subjects and supervisor was English for both groups, the ESL group were anyway at a disadvantage in this respect. For this reason also, the protocols from the SFL group were much more detailed. Nothing could be done to remedy the basic inequality, but some of the inadequacies in the introspective data resulting from the restricted communication of the ESL group, were addressed by means of an additional formal interview on the subject of their approach to vocabulary learning. On-line performance data for all subjects was recorded by the program in exactly the same way (see Appendix IIb).

The taped record of one of the final self-assessment interviews (S11 from the SFL group) was lost due to a defective tape recorder. Otherwise all data is complete.

Both groups used the identical program, the only differences being in the source texts that were used, and in the on-line dictionaries. The ESL group used the monolingual Collins' Electronic English Dictionary and Thesaurus, the SFL group used the bilingual Microtak Spanish Assistant 5, and Collins' On-line Spanish Dictionary. Texts were pre-selected for both groups (see Appendices IB and IC), being judged to be of equivalent levels of difficulty.

Section 2: Learning Rate

In Chapter One the figure of 8 words an hour was proposed as a minimum target rate, based on the size and general nature of the L2 vocabulary learning task as proposed in the literature on vocabulary teaching (see Chapter One). In the pilot studies reported in Chapter Four, learning rate was defined as retrieval rate. The rationale for this was that correctly retrieved items are assumed to have been selected and processed, (translated/annotated and/or grouped), and that the measure reflects a process definition of learning rather than a product or end state. Although correct retrieval of an item on one occasion does not guarantee that it will always be correctly retrieved, (eg: later attempts may be interfered with by the quantity or nature of intervening selections), it was considered that any correct successful cued-production is evidence that the item is at the threshold level of knowledge, and available to be retrieved even if actual retrieval is subject to performance failures arising from the nature of the psycholinguistic processes of word production (ie: spreading activation - see Chapter Three).

This supposition is supported for the majority of subjects in this test programme, by evidence from unsuccessful attempts to re-retrieve items after they had been successfully retrieved once. In the majority of these attempts the target item is clearly identifiable, albeit distorted by spelling errors, (Table 5.1 gives some examples), suggesting that these items are indeed at threshold level.

Table 5.1: Examples of Unsuccessful Re-retrieval Attempts

Subject	Target Item	Retrieval Attempts
S12	colillas desagradado se divisara	petillas/cenillas/corillas desagredado se divieron
S10	padecido emparedados	parecido emparadedas/emparedas
S14	caballeros envuelto emparedados	cabollores/cabellores envuelto emparededos
S20	ploughed puncture	ploughted/plounghed puncted
S15	attaching muddy	ataching muggy/muggs
S19	bumping	punping

However, another kind of error which figured prominently in the re-retrieval attempts of some of the ESL group, casts some doubt on the appropriateness for them of this retrieval-based definition of learning. This error is the substitution of one target item for another. Where it is the case that retrieval of an earlier-selected target item is consistently over-ridden by one or more subsequent ones, it could be concluded that the selection and retrieval processing devoted to the earlier one had failed to establish it at threshold level. In fact, there are only 2 cases in approximately 39 occurrences of this kind of error (S16 & S19) where the intended target word was not eventually re-retrieved, but the implications for assessment of the learning outcome remain. Although for most items (and most learners) it can be assumed that a single completion of the select-analyse-retrieve cycle is enough to ensure threshold knowledge, exceptions may occur where learning conditions have produced the substitution of target items. The possible nature of such conditions will be discussed in Section 4 in the context of learning approaches; for this section I will adhere to the rate of initial correct retrievals as a good enough approximation for a definition of learning rate.

In the pilot studies the retrieval rate achieved by subjects was calculated in words per hour (wph) by dividing the number of items which were correctly retrieved by the total study time (in minutes) and then multiplying the result by 60. None of the subjects in those studies achieved the target rate of 8 wph. An important question to be considered in this test programme is whether the rate is in fact feasible.

2.1 Feasibility of 8 Words per Hour

The Table below (5.2) compares subjects' retrieval rates in the test programme, indicating the total time they spent on the study (together with the number of days it was spread over), the number of items they selected in that time, the number of items correctly retrieved at least once, and the retrieval rate in words per hour. The subjects are ranked from fastest to slowest.

Table 5.2: Retrieval Rates

Subject	Time in mins (spread of days)	Items Selected	Items Correctly Retrieved	Retrieval Rate wph
S12	208 (27)	37	30	8.7
S14	229 (14)	28	28	7.3
S13	250 (28)	25	25	6
S11	258 (31)	35	25	5.8
S10	216 (27)	21	21	5.8
S9	240 (23)	27	23	5.7
S17	287 (35)	23	23	4.8
S20	234 (35)	18	18	4.6
S15	268 (35)	21	17	3.8
S19	120 (14)	11	7	3.5
S16	209 (28)	12	9	2.5
S18	155 (28)	14	4	1.5

One of the subjects (S12) achieved the minimum target rate, demonstrating that it is feasible. One other (S14) came reasonably close. The fact that the rest averaged only 4.6 wph, however, suggests that it is difficult, at least in the context of this kind of autonomous learning strategy. Taking into account that an item must be selected before it can be retrieved, only one other subject (S11) selected enough items to even make the target rate possible within the time they spent.

How difficult was the target rate for the majority who did not achieve it? One way to look at this is to ask how much of an improvement in performance would be evident if these subjects had achieved the target rate. If we take the time that each subject spent as fixed, we can estimate the increased amount of either selection or retrieval, or both, which would have occurred if the target rate had been achieved. From this we can get an idea of how much extra processing each would have had to do in the time. The situation is complicated by the fact that for some subjects all additional items would have had to be both selected and retrieved, whereas for others some of the additional processing would have been the retrieval of already-selected (and presumably processed) items. Nevertheless a broad picture of the extra work each subject would have done is discernible by adding existing selections and retrievals together and expressing the additionally-needed ones as a percentage of the result. Table 5.3 gives the details. We can see from this that to have achieved (and maintained) the target rate would have meant over 30% more

processing for all but 3 of the subjects (S14, S11 & S9), and over 50% more for all the ESL subjects, including over 100% more for the 2 weakest of them (S16 and S18). It is clear, therefore, that the target rate was well beyond achievement for many of these subjects working as they did, and some account is required of why this should have been so.

Table 5.3: Levels of Selection and Retrieval Commensurate with 8 wph

Subject	Existing Selections + Retrievals	Additional Selections for 8 wph	Additional Retrievals for 8 wph	% Additional Processing for 8 wph
S12	67	-	-	-
S14	56	3	3	11
S13	50	8	8	32
S11	60	-	9	15
S10	42	7	7	33
S9	50	5	9	28
S17	46	15	15	65
S20	36	13	13	72
S15	38	15	19	89
S19	18	5	9	78
S16	21	16	19	166
S18	18	7	17	133

2.2 Relation Between Processing Rate and Learning Rate

The obvious question to ask is why was S12 so much more successful, (in terms of retrieval rate at least), than the others? Was it simply that he worked faster or did he also have a more efficient strategy? A comparison of the number of actual processing events initiated by S12 with that of S16 (Table 5.4), for whom the overall study time and spread of sessions was virtually identical, shows that a large variation in general processing rate is possible between individuals, even where the overall shape of their interaction with the system appears to be the same. (An event, as recorded here, refers to a keystroke or mouse click) . S12 generated 700 events (3.4 events per minute) compared to S16's 303 (1.4 events per minute) - more than twice as fast. In addition, although the distribution of events across the 3 program modules is broadly similar, a difference can be seen in the manner of their use of the selection and lexicon-building modules. S12 initiated nearly twice as many sessions in these modules, with a corresponding amount of switching back and forward between them.

Table 5.4: Comparison of S12 & S16 Rate of Processing.

Subject	Time	Spread	Total Processing Events	Selection % events (No. of sessions)	Notes & Grouping % events (No. of sessions)	Testing % events (No. of sessions)
S12	208 mins	27 days	700	13.% (9)	36% (18)	50% (4)
S16	209 mins	28 days	303	10% (4)	41% (11)	48% (4)

This higher level of inter-module activity followed on from the fact that S12 selected 3 times as many items (see Table 5.2), up to the point where he had to start deleting some from the category boxes in order to make room for more (see functions of the program, Chapter Four). This necessitated a number of extra checks back to the notes and groups, resulting in greater complexity, as well as speed, of overall processing, as intended by the system design.

Although this data helps to explain the difference in retrieval rate between 2 subjects at either end of the performance table, the relation between rate of processing and retrieval does not hold generally. Some subjects, ranking high for retrieval, were much lower for rates of processing (eg: S14, S13, S10). For others (S20, S19) the relation was the other way round. Table 5.5 compares rank orders for retrieval and general processing (events per minute), showing that there is not much correlation between them. Features of learning strategy other than speed or quantity of processing, must therefore contribute to differences in the retrieval rate.

Table 5.5: Retrieval vs Quantity and Rate of Events

Subject	Retrieval Rate (Rank Order)	Total Events	Events per minute (Rank Order)
S12	8.7 (1)	700	3.4 (1)
S14	7.3 (2)	398	1.1 (7)
S13	6 (3)	317	1.3 (9)
S11	5.8 (4)	590	2.3 (4)
S10	5.8 (5)	232	1.1 (11)
S9	5.7 (6)	478	2 (5)
S17	4.8 (7)	384	1.3 (9)
S20	4.6 (8)	575	2.5 (3)
S15	3.8 (9)	493	1.8 (6)
S19	3.5 (10)	396	3.3 (2)
S16	2.5 (11)	303	1.4 (8)
S18	1.5 (12)	102	.7 (10)

2.3 Achievement:Target Ratio

Achievement:target ratio is the percentage of the items each subject selected (column 3 in Table 5.2) which were subsequently correctly retrieved (column 4). It is evident that many subjects did not achieve 100%. These subjects are not necessarily those with the lower retrieval rates. Both S12 and S11, for example, rank low for their achievement:target ratio but high for retrieval rate. Also, those who did retrieve all their selected words were not necessarily those with the higher retrieval rates, S17 and S20, for example, rank fairly low for retrieval. What, then, is the significance of the achievement:target ratio for an account of differences in rate of learning?

Although there is no direct connection between rate of processing and retrieval rate, there may be one between rate of processing and target achievement. Five of the subjects (S14, S13, S10, S17, S20) managed to retrieve 100% of their target items, and if we look at Table 5.6, which rank orders subjects by processing rate, we find all but one of them (S20) in the bottom half of the table.

Table 5.6: Comparison of Target Achievement with Processing Rates

Subject	% Target Achievement	Events per Minute (in rank order)
S12	81	3.4
S19	65	3.3
S11	71	2.3
S20	100	2.5
S9	85	2
S15	81	1.8
S14	100	1.7
S16	75	1.4
S13	100	1.3
S17	100	1.3
S10	100	1.1
S18	29	.7

A link between success at retrieving target items and lower rates of general processing (ie: fewer events per minute) would be logical if it were due to more effective processing, eg: fewer clues and guesses required per retrieval, regardless of the overall time taken. Table 5.7 shows clearly that 4 of the subjects who had

100% achievement:target ratios, achieved this level with fewer attempts required to achieve successful retrieval.

Table 5.7: Target Achievement Compared to Successful Retrieval Attempts

Subject	% Target Achievement	Total Attempts	Total Successful (%)	Attempts per Success
S17	100	54	43 (79.6)	1.3
S10	100	42	32 (76)	1.3
S13	100	45	32 (71)	1.4
S14	100	71	49 (69)	1.4
S9	85	57	32 (56)	1.8
S12	81	101	56 (55)	1.8
S11	71	79	40 (50)	1.9
S20	100	78	38 (47.4)	2
S16	75	51	22 (43)	2.3
S15	81	86	36 (41.8)	2.4
S19	65	36	13 (36)	2.8
S18	29	19	6 (31.6)	3.1

Here, the four '100% target achievers' identified in table 5.6 occupy the topmost positions for effectiveness in retrieval. The exception, S20, owes his low success rate to a large number of minor spelling errors. If we extend this analysis to the whole process, not just the retrieval part of it, we can show that effective processing, in terms of lower numbers of events per successful item (Table 5.8) generally reflects high target achievement (significant at the .01 level using Spearman's rank order correlation).

Table 5.8: Effective Processing Compared to Achievement:Target Ratio

Subject	Events per Item Retrieved (rank order)	% Target Achievement
S10	11	100
S13	12.7	100
S14	14	100
S17	16.6	100
S9	20.8	85
S12	23	81
S11	23.6	71
S18	25.5	29
S15	29	81
S20	31.9	100
S16	33.6	75
S19	56.5	65

But it is clear that time spent retrieving current items, even if the retrieval is very effective in terms of numbers of events, still has to be balanced against the rate of selection of new items if this effectiveness is to be reflected in the overall learning rate. If learners adopt strategies aimed at maximising their retrieval score (such as not leaving a test session until they have achieved 100%), they are not necessarily contributing to the improvement of their learning rate. Tracking the number of items selected and the score for each test session for some of the subjects shows this. In Table 5.9 (below) the scores are slightly different from the achievement:target ratios discussed above, because they incorporate decrements resulting from the occasional failure to re-retrieve an item.

Table 5.9: Items and Scores at the End of Each Test Session

Subject	Test 1 Items	Score%	Test 2 Items	Score%	Test 3 Items	Score%	Test 4 Items	Score%	Test 5 Items	Score%
S14	6	100	13	100	13	92	28	100		
S13	10	50	25	36	25	100				
S11	8	100	20	70	20	65	35	69		
S10	7	100	12	100	21	100				
S17	4	100	10	100	23	100				
S20	4	50	10	30	18	84	18	89.5	18	100

As we have seen, S14, S10 and S17 were among the best at retrieving their target items, and each completed 3 sessions with the maximum score. However, S14 lost the opportunity to increase selection in session 3, S10's overall rate of selection was slow, and S17 started from a low base (only 4 items selected in session 1). Their success was not, therefore, translated into correspondingly high retrieval rates. S13, S11 and S20 also achieved 100% at some point in the process, S13 doing so at the expense of an increase in selection in session 3, whilst S11 maintained a good rate of selection in sessions 2 & 4 but was unable to match it with equivalent retrieval. S20 only achieved his maximum score at the expense of selection in sessions 3, 4 & 5. For these reasons, none of these subjects optimised their retrieval rate either.

Correct retrieval is thus only one of the learning goals which the strategy supports, and concentrating on it clearly has an effect on other performance measures and on the overall rate of learning, which does not necessarily optimise them.

Nevertheless, some learners will evidently continue to focus on the 100% score as their main goal, regardless.

2.4 Target Orientation and Process Orientation

This account goes some way towards explaining the connection between low rates of general processing and high target achievement. We might characterise it in terms of an identifiable strategy: 'target-orientation', which aims to score 100%. The importance of this strategy as a learner characteristic is that it may hold down the overall learning rate, and result in performance outcomes belying potential for some of the more successful retrievers. It can be contrasted with an alternative 'process-orientation', which aims to maximise selection and thus favours a higher learning rate. Process-orientation, however, does not guarantee learning rate, as some subjects demonstrated. For S11, for example, an increased rate of selection in later sessions coincided with a decline in her ability to retrieve the items. For S9, a high rate of selection in the earlier sessions meant starting with a low level of retrieval and never catching up. Nevertheless, both these subjects ended up with higher learning rates than S17 and S20, who did achieve 100%. In addition, S12's adoption of a process-oriented strategy, combined with the speed of his processing, and whatever features of quality there were in his learning approach, appears to have contributed much to his achieving the target rate.

2.5 Summary of Section 2

Although only one of these subjects achieved the target rate of 8 words per hour, this shows that it is feasible. However, the majority of subjects would have needed increases of between 30% and 100% in the amounts of processing they did, to have managed the target rate in the time that they spent. The performance of the subject who did achieve it was characterised by a high rate of processing and a strategy which I have called process-oriented because it balances rate of selection with target retrieval. Five subjects achieved 100% target retrieval but this appeared to be at the expense of maintaining a selection rate commensurate with achieving the target

retrieval rate. The strategy of maximising target retrieval without regard for selection rate I have called target-orientation.

It may be concluded that the definition of learning rate as retrieval rate is appropriate to the process definition of vocabulary learning previously developed in this thesis. This is because it is able to reflect benefits for the learning outcome of tolerating less than 'complete' knowledge of target items whilst continuing to seek to expand them (re: Twaddell 1974 in Chapter One). The target rate of 8 words per hour may also be considered feasible, subject to the learner adopting an appropriate strategy for achieving it. The quantitative aims of the learning strategy and the design of the system embodying it are therefore validated in principle, although it is clear that the relation between process-orientation and a higher learning rate ought to be made explicit to the user. This could be done by giving extrinsic feedback on the learning-rate as well as on target achievement (as is currently done in the score). Learning rate is not, however, the whole story, and is inadequate to characterise qualitative aspects of learning outcome or learner strategy. Both high and low levels of performance need to be explained in terms of the types of processing the subjects do, as well as the quantity or rate of it. This issue is addressed in the next two sections.

Section 3: Deep Learning Approaches - the Contribution of the Grouping Strategy

At the level of the unconscious processing postulated by the psychological model (see Chapter Three), grouping target words together should make them easier to retrieve, because it creates a network of associations which increases the range of sources from which a target word may receive activation. At the level of the learning strategy on which the program design is based, effective grouping constitutes a deep learning approach and should result in the words being better known, because it involves making explicit the system of abstract relations which determines how they combine. These are, then, the two ways in which grouping as a strategy is expected to contribute to learner performance and learning outcome.

As was pointed out in Chapter Four, the program design does not require or teach the strategy, it only supports and encourages it, by limiting the extent to which alternative strategies can be relied upon as the target word list grows. The critical number of items is 25. Once the target word list is above 25, note- prompts and clues for items which have been 'finished with', have to be abandoned, making the learner dependent on grouping information to prompt retrieval. In this test programme all subjects were encouraged in the induction session to group their target words, but not all of them continued to do so once past the initial sessions. Some possible reasons for this are suggested in the discussion below. As the majority of subjects did not select more than 25 items during the period of the programme, there was no additional incentive for them to group items arising from constraints of the type I have just described.

To the extent that learners did adopt the strategy, this was manifested by time spent in the lexicon-building module creating new groups, assigning items to them and reviewing the contents of them. How much of this they did can be measured in terms of numbers of relevant events, but how such processing subsequently affected retrieval aspects of their performance is not possible to judge from the small amounts of data involved. In the discussion below, this quantitative measure of the processing done is instead compared with a qualitative evaluation of the groups which resulted from it. The value of the activity, for each subject, can then be assessed on the basis of whether the associations created could represent an enhancement of the subject's knowledge, and whether they could constitute a structural basis for further enlargement of the target item set.

In addition to the lexicon-building module processing, there is a wider application of the grouping strategy, available to those learners who adopt the approach, in the selection and retrieval of items. At selection time, it involves principled initial categorisation of the items selected. At retrieval time, it concerns preferential use of the 'groups' prompt to assist recall. Again, because the target list for most learners did not grow to a size where these strategic aspects became critical to their learning objectives, evidence for the effect of such processing on learning outcome is

minimal. However, the fact that some subjects were aware of grouping as a strategic element in the overall task is evident from retrospective comments made during the self-assessment exercise at the end of the programme. For those subjects for whom it is relevant, an assessment will be made of the extent to which the strategy was an issue for them, and the potential for development that their use of it showed.

3.1 Grouping Events in The Lexicon-Building Module

Activities which are counted as events in the lexicon-building module are: clicking on an item to make it the current one, clicking on the KWIC function, clicking on the dictionary function, saving a note, creating a new group, assigning the current item to a group, reviewing the items in a group. The last three of these can be considered as grouping-related events. Table 5.10 shows how many events of the last 3 types were initiated, and the percentage this represents of total activity in the lexicon-building module.

Table 5.10: Grouping-Related Events in The Lexicon-Building Module

Subject	Create-New-Group Events	Assign-Item-to-Group-Events	Review-Items-in-Group Events	Total Grouping-Related Events	Total as % of Module Events
S17	9	16	26	51	37.7
S20	9	22	26	57	20
S10	5	9	3	17	19.5
S18	2	5	0	7	17
S11	9	25	12	46	16.9
S14	6	12	4	22	16.2
S16	6	13	1	20	16.1
S19	5	12	5	22	13.3
S9	6	14	3	23	12.1
S15	4	6	8	18	11.7
S12	6	11	2	19	7.5
S13	2	4	1	7	6.2

This data shows a considerable variation across subjects. It also reflects the fact that all subjects devoted most of their time in the lexicon-building module to processing activities other than grouping. One possible reason for this is that grouping logically depends on knowing something about the target items (eg: their meanings or typical use patterns) and so has to wait until learners have used the

138

information resources and perhaps noted what they have found out. At that point it is then in competition with the possibility of testing the items just analysed, or of selecting new ones, both of which may be more motivating in the early stages of using the program. Grouping in the early stages is therefore something of an afterthought. Nevertheless, a certain amount was done, especially by S17 who integrated the strategy into a quite distinctive approach to the learning task as a whole. This approach, which reflects her target-orientation (see the previous section), will be discussed in more detail in 3.2.2 below.

Also evident from the above data is the very small contribution that grouping-related events made to the performance of S12 & S13, two of the subjects with the highest retrieval rates (see Table 5.2). Grouping is not, therefore, necessary for optimising retrieval rate at this stage in the use of the program. Indeed, if the quality of the grouping is low or indifferent, it could clearly slow down the overall learning rate. The qualitative analysis below should establish whether this may have been the case for any of these subjects, and also cast light on some of the underlying learning strategies.

3.2 Qualitative Analysis of Groups

Chapter Four introduced a framework for the qualitative analysis of groups, which divided them into functional and non-functional, with semantically-defined titles considered as the most likely to support useful associations. A further distinction was made between relations which derive from the contextual role of the items (called 'lexical sets') and those which are structural in the sense that they pertain to semantic primitives within the items concerned ('sense groups' and 'semantic fields'). Although both these types of group were considered equally functional, provided they were specific enough, in the following analysis it will be seen that some lexical sets contain words which have a somewhat tenuous connection out of context (eg: the association that some SFL subjects made between the words for 'combed' and 'fingers'). Such associations do not exemplify significant lexical relations between the words concerned, they only record a particular image. They

139

would not therefore be expected to enhance the learner's knowledge of the words (although they might aid recall of them in the short term).

By far the biggest categories of group, for both sets of subjects in the test programme, were the semantic field and lexical set. Because the subjects in each set were working with the same texts and tended to select many of the same words, certain of the more obvious groupings occurred to several of them. In many cases, there may have been no more than one item in the text appropriate to a particular group title, but all the subjects were shown how to add 'extra' words (ie: words from outside the text) in order to further exemplify the particular feature they wished to mark. In fact very few of the subjects utilised this facility. This will be commented on further in 3.4. Other types of grouping were more speculative, either not adequately specifying a recognisable lexical feature, or else specifying one for only one item, so that there was no association created. Formal groupings are considered to be useful if they reflect specific features or combinations of features of items, such as reflexivity, tense, unusual spelling, but not if they only serve to characterise the items as members of very general groups such as 'perfect' or 'ER' endings etc. The number of items assigned to a group is also considered to be relevant to its usefulness, (the optimum number being between more than 2 and less than 7), and so also is the degree of cross-referencing (items assigned to more than one group). The number of potential associations generated by a group is equal to the number of pairs of items that the group contains. These factors are synthesised into an assessment (below), in which each subject's grouping is described and evaluated for its usefulness either at the local level, ie: for recalling the particular items involved, or at the global level, ie: for learning about lexical structure. A comment is then made on the extent to which the interaction appears to have supported the realisation of a deep learning approach for this subject.

3.2.1 Analysis by Subject

In the analysis below, each subject's groupings will be described, and assessed according to what proportion of processing in the lexicon-building module was

involved, and the likely contribution that the activity made to the learning process. In the summary in 3.2.2 this evaluation will be contrasted with the subject's performance in retrieval rate and achievement:target ratio. It will be seen that usefulness of grouping at the local level can be associated with target achievement for most learners, although the overall quality of the grouping activity is not very high.

S14:

This subject's grouping constituted 16.2% of his lexicon-building module processing. He grouped 11 of his 28 selections, only 1 of them more than once. He created the semantic field titles 'bad feelings' ('le molestaba-it bothered him/padecido-endured') and 'advances' ('se adelantaba-went on/se acercó-came forward'), and the lexical sets 'room' ('el suelo-the floor/el techo-the roof/la salita-the small room'), 'human body' ('los dedos-fingers/corpus-body/*el mano-hand), and 'hair' ('se peinó-combed/cabello revuelto-messy hair'). 'El techo-the roof' in the 'room' group, was an extra non-contextualised item. A similar attempt to include the word 'mano' as an extra item resulted in his giving it an incorrect definite article. (In fact the word exists in the text with its correct - feminine - article, but he failed to notice it). The ramifications of this are not obvious. If the effect of grouping is to make items easier to retrieve, then it might be that this has the undesired result of reinforcing the incorrect form '*el mano'. On the other hand, 'mano' is one of a group of 'o' -ending nouns with feminine articles which typically cause problems for learners, and which the subject will undoubtedly be made aware of during his studies. It is possible that the association of it with other words in this text might thus serve to bring it back to his attention and assist in his re-learning of the correct form. His only formal category was a lexically informative association of the expressions 'desenvoltura en los demás/cabello revuelto/ envuelto emparedados', which was to put them into a group entitled 'volver', relating them both morphologically and to an underlying semantic primitive connected with 'turning'. This is an example of the optimal kind of processing that was envisaged in the program design, as it involved integrating information from both dictionary and concordancer, to produce a structurally

informative and psychologically salient set of associations. In quantitative terms his processing should have been valuable, covering nearly 40% of his items, with 12 potential associations. Qualitatively too, it contains structural aspects which might be expected to enhance his knowledge. In the final self-assessment exercise he made a number of references to the groups, recalling 'salita/el techo', 'padecido/molestaba', 'el suelo/salita', 'los nudos/flojos', 'envuelto emparedados' in a 'grammar' group, and 'los dedos' grouped as 'body parts'. He also remembered having grouped 'desenvoltura' in the 'volver' group, and although he could not remember its meaning he was convinced he would know it if he saw it in context. 'Se acercó' and 'se adelantaba' he also remembered grouping, although he could not remember their meanings. This subject's group-related processing could therefore have been highly functional for his learning outcome, and possibly even more so in the event of him increasing his target word list. His application of the grouping strategy is therefore assessed as <u>globally useful</u>. The interaction is considered to have contributed towards the realisation, for him, of a deep learning approach.

S17:

37.7% of her processing in the lexicon-building module was group-related. 10 of her 23 items were grouped, 4 of them more than once. None of her groups contained more than 4 items, but 3 contained only 1 item. The groups included a semantic field entitled 'feelings' which she began with the word 'startled', which is consistent with the group title, but then went on to add 'aware' (defined in the dictionary as 'having knowledge' and cited twice in the corpus followed by 'that' and 'of') which makes a rather tenuous and syntactically confusing link. She also created the title 'body' but could only find 1 word to put in it ('jaw'). Her other titles were lexical sets: 'Repairing' ('attaching/ puncture'), 'Road' ('bumping/ ploughed'), 'Country' ('ploughed') and 'Field' (no words), and a narrative-based group, 'Accident' ('screeched/ struck/swerved') in which the middle item ('struck') only belongs by virtue of being causally linked to the accident in the narrative ('..he struck a match..'). Her formal categories were restricted to 'ED words' (4 items) and one attempt to register a pronunciation feature, 'OUGH=ow' which contained

only the item 'ploughed'. In quantitative terms her grouping represents functional processing, as it covers nearly half the target list and generates at least 11 potential associations between individual pairs of wards. In qualitative terms, however, the group titles do not indicate much processing of structural relations between her target words. They demonstrate an intention to generate contextual rather than structural links. This may have been potentially useful to her learning outcome with this small number of items, but it is unlikely that such benefits would continue to operate once the number of source texts had been expanded and the target list had grown. Her application of the strategy is therefore assessed as <u>locally useful</u>, and the support given to her by the design in the realisation of a deep learning approach as only partially evident. Some model of a more generalisable kind of semantic relation between words (eg: a decomposition of a set of emotion-related words such as Rudzka's example given in Chapter One) might have gone further towards helping her develop her awareness of lexical structure and enabled her to apply the strategy more effectively.

S20:

20% of his processing was group-related. He grouped 12 of his 18 items, 4 of them more than once. The largest group contained 5 items, whilst 2 groups contained only 1. He created the semantic field 'feelings' beginning with 'startled', and added 'thrilled' (from his 2nd text) which is a genuine member of the set. He then went on to include 'regards' (in the salutary sense) which is not an obvious relation, especially considering its totally different syntactic role. He created another semantic field entitled 'sensation' into which he put 'startled' together with 'nip' (defined in the dictionary as 'to affect with a stinging sensation'), which is also a dubious association because it glosses over the difference between physical and mental experience. Other groups intended to be semantic fields were: 'action' ('ploughed/screeched/ startled/swerved/ bumping'), which contains a mixture of contextual and structural relations, 'fast movement' ('swerved/screeched'), and 'faith' ('race riots'). The attempted refinement of 'feeling' to 'sensation', and 'action' to 'fast movement' suggests that he was trying to think structurally, but his inclusion of 'screeched' in the movement group is a contextual connection. His

formal categories included 'ED' words in which he mistakenly included the word 'puncture' (nb: he subsequently made 4 attempts to retrieve the form 'punctured'), 'ING' ('bumping/attaching'). '*isiomatic expression' ('I'm looking forward to seeing you') and 'form' ('struck'). Quantitatively his group processing is useful, as it involved 66% of his target set and generated 22 potential associations. Qualitatively too, his focus on semantic fields suggests that his knowledge may have been enhanced. The range of his titles demonstrates an intention to find a variety of relations, but, as the last example ('form') shows, he was not always able to specify a significant feature. Nevertheless there is reason to believe that he intended this processing to assist his learning effort, and that it might have been seen to do so if he had gone on to select a larger number of target words. His processing is therefore assessed as <u>locally useful</u>, and the contribution of the interaction towards developing a deep learning approach is only partially evident. Like S17, he would probably have benefited if the system had been able to supply him with examples of semantically-decomposed concepts and the sense-relations between them.

S10:

Her group-related processing amounted to 19.5% of her activity in the lexicon-building module. She grouped 8 of her 21 selections, 1 of them more than once. All her groups contained 2 items except for 1 with a single item. She created the lexical set 'cigarette' ('ceniza/colillas') and the semantic fields 'discontent' ('padecido-endured/se azoró- got alarmed), 'proximity' (se acercó-came forward/se adelantaba-went on), and 'time' ('larga-long'). Her only formal category was a pair of past participles, 'envuelto/padecido'. Quantitatively her grouping could not have had much effect on her performance, as she covered less than half her target items and only generated 4 potential associations. She made only one reference in the self-assessment exercise, to having grouped 'se azoró' with 'padecido'. However, the processing she did involves significant structural similarity, and would be clearly be available to be built upon in the event of her target list increasing in size. Her processing is thus assessed as <u>locally useful</u> and the support provided by the interaction for deep processing as partially evident and likely to become more so as the target-item list increased.

S11:

Group-related processing made up 16.9% of her lexicon-building module total. She grouped 21 of her 35 selections, though only 3 of them more than once. Her groupings are characterised by the combination of items having different (and sometimes fortuitous) types of relation, apparently more influenced by her imaging of what is described than by any structural analysis of the items involved. For example, the group 'combing' ('revuelto-disordered/se peinó-he combed/dedos-fingers') records an image created in one sentence of the text but is not a proper lexical set because the topic relation between 'revuelto' and 'dedos' is incidental. Another example is her group entitled 'hate' ('desagradado-displeased/revuelto-disordered/ padecido-endured/odiaba-hated'), which is an attempt to identify a semantic field but which also incorporates an incidental link with 'revuelto'. Another is the lexical set entitled 'floppy' ('flojos-loose/la desenvoltura-the free and easy manner/ nudos-knots'), in which her imagination is used to associate words which have little semantic or lexical similarity. An even more personalised example is the group 'cowboys' containing 'fusilamiento-shooting/colillas-cigarette butts'. More potentially useful groupings were the collocation 'chewing' ('mascaba-was chewing/chicle-gum'), and lexical sets 'smoking' ('ceniza/colillas'), 'room' ('el suelo/la salita') and 'newspaper' ('un periodista-a journalist/prensa/press/ periodico-magazine'). The last item was an extra one added from her existing lexical knowledge. Her single formal group was the set 'perfect' ('padecido/atestada/ cometido/aludido'). On a quantitative basis this subject's processing could have contributed substantially to her performance, seeing that she covered almost two thirds of her items, and generated at least 23 potential associations. Qualitatively however, many of her links were somewhat eccentric and dependent on the specific context. She clearly embraced the strategy, including the role of extra non-contextualised items (see 'periodico' above), but used it more for free-association than for structural analysis. Given the imagination she put into the processing it is entirely possible that it could have assisted her retrieval performance, but there is not much evidence that it could have enhanced her knowledge of the lexical structure. It is also likely that with the addition of more target words and source texts some of the associations would have proved to be

more confusing than helpful. Because of the sheer extent of elaboration of target-items, therefore, her grouping is assessed as <u>locally useful</u>, and the support of the interaction for deep learning as there in embryo. But it is clear that the system would have to give her more explicit help in the analysis of structure in the target lexicon, if she were to develop a more linguistically-based appreciation of the relations between words.

S9:

Grouping made up 12.1% of his lexicon-building module processing. He grouped 12 of his 27 selections, 2 more than once. His groups were characterised by syntagmatic relations, eg: his titles 'running his fingers through his hair' ('se peinó-he combed/los dedos-fingers/ cabello-hair'), and 'hate of free and easy manner' ('demás-the rest/desenvoltura-sloppiness') simply reflect combinations of words used in the text. He created the 'smoking' set ('ceniza/colillas'), and a personalised group entitled 'hall' ('atestada-packed/emparedados-sandwiches') which associated 2 of his images from the narrative. His formal categories included the same set of words derived from 'vuelto' as S14 (see above), but for him the group was given the title of the meaning-in-context of one member of the group, 'wrapped', which was possibly confusing and did not encapsulate the morphological similarity between the words. His other formal group involved the very useful grammatical association 'reflexive preterites' ('se peinó/se azoró/se acercó'). Quantitatively, this processing should have been useful, covering nearly half his target items and generating 12 potential associations. However, like S11, the links he created were not informed by an understanding of structural similarity, but more by his memory for phrases and images in the text. For example, in the self-assessment exercise he recalled putting 'revuelto' in the group 'wrapping' and therefore thought the meaning must be similar. He also thought he had grouped 'aguardaba/he was waiting' and 'advertirles/to show them' together meaning 'wait'. As with S11, his grouping may have helped him retrieve some of the items, but it does not provide a basis on which to enlarge the target set and would probably cease to be beneficial once other source texts became involved. It can be assessed as somewhere between <u>locally useful</u> and <u>not useful</u>, but there is only slight evidence of the role of the design in helping him

to develop deep learning approaches, as the most beneficial way for him to develop his syntagmatic associations would have been through the use of the concordancer, and the interaction does not have any way of promoting this if the learner does not take it up for themselves.

S16:

Group-related activity was 16.1% of lexicon-building module processing. She grouped 9 of her 12 items, 4 of them more than once. She created the set 'bicycle words' ('pump/ puncture/riding/bumping'), and also a semantic field title 'sounds' into which she put 'screeched' but did not add to it. Her formal categories , included 'ED words' ('dropped/ploughed/swerved/startled'), 'ING words' ('riding/ bumping'), and the spelling feature 'ough' ('ploughed'). Quantitatively this processing was potentially useful, covering 75% of her targets and creating 13 possible associations, but the groups were not sufficiently semantically rich to make a significant difference to her knowledge or retrieval performance. Her grouping is thus assessed as somewhere between locally useful and not useful depending on other features of her general approach, such as overall comprehension of the source texts. In this case also, there is only slight evidence (in the 1 semantic group she did create) that the interaction has supported the development of any deep learning approach, but it may be that as she was one of the slowest of the subjects overall, more time and further processing might have shown an improvement in this respect.

S15:

His grouping was negligible, making up only 11.7% of his lexicon-building module processing. He created only 1 group containing more than 1 item, a spelling-feature set entitled '/U/' words, containing 'bump/puncture'. This processing is unlikely to have influenced his overall performance and is therefore assessed as not useful. There is no evidence with this subject, that the design has supported any kind of deep approach as far as grouping processes are concerned, and little indication of how it might be modified in order to do so.

S18:

Her relatively high figure for group-related processing (17%) is a reflection of her limited overall processing time (see Table 5.2). As I mentioned earlier, all subjects did a certain amount of grouping in the initial sessions, but this subject did not proceed from there. She grouped 5 of her 14 items, none more than once, creating the lexical set 'bicycle words' ('puncture/ riding'), and the formal category 'ED words' ('screeched/startled/ swerved'). The groups only generated 4 possible associations, and there is no reason to think that it was of sufficient quality to have an effect on her knowledge of, or capability of retrieving, these items. Her grouping is assessed as <u>not useful</u>, and the support for deep learning as not evident.

S19:

He did proportionately little group-related processing (13.3%), even allowing that his percentage was inflated for the same reason as S18. He grouped 7 of his 11 selections, 2 of them more than once, and created the general semantic field 'actions' ('bumping/halt/swerved/startled'). The inclusion of the last item ('startled'), however, shows that his grasp of structural similarity in these words was not clear. As well as the obvious 'ED' and 'ING' words, his formal categories also included 'LY' in which he registered 'suddenly'. Because of the paucity of his data, there is no reason to think that any of the structural associations he generated might have eventually assisted him in retrieval of target words. His assessment for grouping is therefore that it was <u>not useful</u>, and the interaction's support for deep learning as not evident.

S12:

His very high general processing rate (see section 2) meant that the moderate amount of grouping he did constituted only 7.5% of lexicon-building module processing. He grouped 8 of his 37 selections, 2 of them more than once, creating the obvious lexical set 'smoking' ('ceniza/colillas') and the field 'suffering/annoyance' ('molestaba/padecido/desagradado'). He also recorded 2 collocations, one under 'tying' ('nudos/flojos') and one under 'chewing' 'mascaba/chiclé'). His formal associations were 'past participles' ('atestada/

desagradado') and 'imperfectos' ('molestaba'). This processing, covering less than a quarter of his target words and generating only 9 potential associations is unlikely to have had much influence on his learning outcome. The fact that 2 of his target words ('mascaba/odiaba') belonged to the 'imperfectos' group but did not get added to it suggests that his attitude to the activity was less than whole-hearted. His grouping is assessed as not useful, and the support of the design for deep learning as not evident. In view, however, of this subject's success in achieving the objectives of the strategy as a whole (see Section 1) it may be that the failure of the interaction to support grouping was compensated for by the promotion of some other form of deep learning approach. This will be investigated in the next section, where the relation between performance outcome and learning approach is considered.

S13:

She did only 6.2% of group-related processing in the lexicon-building module. Only 4 items were involved, none more than once. She created the 'cigarette' set ('ceniza/colillas'), and also the formal category 'past participle' ('atestada/envuelto'), but failed to add other items ('padecido/aludido') which would have expanded the set. In the self-assessment exercise she remembered that 'ceniza' and 'colillas' were a pair, but couldn't recall if she had made a group for 'desenvoltura' and 'envuelto'. There is no reason to think that this processing might have had a significant effect on her learning outcome and it is therefore assessed as not useful. As was the case with S12, this subject was relatively successful in the general performance aspects of the interaction, and consequently the judgement that the interaction did not support deep learning for her has to be weighed against the possibility that it did so in some way other than through the grouping strategy.

3.2.2 Summary

Table 5.11 shows the subjects' grouping, ranked according to the evaluation above, compared to other aspects of their performance such as retrieval rate (% of the target rate) and target-achievement.

Table 5.11: Grouping Compared to Other Performance

Subject & Group	Grouping (evaluation)	Retrieval rate (% of min. target rate)	Target achievement (%)
S14(SFL)	globally useful	91	100
S10(SFL)	locally useful	72.5	100
S11(SFL)	locally useful	72.5	71
S17(ESL)	locally useful	60	100
S20(ESL)	locally useful	57.5	100
S9(SFL)	locally useful/ not useful	71	85 ·
S16(ESL)	locally useful/ not useful	31	75
S12(SFL)	not useful	100	81
S13(SFL)	not useful	75	100
S15(ESL)	not useful	47.5	81
S19(ESL)	not useful	43.7	65
S18(ESL)	not useful	18.7	29

Only 1 subject, S14, seems likely to have unambiguously enhanced his learning outcome via group-related processing in the lexicon-building module. His combination of a relatively moderate (16.2%) proportion of total processing covering 40% of his items, with a higher-than-average level of quality in terms of the structural relevance of his group titles, enabled him to maintain his learning rate whilst providing a basis for continuing enlargement of the target word set.

Of the others, three (S17, S20 & S11) committed themselves to the strategy to the extent that they covered half or more of their target words and attempted to generate a significant number of useful associations, but the looseness and context-dependent nature of their linking would appear less likely to enhance their knowledge of target items, and would probably serve retrieval of them only in the short term. These learners may have traded off, to some extent, a higher learning rate for the time spent in inconclusive grouping activities, although a closer analysis of S17's processing suggests that she developed an interesting, original, strategy of her own. S17 used a clue only once in 54 attempted retrievals, otherwise depending on the given context (ie: the automatically-provided KWIC citation with the target-item gapped out, which is the initial prompt for the retrieval of an item - see the description of testing module functionality in Chapter Four). In other words she did

not use notes or group clues, although, in the lexicon-building module, she had done a fair amount of processing of both types. In fact, as we have seen from Table 5.10, her grouping-related activity in this module was the highest of all, which raises the question of how this processing fitted in with her context-based approach to retrieval. Like S20, S17's high level of group-related processing in the lexicon-building module was mainly due to the number of times she reviewed the contents of existing groups, a tactic which was rarely used by other subjects (with the exception of S11). A review of the items in a group can presumably be used for a number of purposes, to rehearse the items themselves, to re-assess the appropriateness of the titles, to see if other links exist which have not yet been created etc. The difference in the way the S17 and S20 used the group review is that S17 tended to review all her groups at the same time, one after the other, whereas S20 interspersed group reviewing with other processing, such as clicking one of the items in a group to see what notes he had written for it. S17's strategy here may have been to use the groups to mentally rehearse the constituent items, like a form of recognition test. If this was the case, then she must have had very good comprehension and recall of the originating text, because she was apparently able to relate each decontextualised item in a group not only to its meaning out of context, but also, without further processing, to its original context. (In fact it is true to say that her general level of English was the highest of the ESL subjects). This strategy, which represents a deep learning approach in the understanding-of-the-text sense discussed in Chapter Three, (although it does not necessarily result in explicit awareness of lexical structure within her target-items), was obviously successful as far as target achievement is concerned (Table 5.6). However, the time spent in the effort to retrieve items without using clues may have considerably slowed down her learning rate (Table 5.2). This subject's approach demonstrates that learners may find unexpected ways to adapt the functionality of a program to their own preferred ways of working.

Finally, two learners (S12 & S13) clearly abandoned the grouping strategy at the first opportunity. This may have proved beneficial for their retrieval rate, and in S13's case did not appear to harm target achievement either. Other aspects of these

learners' performance and outcome will have to be investigated in order to establish whether the 'non-grouping' strategy is a viable alternative, or whether there is any way in which their learning suffered as a result of their decision not to group their target words.

In general, this analysis demonstrates that learners at this level of competence find grouping difficult, and that the structural awareness they exhibit is low. Very few subjects, for example, created genuine collocation groups (one instance was S11 & S12's 'mascaba/chicle-chewing gum'). Some of the ESL group might have been expected to register the association of 'struck /match', or 'screeched/halt' or the phrasal nature of 'bumping/along', 'came/up','turned/out' etc., but in the event none did. In part this failure relates to a general failure to use the concordancer (this will be discussed in the chapter summary), but it may also be the case that the program design encourages 'paradigmatic' thinking at the expense of collocation. It is also noticeable that whereas almost all the SFL subjects expressed their group titles in L1, the ESL subjects uniformly used L2. This may have been a function of English being the language in which both groups communicated with the programme supervisor (usually there was some discussion around the group-creating procedure), or it is possible that the ESL group, by virtue of their circumstances, were simply more used to operating in L2. The advantages of using L2 are obvious, whilst the problems of thinking up titles do not seem to be any worse in L2 than in L1 (witness S9's 'hate of free and easy manner'). Clearly it is necessary to encourage learners to use L2. The need for tutorial support for this activity is evident, because of the potentially counter-productive effect of time spent generating associations which may not only be non-useful but may even be confusing. The fact that many subjects created the same groups, eg: 3 of the ESL subjects created a group entitled 'feelings', 5 of the SFL subjects created 'smoking/ cigarette' etc. suggests that tutorial help is feasible. The key factor is that they tended to select the same words, and there is a limited number of ways that a given set of words can be semantically grouped. If it turns out to be possible to predict which words a given group of learners might select from a given text, then it might also be feasible to predict and assist the generation of useful groupings.

3.3 Summary of Section 3

Deep learning approaches as represented in the use of a grouping strategy did not play the role in determining performance outcome which the system design intended. This was because most subjects did not reach the stage at which this strategy would have become necessary to their learning aims, rather than merely feasible. However, it did feature to some degree in all subjects' performance and assessments of differences in both quality of group-related processing, and its relevance to overall learning strategy are possible.

One subject, S14, demonstrated an approach to grouping which combined an awareness of the potential of the activity for learning, with a quality of processing which suggests that his knowledge of both specific and general structural characteristics of target items was enhanced. This subject's success in terms of target achievement and a comparatively high learning rate are taken as concomitant with this assessment of his grouping, and as evidence that the strategy is linked to quality in the learning outcome. Three others (S17, S20 & S11) did enough processing in quantitative terms to affect their learning outcome, but did not meet the qualitative requirements for that effect to be sure of being beneficial. For these subjects, processing time spent in grouping may have been traded-off against success in improving their retrieval rate, as indicated by the no-clue strategy adopted by S17, and the virtual abandonment of grouping by two of the subjects (S12 & S13) with higher retrieval rates.

Evidence of subjects adopting a wider interpretation of the grouping strategy is thin. Categorisation at selection time was not developed along these lines, and the use of group-prompts at retrieval time does not demonstrate any significant favouring of this kind of prompt over the use of notes. Generally though, learners who did more group-related processing in the lexicon-building module, tended to use notes (and, consequently, L1 translation) less, which suggests that a wider grouping-based (and L2-focused) learning strategy is possible, and that, under conditions which restrict the usefulness of other strategies, it could be developed. The combination by one of

the learners, S17, of a retrieval strategy which did not use prompts, and a high level of group-related processing in the lexicon-building module, showed that there are other ways that grouping can be used, and that the strategy as defined is not the only approach into which group-processing can fit. The implication of these findings for the design of the system is that it should be supportive of a range of learner behaviour, but the generally low quality of subjects' groups, the predominance of context-related lexical sets, and the absence of collocational associations, indicates that tutorial support is needed in certain aspects of the task. The balance required is between supporting learner autonomy and adapting to learner needs. There is some evidence that learners of the same general background (eg: from the same course) may tend to select the same target words and to group them under similar titles, which indicates a possible role for a general learner model in the provision of suggestions and examples to help those for whom the grouping activity is not intuitive. This will be discussed further in the next chapter.

Section 4: Deep & Surface Approaches Related to Performance Outcome

A surface/deep framework was defined in Chapter Three as a way to describe the learner's perception of structure in the learning task and learning content. A surface approach is one which focuses on features of the task in isolation, in this case on the meaning of isolated L2 target-items, normally expressed as L1 translation, a deep approach looks for underlying principles connecting the task features into a system, in this case the system of lexical relations tying target-items to each other and to the meaning of the text in which they occur. The aim of the program design is to help the learner to develop a deeper level of structural awareness by focusing on processes which involve the use of structural lexical features to aid retrieval.

The purpose of this section will be to establish whether the approaches of the subjects in the test programme, classified on the basis of this distinction, reflect the performance outcomes described in the previous two sections. The data which will be examined is:

i) Subjects' comments on how they approach the general problems of vocabulary-learning (collected in an interview in the case of the ESL subjects, and during target-word selection processing in the case of the SFL group).

ii) Subjects' comments on the meaning and categorisation of target-items.

iii) Subjects' notes written in the lexicon-building module.

iv) Patterns of retrieval in the testing module.

v) Subjects' comments during the self-assessment exercise.

Because the learning approach is not a stable characteristic of the learners themselves, but of their perception of the task, and because the surface task as represented by the different program modules varies, it is likely that these different data sources will produce superficially different results for the same learner. However, the underlying issue is their appreciation of lexical structure, and there may be consistency in this, detectable as a tendency, in each subject's approach to the task, towards one or other end of the surface-deep scale. Once each subject's general tendency has been identified, it will be compared with the performance outcome discussed in the previous sections. If the overall design of the program is to be validated, we would expect to find that those learners adopting the most markedly deep approaches also register the highest quality learning outcomes in terms of measures such as retrieval rate and target-achievement ratio.

The assessment of learner comments as evidence of deep or surface approach is essentially subjective, but insofar as it is based on well-defined criteria for what constitutes attention to structure in the learning content, it can be subjected to validation by independent judges. Because of the complexity of the qualitative description involved, the assessment was conducted in two stages. Firstly a pilot phase, to test and improve the classification criteria, and secondly the validation of the data on the basis of these criteria. The pilot judge (an academic with experience

in the field of technology and language learning) was asked to assess examples of learners' comments and classify them as deep or surface, on the basis of the following distinctions:

Surface - vocabulary-learning is viewed as a process of *memorisation, target-items* are referred to in terms of L1 translation and/or intra-word features.

Deep - vocabulary-learning is viewed as a process of comprehension, *target-items* are referred to in terms of contextual meaning and/or inter-word relations.

The judge was asked to note, whenever appropriate, reasons for classifying as they did, and to indicate wherever they found it impossible to make a straightforward deep/surface assessment. A set of detailed criteria were produced, on the basis of this assessment. These criteria, (described in the introduction to each sub-section below), were then used as the basis for assessment of deep and surface approaches, in the analysis below. The same criteria were also given to a second, external, judge (an experienced language teacher), who was asked to evaluate the same data. The agreement between pilot judge and researcher was 65%, which rose to 93% between the external judge and the researcher, following clarification of the criteria given. The latter judgements were used in the following analysis. (All three assessments are detailed in appendix III). The evaluation of deep and surface characteristics in this analysis is therefore considered to have been objectively validated.

4.1 General Approach to Vocabulary Learning

Evidence to characterise subjects' overall approach to the issue of vocabulary-learning was sought in an interview (for the ESL group), and in general comments (for the SFL group). Surface approaches were expected to be associated with a focus on the difficulties of memorisation, and the devices needed to overcome these problems. Deep approaches were expected to show themselves in references to

156

register, changing roles of words, and the need to relate new words to existing knowledge. The criteria for assessment of this data are as follows:

Surface: The Subject views vocabulary learning as mainly to do with memorisation of individual word meanings. They refer to lists, to repetition, and the need to associate L2 items with their L1 equivalents.

Deep: The Subject views vocabulary learning as mainly to do with understanding the context that L2 items are encountered in, and/or the relations of meaning that exist between these items. They refer to context, to gist, to the relation of L2 items to other L2 items, and to issues of lexical structure and register. Table 5.12 gives some examples of subjects' comments together with an assessment of them as deep or surface:

Table 5.12: Interview Data on Approach to Vocabulary Learning

Subject	Reference	Comment	Deep	Surface
S17 (ESL)	How do you go about learning vocabulary?	..vocabulary is difficult when I'm talking..for example: these (booklets, pamphlets etc. on a rack) are 'leaflets' ..but some are 'magazines'..I try to use a general word..	√	
		..I try to repeat (write) new words on paper 5 times..it works		√
S18 (ESL)		..If I don't write it (straight away) I don't remember.. ..I write it down 10 times in a notebook..it works!		√
		..I never use words which I don't know what they mean..with colleagues it doesn't have to be academic..for someone with more culture I am more careful..	√	
S20 (ESL)		..I create words from Portuguese (but the pronunciation is difficult)...it takes longer to use an (L1 dictionary)..though I think it's better..one words pulls another..I try to use words in conversation..	√	
S19 (ESL)		..I write them down then I put them in my mind..I'm quite good at remembering..		√
		..when I go to the toilet I take the English dictionary..it's not good to use a Portuguese dictionary..it makes confusion..you need to think in English.. you need to read books.. newspapers every day	√	
S15 (ESL)		..no system for learning words.. sometimes I look in the dictionary..I'm lazy		√
S16 (ESL)		I have a note book for English and French..I check words and try to speak with those words...I (used to) (re-wite) them once a week..it is very useful to repeat all the time..after 6 months people were very surprised because I have improved..		√
		I used to know every irregular verb..I cover the columns and test myself..I never use auxilliary verbs..I know them and when to use but not how to use.		√

S10 (SFL)		I don't like to jump into the dictionary.. I look up eventually.. always like to get a general meaning then go for the words..the next time I see a word - if I've struggled to get it I can use it confidently	√	
		..You can do a summary without knowing the meaning of 10% of all words..you can pick up a dictionary & look up - then you don't remember..	√	
S12 (SFL)		..lots of different texts..underline words..look them up..going back 3 or 4 times..hoping it will have sunk in..word lists not very useful..	√	?
S14 (SFL)		..I don't always rely on a dictionary.. there are lots of colloquial meanings.. you can get the general meaning..	√	
S13 (SFL)		..I try to get the root of the word..break it up..look for the infinitive..then you can know what the rest of it means..	√	.
		..It doesn't matter what individual words mean..I must read the whole text to get an idea of what this part means..go back to the words when I've established what the rest of the text is about..	√	
S9 (SFL)		..I tend to look for the gist first...then look up more words rather than fewer..	√	
S11 (SFL)		..the answers (vocabulary items) I just look up slip out of my mind the fastest		√

Some of these examples clearly refer to surface concerns with memorisation (S17's second comment, S18, S19, S16 etc.). These are, notably, ESL subjects whose awareness of the scope of the vocabulary-learning problem is arguably sharper than that of their SFL counterparts, because of their circumstances. (Many of them also described strategies for getting people to help them with vocabulary in everyday conversation, and claimed to make a point of trying to use new words as soon as possible. This may have been a case of them merely repeating what their teachers had told them to do, because there was not much evidence of them doing the same with the words they learned during this programme!). There are hints of the same concern with memorisation in S12's approach, although he departs from the others in his attention to words-in-context rather than decontextualised lists - for this reason his approach is difficult to classify, having both deep and surface elements in it. Other examples are obviously deep (eg: S17's first comment, S20, S14, S13, S9 etc.), in that they focus explicitly on the existence of relations and associations of sense, syntax or discourse, and express the intention to get to grips with this rather than simply memorising the words concerned. Further difficulties in classification arise with examples like S10 (3rd comment) and S11, who appear to be critical of surface approaches without necessarily indicating a deep alternative. This data, nevertheless, suggests that the ESL learners are more prone to surface approaches to

the overall task than the SFL group, possibly because of their less academic background and circumstances. We can see clearly, however, that it is possible for learners to exhibit both approaches within the same task (eg: S17, S12), wherein the surface elements may be there to service a deep intention. An approach is classified as surface only if there are almost no deep characteristics evident.

4.2 Comments on Meaning and Categorisation of Target-Items

The data here is in the form of comments made, or answers to questions given, during processing in the selection module related to inferring meaning and categorisation of items in the source text. This data (Table 5.13) illustrates differences between deep-level reference to understanding of text, grammar or register whilst considering selection of a word in the source text, and surface-level reliance on association, whether L1 or L2, in the inference of meanings of target-words. Categorisation involves classifying items as important because of their meaning, or form (spelling or grammar), or because they have a significant contextual role. The data include comments that subjects made when asked to say why they had put selected items into MEANING, FORM or CONTEXT boxes. The criteria are:

Surface: The subject is mainly concerned with finding an L1 translation for the target item itself, from memory or from literal translation without taking sense and/or grammatical structure into account, or else tries to relate it to other L2 words on the basis of a superficial resemblance (eg: spelling or sound). The subject uses simple formal classification criteria (simple verb tenses, regularity, gender, plurality etc.).

Deep: The subject tries to sensibly interpret the meaning of the item's immediate context, or else to analyse it grammatically, or relate it to other L2 words on the basis of meaning or the register of the text. The subject uses semantic or collocational criteria, or more complex formal criteria (verb mood/aspect, spelling rules, affixation).

159

Table 5.13: Approach to Inference of Meaning and Categorisation at Selection Time

Subject	Reference	Comment	Deep	Surface
S17 (ESL)	'..race riots..'	..it's about foreign/ black people - I have to put it together..	√	
	'jaw' ('..her jaw dropped..')	..it's not a verb..something on the face..the same for the animal in the film?	√	
	'drive' ('..the advert stresses initiative and drive..'	'drive'..is different from drive a car..	√	
S18 (ESL)	'attaching'	..like attack/pressing?		√
	'startled' classifying for form	..it is past (tense)..		√
S20 (ESL)	'..fit into..'	'fit' means suitable..it has other meanings..it's a phrasal verb	√	
S19 (ESL)	'..find the work interesting and challenging..'	'challenge' is like 'discover'?	√	
	'startled' classifying for form	.. I put it in form because it's like 'start'		√
S16 (ESL)	'riding' classifying for form	..because it is a small word..		√
S10 (SFL)	'atestada' (packed)	..I get the impression something has been started/triggered	√	
	'molestaba' (bothered)	..obviously not a pleasant word	√	
	'colillas' (cigarette ends)	..not sure if its a conjugation or a noun..it's a noun ending..	√	
	'señaló' (indicated/announced ..also fix/settle)	..I knew I had seen it before..there's another meaning.. to fix..I remember the context of the other meaning but not this meaning here..	√	
	'envuelto' (wrapped up) classifying for form	..part of 'volver'..shall I put it in form?	√	
S12 (SFL)	'el aludidio se azoró' (the person referred to was alarmed)	..he aluded to got alarmed..		√
S14 (SFL)	'levantó'/'molestaba'	..We've just done imperfect & preterite..levantó and molestaba catch my eye...	√	
	'envuelto'	..a stem-changing verb with perfect ending	√	
S13 (SFL)	'desagradado' (displeased)	..agradable means 'pleasant'..desagradado is the opposite..	√	
	'fila' (row)	..I don't know it but it doesn't seem important..I think I can get away with it..only if it's going to help me with the text..	√	
	'la larga espera que habian padecido' (the long wait they had suffered)	..the sentence is in my head..if it comes up in the test I'll remember it - without knowing what it means..		√
	'padecido' (suffered - from infinitive 'padecer')	..from 'decir'..looking at the root..*pp. of 'decir'..	√	
	'guardara las formas' (observe the niceties)	..look after?..I'm thinking in Italian..		√
S9 (SFL)	'corbatas ...flojos' (loose ties)	..their ties were floral?		√
	'el ministro odiaba la desenvoltura' (the minister hated the casualness/ sloppiness)	..minister dared the desenvoltura.. ..minister hated the unwrapping..		√
	'dejadme' (leave it to me)	..that's an imperative..he's obviously quite irate..	√	
	'sea' (subj. verb 'to be')	..that's a subjunctive..he's trying to influence him? ..all the time they've been trying to get a question at him..	√	
S11 (SFL)	'envuelto' (wrapped)	..it's part of last year's vocab..'muy basica'		√
	'constestar' (to answer)	..I can't contest the questions..		√

160

Most of the comments exemplifying deep approaches at the selection stage come from the SFL group, although some of them showed surface tendencies too. S12, S9, S13, S11, for example, are shown here exploring superficial associations based on formal similarity and literal paraphrase. For the ESL group, S17, showed a consistently deep approach, relating words to context, grammatical structure and alternative meanings in L2, except for her categorisation criteria which tended to be superficial. Data for categorisation is in fact generally thin, with evidence from only one of the SFL group. As was mentioned in Section 3, this part of the interaction was not fully understood by learners, and their performance levels did not reach the point where it should have become more relevant. SFL learners such as S10, S14, S13, S9 show deep perception across a range of structural possibilities, eg: discourse-related, word-sense-related, grammatical structure etc., although S10 and S9 illustrate that in some cases their assumptions, whilst structurally-oriented, were factually inaccurate. 'Atestada' (S10's 1st comment) has no connection with 'triggering', although the subject has correctly perceived the mood of the discourse. '*Decido' is not the past participle of 'decir' as S13 surmises (4th comment), although 'padecido' is a past participle. These contradictions demonstrate that learner approach, and knowledge of content, are distinct and complementary rather than causally related phenomena (see Chapter Three), and thus confirm the need for separate sources of data on each.

4.3 Subjects' Notes on Target-Items

The data considered here consists of the subjects' notes and translations for their target-items, and some of the comments they made whilst writing them. These are assessed according to whether they rely on L1 or L2, and whether they contain any direct or indirect reference to grammatical structure or meaning in context. A sample from each subject is given in Table 5.14. The criteria are:

Surface: The subject is concerned with recording a single L1 translation, or else with recording a superficial feature of the item, or an irrelevant association of their own (ie: one that doesn't relate to the meaning or structure of the item).

Deep: The subject is concerned with L2 synonyms, homonyms, collocations and sense relations, or else with recording complex or multiple forms of the target item (eg: past tense as well as infinitive, if it is a verb), or an aspect of meaning in context

Table 5.14: Notes and Translations

Subject	Target Word	Sample Note/Comment	Deep	Surface
S17 L1=Spa L2=Eng	'aware'	Informed, knowing, knowlegdeable Informado,conociendo, conocimiento	√	
S18 L1=Spa L2=Eng	'aware'	notificar information general		√
S20 L1=Por L2=Eng	bumping	hit strike slam against some thing acao de bater em alguma coida	√	
S19 L1=Por L2=Eng	'aware'	agonizante/aterrorizado		√
S15 L1=Por L2=Eng	'ploughed'	like a imperfect field, not plain	√	
	'nip' ('..there is a nip in the air..')	' can I use 'there is a nip from smoking'?	√	
S16 L1=Fre L2=Eng	'aware'	knowledge politically ...connaissances en politique		√
	Looking up 'struck' (from 'he struck a match') in dictionary	'..to 'hurt' a match..'		√
S10 L1=Eng L2=Spa	'padecido'	suffer, endure, put up with		√
S12 L1=Eng L2=Spa	'padecido'	padecer-to suffer		√
	'acusador' (adjectival form)	'..what part of language is 'acusador?'	√	
S14 L1=Eng L2=Spa	'se la puso' (he put it on)	pret. poner. chaqueta	√	
	'se sequia' (went on)	'..I'll put 'pursue' because the verb is like the French...'	√	
S13 L1=Eng & Ita L2=Spa	'la salita' (the little room)	la .... era piena		√
S9 L1=Eng L2=Spa	'puesta'	dressed		√
	'desenvoltura'	'..because he was picking them up on their behaviour..'	√	
S11 L1=Eng L2=Spa	'desenvoltura'	free and easy manner		√
	'habian padecido'(had endured)	'..it used the perfect form with 'haber..'	√	
	'puso' (put - preterite of 'poner')	'.I get muddled with poner.. pedir..podir..'	√	

Although subjects used a range of methods (L1 paraphrase, L2 synonym, grammatical reference etc.), most tended to stay with one or other technique - the samples above are representative. This time it is the ESL group who adopted the deeper approaches, mainly because of what was for them a default decision to operate in the target language, whilst the SFL group tended to use L1. S17 & S20 illustrate the employment of a range of L2 synonyms as well as an L1 translation. S16 does the same, but in her case the translation is completely inappropriate for the contextual meaning. Of the SFL group, only S14 adopts a deep approach, utilising grammatical information and the co-text. The other SFL subjects fall back on L1 translation (in S13's case her L1 is Italian).

4.4 Patterns of Retrieval

An aspect of performance which reveals a surface approach to remembering the meanings of items was discussed in pilot study 2. This was the tendency to recall words as L2-L1 pairs (see Chapter Four Section 4.2). It was shown in that study that L2 target-items may become disassociated from their L1 meanings, and that the result is a kind of trial-and-error process in which the learner tries to re-associate them by making guesses from a mental list. A similar process can be observed during item retrieval in the testing module, where it is manifested by the inappropriate production of target-items, ie: the subject retrieves a target-item which is not the one currently being tested. (For example, S13 retrieved 'ceniza' when the word required was 'colillas', S16 retrieved 'ploughed' when the word being tested was 'struck'). In some cases these errors may be the product of specific associations, eg: S13's confusion was between 2 items which had been put into the 'smoking' group, but in other cases there is no connection and this is taken to be indicative of the disassociation of items from their meanings and other features. Such is the likely outcome of a surface approach. With a deep approach, the context of the target-item, or the L2-related evidence provided by available clues (notes involving synonyms or paraphrases, relevant groups) should serve to activate some connections to the item being tested, so that what is finally retrieved, even if not wholly correct, can at least be seen to bear some relation to it. Subjects for

whom the retrieval of unrelated target-items happened most can be identified from the program's feedback record. This records whether the string of letters retrieved is the same as, or near to, one of the words in the target list.

Table 5.15: Retrieval of unrelated target-items

Subject	Total Retrievals	Retrievals of unrelated items	%
S10(SFL)	42	0	0
S9(SFL)	57	0	0
S14(SFL)	71	0	0
S12(SFL)	101	1	.9
S11(SFL)	79	3	3.8
S13(SFL)	45	2	4.4
S15(ESL)	86	4	4.6
S17(ESL)	54	4	7.4
S20(ESL)	78	7	8.9
S18(ESL)	19	4	21
S16(ESL)	51	12	23.5
S19(ESL)	36	12	33

Table 5.15 shows, for each subject, the number of retrieval attempts which were the same as, or near to, a target-item which was not the one being tested, and thereby unrelated to the item being tested. Subjects are ranked in order lowest to highest for percentage of this kind of error. Those with high errors of this type are likely to have dissociated the item from its meaning or context, evidence of a surface approach. This data tends to confirm the emerging pattern throughout this analysis, which is for the SFL group as a whole, with S10, S12, S14 in particular, to do better on the evidence of deep learning approaches, for 2 of the ESL learners (S17 , S20) also to do well in this respect, but for the rest of the ESL group (S16, S18, S16, S19) to demonstrate evidence of predominantly surface approaches.

4.5 Approach to Self-Assessment

In the self-assessment exercise, subjects were asked to comment on each of the words in their final target list (presented in order of selection) in terms of what they now remembered about its original context, whether they knew its meaning, and whether they could use it themselves. The assessment was interpreted as evidence of learning outcome and the comments as indicating depth in the learner's approach to the learning content. The learning outcome can be represented quantitatively,

assigning a score of 2 for each 'yes' judgement, 1 for each 'maybe', 0 for each 'no', and representing the total as a percentage of the maximum score (ie: what they would have scored if they had answered 'yes' to all 3 categories for every item). This data again reflects the general pattern distinguishing SFL learners from ESL (with the exception of S17, S20). The comment-data collected during this part of the programme came mainly from the SFL group, who were interviewed in their native language (English), and from S20 and S17 of the ESL learners, who were possibly the most fluent of the group. The criteria used to evaluate the self-assessment comments were:

Surface: Subject only refers to the meaning of the item as a single L1 or L2 word or concept, or to a very general idea of its originating context, or to a superficial syntactic role.

Deep: Subject refers to the item's role in the wider context of language (eg: as part of a register), or to its specific co-text, or to a complex syntactic role, or to its possible senses and uses.

Table 5.16: Self-Assessed Learning Outcome

Subject	Score	Maximum Score	%
S10	122	126	96.8
S17	105	138	76
S9	118	162	72
S20	74	108	68.5
S11	139	210	66
S14	110	168	65.5
S12	145	222	65
S13	72	150	48
S18	38	84	45
S16	26	72	36
S15	36	126	28.5
S19	14	66	21

In Table 5.17 a selection of comments exemplifying mainly deep approaches is shown. S13 and S9 are the exceptions. The data from the other ESL subjects (S15, S16, S18, S19) was very thin, mainly consisting of unsuccessful surface-level

attempts to recall single meaning-concepts for each target-item (reflected in their scores in Table 5.16).

Table 5.17: Comments on Self-Assessment

Subject	Target Word	Comment	Deep	Surface
S10	'señaló'	I've been coming across it a lot in the news..it's press language..	√	
	'guardara las formas'	..would come in handy in an essay..it was subjunctive.. must have been a 'que'..	√	
	'desenvoltura'	..yesterday, in an article about drugs..	√	
	'enfrentó'	I saw it yesterday..I dared to use it in my summary	√	
S12	'emparedados'	'..revueltos..'	√	
	'flojos'	'..nudos..'	√	
	'molestaba'	' no guardara las formas '	√	
	'seca'	'..con voz seca..' (with a dry voice)	√	
	'espera'	..long wait they had suffered - padecer!	√	
	'cabello' (hair)	'..hair..greasy..'desenvoltura'?	√	
S14	'colillas'	..'el suelo' (floor)...cigarette ends	√	
	'las miradas' (the glances/gazes)	..of the journalists.. 'periodistas' was one of the first words I learned in Spanish	√	
	'cabello'	..scruffy hair..'se peino'?	√	
	'odiaba' (he hated)	..opposite 'gustar'	√	
S13	'prensa' (the Press)	..similar to Italian ..'presa' to be in a hurry..means something different..		√
	'se peinó' (he combed)	'..pettinarse..'(Italian - to comb)		√
	'lamentó' (he lamented)	.I knew what it meant so I didn't pay much attention to the context..		√
S9	'emparedados' (sandwiches)	..I'd use 'bocadillos'..		√
-	'puesta' (from 'la chaqueta puesta..' with a jacket on..)	.sus propios puestas..does it mean 'own'?		√
S17	'aware'	..realise..he was riding a bicycle and he aware...he was riding in a (ploughed) road..	√	
	'jaw'	..the body..the lady had (dropped) the jaw..the sensation was surprise..	√	
	'earnings'	..something you get from the job..high potential earnings	√	
	'crawls'	..when you go to visit place..the context means pub..in the same day..crawl of pub..can you say shop crawl?	√	
S20	'aware'	.I tried once but did it in a wrong way..I try to use this as 'knowledge'..in a composition	√	
	'bumping'	..he aware the bicycle was bumping	√	

S10's learning outcome, as reflected in the confidence of her assessment, was greatly superior to anyone else's. Her approach, revealed in the comments, shows that she related the items she learned during the programme, to a wide conception of their relevance to language-learning in general. The other 2 SFL subjects (S12, S14), and the 2 ESL subjects (S17, S20) demonstrate deep approaches in their explicit references to features of the L2 co-text in which the target-items were

encountered. These comments can be compared with those of S13, who tended to relate more to her L1 (in her case Italian as well as English).

4.6 Approach Related to Performance

The qualitative assessments made in this section can now be compared with the quantitative measures from Section 2 (Table 5.18). Subjects are rank ordered in 3 groups according to learning rate multiplied by target achievement.

Table 5.18: Qualitative and Quantitative Evaluation

Subject	Learning Rate (wph)	Target Achieve-ment %	Self-Assess-ment %	Vocab general approach	Approach to Selection and Notes	Self-Assessment comments	Unrelated retrievals%
Top Group							
S14 (SFL)	7.3	100	65.5	Deep	Deep	Deep	0 low
S12 (SFL)	8.7	81	65	Mainly Deep	Mainly Deep	Deep	.9 low
S13 (SFL)	6	100	48	Deep	Mainly Surface	Mainly Surface	4.4 low
S10 (SFL)	5.8	100	96.8	Deep	Mainly Deep	Deep	0 low
S9 (SFL)	5.7	85	72	Deep	Mainly Deep	Mainly Surface	0 low
S17 (ESL)	4.8	100	76	Mainly Deep	Deep	Mainly Deep	7.4 low
S20 (ESL)	4.6	100	68.5	Deep	Deep	Mainly Deep	8.9 low
Middle Group							
S11 (SFL)	5.8	71	66	Surface	Mainly Deep	N/A	3.8 low
S15 (ESL)	3.8	81	28.5	Surface	Deep	Surface	4.6 low
Bottom Group							
S19 (ESL)	3.5	65	21	Mainly Surface	Mainly Deep	Surface	33 high
S16 (ESL)	2.5	75	36	Surface	Surface	Surface	23.5 high
S18 (ESL)	1.5	29	45	Mainly Surface	Surface	Surface	21 high

In general we can see that predominantly deep approaches are associated with the higher retrieval rates and target achievement of the top group. The bottom group have the lowest retrieval/target achievement and the lowest incidence of deep assessment. The boundaries separating the middle group from the others are not so clearly defined, due to the uncertainty of the qualitative measures over a small amount of data. This confirms that, for some learners, both approaches may be

equally evident in the learning process, and that, for some aspects of this interaction, a degree of successful performance may be achieved via a surface approach. The analysis of the extremes of the range, however, demonstrates that the program's quantitative measures of successful learning are reflected by the external categories of deep and surface learning applied to the learning process.

Section 5: The Learner Model

This section will compare values assigned to items in the system's model of the learner with their own self-assessment, in order to test how well the model represents their conscious knowledge of the target vocabulary.

5.1 Sequencing Strategy for Testing

The system's learner model consists of a representation of the processing the learner has done to date. It is the basis on which the program adapts the sequencing of items for testing, to the individual learner. The information it contains is an abstraction from the record of events (keystrokes and mouse clicks), and represents the learner's activity in terms of 3 cumulative measures of different aspects of the vocabulary-learning process (see Chapter Four):

i) General processing events (initial selection of items, subsequent location of those items in their source texts, KWIC lookups, notes-saving events).

ii) Grouping events (assignment of items to groups).

iii) Retrieval events (correct retrievals, incorrect re-retrievals).

When the learner enters the testing module, the program sequences the target items, using the relative values of these 3 measures, and applying the principles of maximising the effects of association (eg: preceding weakly-associated items by strongly-associated ones), and structured reviewing (staged recycling of items that need further practice). These principles were described in Chapter Four. In Pilot

Study 2 a comparison was made between the learner model's representation of the subject's knowledge at the end of the study, and his own assessment. It was concluded that discrepancies in the correspondence between the 2 representations might be useful in indicating something about the contribution of the model to the learning process, and its usefulness in helping to evaluate it. Whether the sequencing actually does contribute to the learning process, must depend on the relevance the learner model has to qualitative aspects of the learner's knowledge of the target words. These aspects are themselves dependent on the quality of processing, from which the model has been abstracted. If the learner has adopted a deep approach, the model should be relevant to the learner's knowledge, whereas if they have adopted a surface approach, it is less likely to be. Relevance could be investigated by comparing the models of learners adopting deep approaches, such as S14, S20, S17 (see previous section), with those adopting surface approaches, such as S16, S18, S19. However, because the models are concerned with relative values of items within the same target list, and do not refer to any objective standard, it is not possible to make the comparison directly. Instead, each model will be compared with the learners' own evaluation of what they know, as indicated by the self-assessment exercise. Broad distinctions in this assessment can be made, between items which the learner appears to know nothing of (ie: no ticks in the self-assessment exercise), those which they seem to know fully (3 ticks), and those which are somewhere in between (1 or 2 ticks). If what the learner model represents is in any way equivalent to what this subjective assessment shows, then there should be some correspondence between the number of ticks an item has, and the ranges which the model places it in. Any consistency found in the type of correspondence for deep-approach learners can then be contrasted with that found for surface-approach learners, to show the relevance the models have to the types of approach the learners have adopted.

Given the complexity of lexical knowledge it is not expected that there would be an exact correspondence between the highly simplified quantitative measures that the model contains and the qualitative judgements the learners make. The most likely discrepancy, because of the relative poverty of the quantitative representation, is

169

one in which the model under-rates the learner's knowledge (ie: items the learner knows are in the low ranges in the model). Some element of this is inevitable, as the model compares items in the same set with each other, regardless of how deeply-processed (and thus well-known) the lowest of them may in fact be. An extreme case of this is found with subject S10. Here, a low level of overall processing with high confidence in self-assessment produces a very large degree of under-rating by the model. More significant, therefore, is the extent of over-rating (ie: items the learner does not know are in the high ranges). This will occur because the processing the learner has done has not resulted in any confidence that they know the items. If this over-rating is found to be the case consistently for the surface-approach learners, but not for the deep-approach learners, then it can be concluded that the model is indeed more relevant for the latter, and that the sequencing based on it is more likely to contribute to their learning processes.

The key factors in the sequencing are low retrieval, high general and/or group processing, and high retrieval (see the rationale given in the re-design of the model in Chapter Four). In the interests of simplification, the 7 categories will be compressed into 3 bands representing: a) low retrieval + low-mid groups/other processing (Chapter Four categories iii, iv), b) high groups/other processing (categories i, ii), and c) mid-high retrieval (categories v, vi,vii).

In the example below (Table 5.19), self-assessment is compared with these factors for subject S14. Columns A1 & A2 show items which have 0-½ ticks (½ being where the subject has indicated 'maybe' for one of the assessment categories), compared to those classed as low retrieval + low or medium general and/or group processing (items in bold are the ones which correspond); columns A3 & A4 show items with 1-2 ticks, compared to those with high general and/or group processing; and columns A5 & A6 show items with 2½-3 ticks compared to those with high retrieval. Items in italics are those which do not correspond.

The main correspondence in this subject's data is between 'known' items (column A5) and those ranked high for retrieval (column A6). There is also a certain amount

170

of correspondence between 'not known' items and items ranked low for retrieval. This reflects the process definition of learning discussed in section 2, wherein the retrieval measure is taken to be representative of other processing which has occurred.

Table 5.19: Comparison of Learner Model and Self-Assessment

SELF-ASSESSMENT	LEARNER MODEL
A1.Items with 0-½ ticks	**A2.Low retrieval/low-medium general and/or grouping**
desenvoltura	**desenvoltura**
se acercó	**se acercó**
se adelantaba	**se adelantaba**
el aludidio	**el aludido**
hubo	**hubo**
comportamiento	**comportamiento**
se la puso	*se seguía*
	cometido
	se peinó
	los dedo
	odiaba
	gritó
	mangas de camisa
	amenazador
A3.Items with 1-2 ticks	**A4.Low retrieval/high general and/or grouping**
se peinó	*cabello revuelto*
cometido	*mascaba chiclé*
padecido	
colillas	
mangas de camisa	
A5.Items with 2½ -3 ticks	**A6.Medium/high retrieval**
la salita	**la salita**
el suelo	**el suelo**
envuelto emparedados	**envuelto emparedados**
caballeros	**caballeros**
flojos	**flojos**
miradas	**miradas**
seca	**seca**
molestaba	**molestaba**
nudos	**nudos**
cabello revuelto	*padecido*
los dedos	*se la puso*
mascaba chiclé	*colillas*
se seguía	
odiaba	
amenazador	
gritó	

Key: Items in bold indicate those where the learner model corresponds to the learner's own self-assessment of the item.

Total correspondence is 15 out of the 28 items (53%). Of the non-corresponding items, 10 (36%) represent under-rating of the learner's knowledge by the model. This reflects the gap between the model's quantitative data compared with the

subject's own judgement. The remaining 3 items are those where the model has over-rated the subject's knowledge. This could be accounted for if the quality of processing done on these 3 items was inferior in some way to that done on the other items. Inferior quality of processing has been previously identified with surface learning, so the degree of over-rating of the subject by the learner model could be taken as an indication of the extent of that subject's surface learning, ie: given the amount of processing they have done, they should feel that they know it better than they do. The small percentage of over-rating in S14's case is consistent with the fact that he has been identified as a predominantly deep learner (section 4.6). We have also seen that he came close to the target rate for retrieval, was a target achiever, and adopted a viable grouping strategy (section 3.2.2). If the same kind of analysis is applied to the rest of the Subjects (with the exception of S10, whose results are commented on above), then the extent to which this over-rating of surface is general, can be shown. Fig 5.1 represents the situation graphically, illustrating the model's tendency to both over- and under-rate subjects.

Fig 5.1: Correspondence Between Self-Assessment & Learner Model

% Under-rating by Model (ie: quality of processing high enough for a small amount to lead to confident knowledge = deep approach)

% Over-rating by Model (ie: a lot of processing has not resulted in confident knowledge = surface approach)
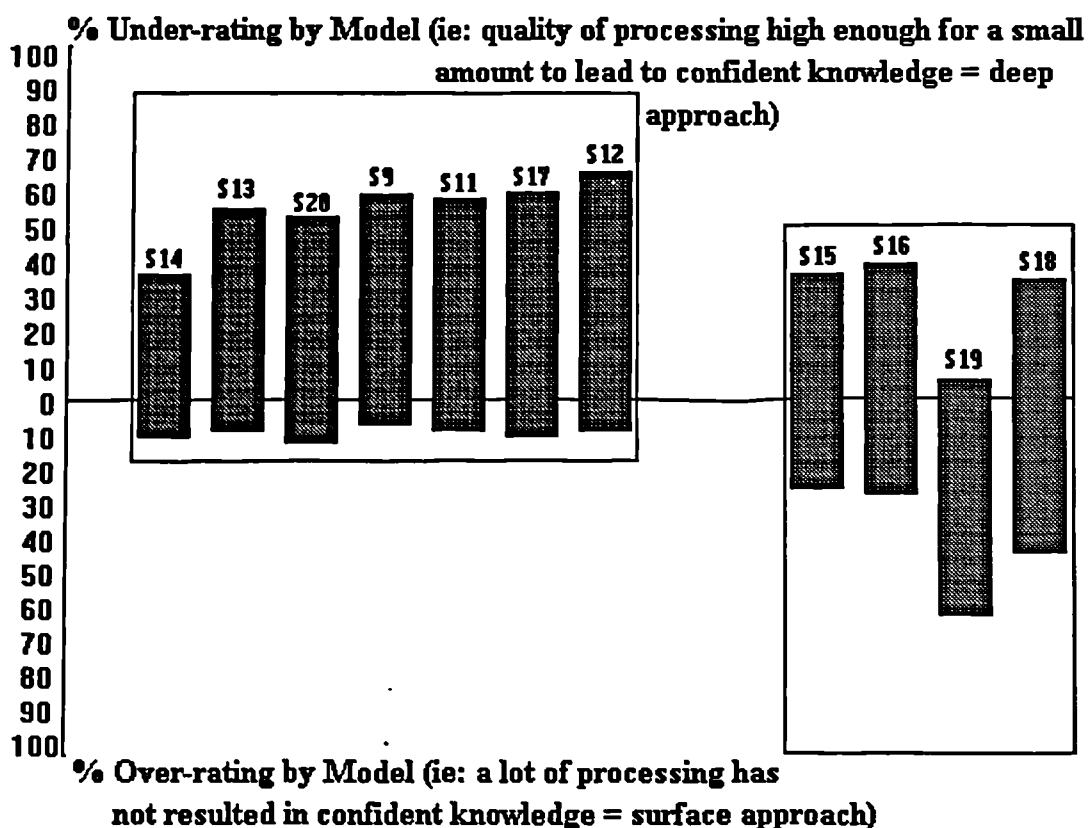
The amount of overall correspondence between learner and model can be judged by the size of the bars in this figure (the smaller the bar the more correspondence). It is notably low, mainly because of the importance that the model attaches to grouping in its assessment, and the generally low level of this kind of processing done by these learners. However, it is evident that the amount of over-rating is considerably lower for the more successful (either in terms of retrieval rate or target achievement) subjects (S14, S13, S20, S9, S11, S17, S12, for whom it averages 7.6%), than it is for the less successful ones (S15, S16, S19, S18, with an average of 38.7%).

As we have seen (Section 4) that better learning outcomes are generally associated with deep learning approaches, we can conclude that the comparison between the system's learner model, and the learner's own assessment of their knowledge, effectively identifies the approach, in the sense that if it under-rates knowledge compared to the learner's self-assessment, then the learning approach is likely to have been deep. The learner model is therefore, at least in co-operation with the assessment provided by the learner, partially confirmed as a representation of the state of the learner's mental lexicon. There is also reason to believe that its relevance would increase if and when the learner's processing came more into line with the strategy assumed by the system design. It is nevertheless important to note that this model is not able to represent the learner on its own, and that it requires the subjective self-assessment data in order to be meaningful. We may conclude from this that some form of self-assessment of the learning outcome should be supported by the system design, as part of the overall interaction.

5.2 Summary of Section 5

Sequencing of items for testing is based on a quantitative model of the processing the learner has done. This model identifies items which are less likely to be retrieved, and those which are likely to activate larger areas of the learner's mental lexicon. A comparison of values ascribed to items in the model with subjective judgements by subjects about their knowledge of those items, suggests that the learner model consistently over-rates the knowledge of subjects who have been identified as adopting surface approaches. Although the model alone is not able to

identify such subjects, the comparison suggests that its representation is more relevant to the quality of the knowledge of subjects who have adopted deep approaches, and that the test-sequences based on it are thus more likely to contribute to learning processes for these learners.

Section 6: Summary of Main Test Programme

This chapter set out to address the questions whether a target learning-rate of 8 words per hour is feasible with this system, whether the interaction supports the development of deep learning approaches, whether the quality of the learners' approach relates to their performance outcome, and whether the system's learner model reflects the learner's approach and the learning outcome. A programme of tests was reported on, in which 12 subjects were observed and recorded, and their interaction data and introspective comments analysed.

6.1 Learning Rate

It was found that defining the learning rate in terms of the rate of retrieval of target-items was both practical and relevant to the quantitative requirements for vocabulary-learning discussed in previous chapters. In addition, it was shown that the target rate of 8 words per hour is feasible, although 'target oriented' learners who focus on objectives such as achieving correct retrieval of all target items, may do so at the expense of learning rate. 'Process-oriented' learners who focus on the learning rate, may do so, on the other hand, at the expense of qualitative aspects of processing, which has consequences for the overall quality of the learning outcome. The learners who were studying Spanish at an academic institution achieved a uniformly higher learning rate than those who were trying to improve their English in situ, for vocational purposes. This may have been due to 2 factors:

i) The SFL group were working in an academic environment, and had a generally higher educational level and more familiarity with text-based tasks. This led them to identify the goals of the program more quickly, and to develop strategies to achieve them. Their higher levels of general reading skills may have given them the

confidence to adopt a more casual approach to both comprehension of the texts and the use of the lexical tools.

ii) The availability, for the SFL group, of a bi-lingual dictionary gave them faster access to a definition for their target-items. Although this resulted in a certain amount of surface processing at the lexicon-building stage, it was more functional for the quantitative goals of the program than was the additional work that the ESL group had to do to decode the English definitions given in their mono-lingual dictionary.

Both factors taken together suggest that the time taken by individual learners to 'work up' to their most efficient level of processing with this program may vary considerably, depending on educational background and familiarity with the particular lexical resources being used. To ease the problem for non-academic learners using mono-lingual or non-learners' dictionaries, it is probably better, at least initially, to provide source texts that are well within their normal reading range and contain fewer new words, rather than more difficult ones which ensure a rich source of new vocabulary. Given a longer and easier induction into the use of the program, it may well be that even lower-level learners could achieve learning rates comparable with the target rate proposed in this study.

6.2 The Grouping Strategy

The grouping strategy did not contribute as much as was expected, for most subjects. This was due to low quantity as well as quality of group-related processing. Practical imperatives involved in selection, lookup and testing, tended to take precedence over the more reflective grouping activities. In addition, the 25-word threshold at which grouping becomes critical to the system's support of learning objectives, was not reached by most subjects. Effective grouping is not intuitive for all learners, and some wasted time with vague and over-general classifications which could not have been expected to assist their retrieval of the items concerned. In addition, the failure of most subjects to use the KWIC tool to

explore collocational features of their target words meant that a whole range of possible syntagmatic associations went unexploited. Again, it is the non-intuitive nature of KWIC output which is responsible, and the program clearly needs some way to support the learner in developing an understanding of what the tool can offer and how to use it. Nevertheless, the subjects who did some useful global or local grouping, tended also to be the ones who adopted predominantly deep approaches, and who maximised their target achievement. This suggests that the grouping strategy is related to a good quality learning outcome, and that this could become more so over a longer period of time, as the effects of global grouping became more marked. This strategy is also not incompatible with a higher learning rate, as at least one of the subjects showed. The lessons from the failure of many subjects to engage in or profit from grouping in these tests is that grouping needs to become critical for learning objectives at an earlier stage, and in a more comprehensible form. Learners need support (suggestions, examples) in generating useful groups, such as examples of semantically decomposed concepts and sense relations, and this must be an objective for future developments of the design. The incidence of selection of the same words and creation of similar groupings indicates that a tutorial approach to developing grouping skills may be feasible.

6.3 Deep and Surface Approaches

The deep/surface distinction in learning approach was found to be applicable to each stage of the interaction, and to subjects' general approach to vocabulary learning, and their subsequent evaluation of their knowledge of the target words. It was found that those identified as adopting predominantly deep approaches, were in general the most successful in performance terms. This was considered to validate the overall design of the program, as it represents an external measure of quality in both learning process and outcome, in which the criteria for assessment could be confirmed by independent judges. Aspects of both deep and surface orientation were present in the approach of most subjects, and the design appears to support successful performance from both perspectives, which suggests that it is to some extent adaptable to the individual learning styles of users. The degree to which it

would go on supporting successful surface learning, however, would probably diminish as the target-item list became larger, because the lack of significant associations within the lexicon would mean that as the time lapse between practices of items grew bigger, so would the likelihood of forgetting them. In other studies of this program in use, not reported here, the target-item list has risen to as many as 120 items in a 2-month period. At that size of lexicon, the average practice session will only cover about 20% of items. Although the sequencing strategy ensures that the less active items are recycled first, and that in time all items will eventually be reviewed, time lapses of weeks may occur before earlier items return. It is hard to imagine that associations at the surface level would persist long enough for the retrieval practice effect to be able to operate in such cases.

6.4 Learner Model

The learner model was found to reflect performance in the sense that it records processing in a form which bears some relation to the learning outcome as assessed by the learners themselves, but it is not able to distinguish between successful and unsuccessful learners per se. It was shown to be more likely to over-estimate the self-assessed knowledge of subjects who had adopted predominantly surface approaches, and therefore to be in accordance with the assumptions of the general design (ie: that deep approaches equal successful processing). It was thus considered to be effective in adapting the interaction to the performance of learners, although ultimately dependent on other forms of modelling based on subjective assessment by the learners themselves.

6.5 Design Implications

Considering that all these subjects began as novices with the program, and that the circumstances of the test programme were far from ideal (especially for the ESL group), it is not unreasonable to suggest that with more practice and a better understanding of the aims of the interaction, most learners could achieve the target rate and the quality of processing which guarantees a satisfactory learning outcome. However, there are a number of ways in which the design needs to be improved:

The design has been shown to support at least 2 different strategies (target-oriented and process-oriented), and to support meaningful ways of measuring learner success in meeting the respective objectives of these strategies (learning/ retrieval rate and target achievement). The importance of maintaining a balance between the 2 strategies, and the apparent connection between target orientation, grouping, and a deep approach, suggests that these 3 criteria should always be available to be contrasted with the current learning rate, for the purpose of guiding the interaction. This means that the system should provide users with a range of scores, etc. which reflect their progress in a wider sense.

The design does not adequately support learners in developing their awareness of lexical structure through the grouping activity. There are two ways in which this could be improved. Firstly, some attention should be pre-paid to the nature of the originating texts and the kinds of optimal associations they contain. It should be possible to make these associations explicit, where necessary, to the learner. This implies an element of pre-selection of text, prediction of likely target items, and addition of tutorial material. There is therefore a trade-off to be made between total learner freedom to choose texts and target words, and the availability of support for what is an essential element of the interaction. This support should focus not only on the paradigmatic semantic relations which are discoverable from the dictionary and thesaurus, but also on collocational relations revealed by use of the concordancer. The latter implies a strong tutorial emphasis on an aspect of processing which the learners in this test programme generally failed to exploit. Secondly, the structure of the interaction should be focused more tightly and explicitly on the purpose and mnemonic benefits of grouping. Initial categorisation of items should be directly related to the kinds of prompts and clues which will be available to help retrieve them, so that the learner makes an explicit choice (rather than a default one) about how they wish to store and eventually retrieve items.

The role of the learner model in helping to determine optimal sequencing of items for testing may be also be enhanced. Firstly, its current operation could be refined

by incorporating self-assessment of the kind used off-line in the test programme, into the interaction. Secondly it could be developed to take into account the learner's strategic profile as target- or process-oriented, and used to help generate tutorial suggestions intended to guide the interaction towards a proper balance.

CHAPTER SIX: SUMMARY AND CONCLUSIONS

This chapter contains a summary of the previous five chapters, covering the background to the research, theory of the design, methodology of evaluation, and empirical findings. Following this, there is a discussion of the implications of the work from the point of view of its contribution to knowledge in the areas of: understanding and characterisation of vocabulary learning processes, theoretical and implementational issues in the design of Computer-Aided Language Learning for vocabulary, and further research into language learning tools and environments.

Section One: Summary

The introduction described how research questions relating to the design of CALL for vocabulary arose out of attempts to develop programs for learners of English as a Foreign Language. These learners, characterised as adult, well-motivated non-beginners, were considered to need support in a learning situation which required productive, self-directed word-learning, in the absence of human tutorial assistance. A CALL design, based on current lexically-oriented teaching theory and methodology, was implemented to support self-access activities aimed at developing new vocabulary items to a sufficient level of internalisation to make them accessible, subsequently, for productive practice in a communicative environment. When the design failed to generate the kind of interaction expected, a number of questions were posed regarding its theoretical and functional principles. These questions formed the basis for the research described in this thesis. The intention of this work has thus been to: a) establish a theoretical justification for the assumption that CALL can support self-directed vocabulary learning, b) develop principles for the design of a program to be used in this way, and c) establish a methodology for its evaluation.

In Chapter One, a review of the literature in the fields of linguistics, second-language acquisition and second-language learning was undertaken in order to establish the nature of the content in L2 vocabulary acquisition, and the processing

requirements that conscious learning implies. This review concluded that there are both quantitative and qualitative objectives in the task, and that vocabulary knowledge involves some understanding of the structure which exists in the relations that individual words have to each other, in terms of their meanings and the ways they typically combine. It was determined that both contextualised and decontextualised processing are implicated in this knowledge, that there is a definable 'threshold level' of word knowledge appropriate to the learning aims earlier described. It was also concluded that the size of the learning task imposes a need for a minimum learning rate, and an opportunistic approach to the identification of target items.

Chapter Two examined the state-of-the-art in CALL for vocabulary, seeking principles for the design of a system which would address these requirements. The investigation revealed that vocabulary-learning systems generally tend to focus either on the selection, or on the retrieval, of target items, but not on the processing that goes on in between. Design features relevant to the incorporation of new items into existing conceptual structures ('lexicon building'), or to the explicit use of structure to make items memorable, were found in only a few systems. These tended either to base their procedures on simple, and often inflexible, models of human cognition, or else around open, and unguided, use of lexical tools. Few systems addressed issues of the adaptability learners need, or the nature of the performance vocabulary learning implies. The review concluded that the most relevant design principles, for the learning aims stated in Chapter One, were those which addressed: learner exploration of rich textual context, lexical search and the use of information resources, productive retrieval based on cloze-type prompting, and adaptable system-structured presentation and review, based on a coherent learner model. It was concluded that the way is open for an approach to CALL-for -vocabulary design which capitalises on theories of the structure of the mental lexicon, utilises on-line tools such as dictionaries and concordancers, and supports individualised and study-based processing of target items, via adaptability to the learner in aspects of the interaction concerned with testing and feedback.

In Chapter Three, system design specifications, derived from the requirements identified in the preceding chapters, were outlined. These described learning activities based around the selection of items from a rich textual context, the organisation of them into groups according to common lexical-structural features, and the retrieving of them from memory, using the same groups and features as prompts. A theoretical psycholinguistic basis for the characterisation of lexicon-building processes was established, and used as the basis for a model of the learner's performance, which could be implemented to adapt the sequence of testing of target items to the individual learner. Some issues of the use of computers to collect interaction data for the purpose of evaluation were discussed, and the need for an additional source of subjective data from learners themselves was proposed. A psycholinguistic basis for the characterisation of lexicon-building processes was established, and specifications derived for a learning activity based around the selection of items from a rich textual context, the organisation of them into groups according to common lexical-structural features, and the retrieving of them from memory, using the same groups and features as prompts. The kinds of approach learners might adopt to this activity were discussed, and the possible effects of these on process and outcome considered. A theoretical framework for describing these approaches, based on a distinction between 'deep' or structure-oriented, and 'surface' or feature-oriented learner perception of the learning content, was described. Requirements for a methodology of evaluation were considered, focusing on the need to assess whether this interaction supports a rich learning experience for the learner, whether the learning content accords with the system's objectives, whether it encourages a deep learning approach, and whether it supports an optimal learning rate. These considerations were then summarised as a set of theoretical design principles for a proposed vocabulary-learning system, addressing individualisation of item-selection and lexicon-building processes, adaptive structured presentation and review of target items for retrieval, and the generation and recording of performance data for the purposes of qualitative and quantitative evaluation of the interaction the system supports.

In Chapter Four a functional architecture for the system was proposed, comprising three modules dealing with selection, lexicon-building, and retrieval respectively. Specifications were then drawn up for general interaction procedures, for the on-line information resources which are required to support lexicon-building, and for the learner model on which structured presentation of test items is based. These specifications were put into context in a description of a 'model interaction' outlining the way that a learner might use the system to achieve their learning objectives. The implemented system and the approach to evaluation were tested in two pilot studies carried out with English learners of Spanish at different levels. These studies confirmed that the interaction is capable of meeting the learning experience and learning content requirements. They also established a framework for the qualitative evaluation of the processing that learners did at the lexicon-building stage, confirmed the initial relevance of the deep/surface analysis of learning approach, and provided an empirical basis for refinements to the design aimed at promoting the grouping strategy, encouraging the development of deep learning, and optimising the potential learning rate.

In Chapter Five a programme of tests, using learners of Spanish and of English, was described, and the results discussed. Findings supported the assumption that learning was the outcome of processing at all three stages in the interaction, and that quality in the learner's approach at selection and lexicon-building stages was associated with successful performance at retrieval. It was established that the target rate for vocabulary learning with this system, of 8 words per hour, in accordance with the literature, was feasible, but also that this was unlikely to be achieved in their first 4-5 hours of using the system. A distinction was therefore made between strategies of target-orientation, in which the learner focused on retrieving all of their target items, and process-orientation, in which they were concerned to maximise the number of items in the target list. These strategies were shown to be relevant to the overall learning outcome, in that a good balance between them was associated with measures of quality in the approach, assessed according to the deep/surface framework. The potential relevance of the grouping strategy to quality in the learning outcome was also established, but it was shown

that the strategy is not intuitive and may thus be non-functional for some learners, if unsupported by tutorial help, in the sense that it may slow down their processing in other respects without providing additional benefits. The level of target-item selection which the design assumed, was too high for these novice users and only two of them reached the point at which the refinements intended to promote the grouping strategy, introduced as a result of the pilot studies, became effective. It was nevertheless shown that item grouping as supported by the system is flexible in the kinds of use to which learners can put it, and that lexicon-building strategies not predicted by the design are possible. Deep and surface approaches were identified in the ways that learners viewed the vocabulary-learning task, how they selected their target items, how they used information from the on-line resources, in their grouping, and in their subjective assessment of what they had learned. Deep approaches were observed in some of the learners, at both synthetic (how they went about comprehending the source texts) and analytic (how they analysed the decontextualised items) levels, and these learners were shown to be the ones who performed best in terms of either target-achievement or learning rate. The learner model, based on a record of processing throughout the interaction and used to determine the sequencing of target items for presentation at the retrieval stage, was shown to better reflect the self-assessment of the learners who adopted deep approaches, in that it did not over-estimate their knowledge as much as it did for those with surface approaches.

The following implications for the design of the system were suggested:

i) The global learning goals of the strategy, relating to the overall learning rate and the optimal balance between target- and process-orientation., should be made explicit to the learner. Feedback could be given in the form of performance measures reflecting the relation between the distribution of time in the various modules and the different categories of outcome (items selected, groups generated, successful retrievals).

ii) Tutorial support is required for the activities of: categorisation of lexical features of target-items, creation of useful semantic groups, and the use of the KWIC concordancer to explore collocation and other syntagmatic features. This could take the form of 'demonstration' procedures or menus of options making clear the kinds of lexical relations which the learner might look for among their target-items.

iii) The lexical tools should be adapted to the requirements of the interaction. They should be accessible simultaneously and in parallel with all other on-screen information relating to lexicon-building. Dictionaries should reflect the network structure of the lexicon, emphasising sense-relations rather than giving extended definitions or examples. Concordancers should be interactive, highlighting statistically significant co-occurrences and allowing elements in the citations to be selected and transferred into notes or groups.

iv) The self-assessment activity should be incorporated into the interaction, providing information for the elaboration of the learner model so that it can more accurately reflect qualitative as well as quantitative aspects of the learner's performance. Surface approaches, which are unlikely to support significant quantitative gains in vocabulary-knowledge over time, could then be diagnosed and the learner's attention drawn to ways in which the quality of their processing might be enhanced.

It was concluded from the results of this test programme that the design does support an autonomous learning process which, provided the learning approach is appropriate, enables learners to acquire target language vocabulary at the threshold level of production, at a rate which is commensurate with the overall size of the learning task. Although the development of optimal strategies in the use of the program did not proceed as expected, for many of the subjects, it was nevertheless considered that further experience working with the system would be likely to result in improvement in performance for most of them. The extent of surface learning approaches, however, and their effect on learning outcome and the system's

modelling, demonstrate that a CALL design, if it is entirely learner-directed, cannot presume that effective learning will result. Where the learner's approach is predominantly surface, distortions will be present in the outcome to the interaction and the system's adaptation to the learner will be inappropriate. Surface learning approaches thus undermine the attempt at rationalist design and modelling, and the need for an instructional element is clear.

Section 2: Characterisation of Vocabulary-Learning

Vocabulary learning has been identified, in this research, with a 3-stage process involving selection, incorporation (or lexicon-building), and retrieval. Acquisition of new vocabulary is thus seen as the outcome of these three types of processing combined, not just as the product of one of them (eg: the understanding of a text, or the finding of a translation in a dictionary, or the 'getting right' of an exercise). Variation in this outcome is inevitable, both in the quality of acquired knowledge for the individual learner, and between different learners, and two key aspects of variation are to be found in the rate of learning and in the degree of development of lexical-structural awareness. Rate has been shown to be meaningful in theoretical terms (ie: there is a need for a minimum rate at which new items can be cycled through the three processing stages, in order that the L2 lexicon can grow to a size comparable with that of the L1, in a time consistent with the amount of study time feasible for an adult second-language learner). It has also been seen to have practical implications in terms of learner strategy, in the need for a balance between targeting particular items for practice, and activating whole areas of the mental network in order to maximise the effects of association and activation. Lexical-structural awareness involves the learner in evaluating word forms from two points of view: the contribution they make to meaning in the overall context, and the nature of their structural features. These studies suggest that the first way of looking at new vocabulary is the more intuitive; learners are predominantly concerned with reading and understanding, and vocabulary does not become an issue until there is a breakdown in comprehension. However, it was a premise that the second way of looking at new items is necessary if these items are intended for

productive learning. The relatively non-intuitive processes involved in analysing text for lexical features, coupled with the historical neglect of the lexical level of description in linguistics, has made this a more demanding task, but principles are available at both paradigmatic and syntagmatic levels, and the application of these principles to the identification and description of target items is an important step in the development of an effective conscious learning strategy. Structural awareness thus plays a role in learners' apprehension of learning content, but, unlike some semantic relations (sense relations, relations of semantic composition), collocations and lexical phrases are not yet well-understood in psychological terms (see Knowles, forthcoming), and the structure which is inherent in this way of looking at language (eg: patterns of distribution and frequency) is not intuitively graspable by learners. Now that tools for identifying occurrence-related phenomena exist (concordancers etc.), and theories of lexical distribution are looking more at discourse and register levels (eg:Hooey 1991), we can expect the methodologies to be developed which are necessary to make this level of analysis explicit and more easily learnable. Much work is required, however, into ways to help learners to characterise these lexical phenomena, to describe types of collocation and lexical phrase, to understand systems of cohesion and the lexical choices which result. Also to develop learning strategies which are geared to the incorporation of larger chunks of language into their mental store.

We have seen that the interaction extends along a continuum from context-oriented, where the identification of a word as a target vocabulary item occurs when overall comprehension of the message of the text breaks down, because of its unfamiliarity and opacity (ie: the word is seen as causing loss of sense in the text), to lexis-oriented, where the word becoming a target item depends on it having features which are familiar from other contexts (ie: it is seen as having possibilities for making sense in itself). This affects the way that a learner reads a text (eg: ignoring vocabulary until the extraction of gist becomes impossible, or else allowing gist to remain vague whilst the roles of salient vocabulary items are assessed). It also affects the way they process target vocabulary once it has been identified (ie: putting effort into elaborating the meaning of the original context, or

into establishing patterns of association out of context). Because the first approach is more focused, and has a clear end once understanding of gist has been restored, it is faster. A further advantage is that it is more likely to support internalisation of collocations and/or lexical phrases. However, the extent of its context-dependence means that the resulting knowlege of the target items themselves may not support the generation of new contexts for production. The evidence from this study, suggests that both retrieval and self-assessment are weakened if reliance has been on context alone. The approach in fact favours learners who rely on implicit aspects of their linguistic knowledge, such as the reinforcement of syntagmatic associations from auditory memory (ie: they remember collocations and phrases by sound, like rhymes or chants). The second type of processing takes longer, involves learning to use resources such as thesauri etc. and is not so clearly bounded (a chain of semantic association may have no logical end). It's main advantages are in reinforcing awareness of word composition and thus assisting in the development of the ability to generate new words, which as Aitchison (op.cit) has pointed out, is an important aspect of lexical knowledge, and in revealing the semantic structure of the target language which may be different from that of the L1. This approach favours learners who adopt a conceptual, and/or formal analysis. We therefore have a distinction, in terms of characteristics of vocabulary-learning, which corresponds to Bialystok's (1988) 'automatic' and 'analysed' dimensions of language proficiency.

Production has been a central issue throughout this study, because it is a defining factor in the threshold level of word-knowledge. Determining the objective of the computer-based strategy as the threshold level is consistent with the process definition of learning, in which the third stage involves productive retrieval, but the necessary conditions for retrieval may vary considerably, from a specific L1 trigger, to a broader range of conditions such as the presence of a particular semantic field, and/or a familiar collocate, and/or a syntactical condition. The ultimate purpose of the strategy is to create a lexicon which will be available when the opportunity for full communicative practice occurs, and the need for retrieval is unpredictable. How the vocabulary acquired in this way should be consolidated

for use is therefore an important question. The conscious intention, on the part of the learner, to get items into use, is important, but the eventual objective is the unconscious triggering of appropriate items by the linguistic and communicative environment. Thus, whilst incidental exposure during communication will play a role in the consolidation of words which are at threshold level, the learner needs to 'stack the odds' as far as possible, by contriving practice wherever this is feasible. An understanding of the role that discourse register plays, in determining the likelihood of word occurrence, is therefore important, as the leaner should be able to seek out texts and situations in which certain areas of vocabulary can be expected. In addition, the learner needs a continuing awareness of the identity of items which are difficult to use, or which may have slipped in some way. The problem this presents, as the learner's vocabulary expands towards many thousands of items, is obvious, but continued use of this program should go some way towards overcoming this, because of the way the program continually re-sequences items for practice, on the basis of the spread of processing levels across the whole target list. This should ensure that all items will eventually be re-presented, even where they have slipped from conscious memory. The incorporation of on-line self-assessment could also be expected to enhance the learner's awareness of the state of their target lexicon, and to reflect on the opportunities they have (or have not) found to get words into genuine practice. Finally, it may be expected that the development of structural awareness, expressed as enhanced skill in classifying items and detecting new types of relation between them, will help to raise general levels of activation in the learner's network as a whole, favouring the more automatic kind of lexical decision postulated by the psycholinguistic model.

The issue of the relation between cognitive models of linguistic processing and the design of vocabulary-learning systems, has also been addressed in this study. The psycholinguistic model of speech production and the mental lexicon, which was used as a basis for the design of the interaction (see Chapter Three), postulates a non-deterministic mechanism for producing words, ie: competing activation levels and local conditions in the mental network combine to generate several candidates for production, rather than specifying a single one. This model was useful in that it supported the general intention to deal with the target lexicon as an integrated

whole, (required by the assessment of the quantitative 'size' of the learning task). The main implication of this, for the interaction, was that learners would be encouraged to produce the words that came to mind, with less concern for the kind of 'correctness' that conventional vocabulary-learning emphasises. However, the 'target orientation' of some learners showed that they preferred to be engaged in 'word-search' rather than 'word-generation', and that they were adapting their unconscious processing mechanisms to targeting specific items in their lexicons, in order to confirm to themselves that these items could be produced at will. Incidental effects of some of these attempts, (slips, blends) indicate that the underlying mechanisms may well be the same ones as operate in continuous speech production, but the fact that they are under conscious control means that the way they operate is different. We therefore need to study conscious vocabulary-learning in its own terms, not simply as a way of expediting assumed unconscious processes of acquisition. The different conscious ways that human beings are able to use their 'word-producing-equipment' indicate that underlying psycholinguistic processes are extremely flexible. Rather than postulating more and more detailed models of unconscious mechanisms, we should perhaps be focusing on what learners think about while they are manipulating lexical data with a learning objective in mind.

Section 3: The Design of CALL for Vocabulary

Sussex et al (1994) suggest that the key to effective learning in a partly unstructured computer-based environment lies in the arrangement and implementation of tools and activities to support learning (op.cit, p.140). This research has demonstrated that learner approach is equally central to effective learning, and that, as well as a theory of the cognitive mechanisms that lexical learning is assumed to involve, CALL programs should also embody a theory of conscious learning, ie: of the effect that different approaches to the task are likely to have on the interaction, and on the learner's performance, and should incorporate effective means of diagnosing learner approach at different stages of the interaction, and promoting deep learning wherever appropriate. Whilst

190

principles of learner-directed learning and the development of learning skills, particularly in the use of lexical tools and information resources, are very important and necessary to the opportunism that extensive vocabulary acquisition depends on, we must not lose sight of the learner's potential to undermine their own efforts by adopting an approach which focuses on superficial aspects of the learning content. Tutorial guidance is necessary for some learners, to help them locate the 'deep path' to optimal performance. This is particularly true for lexicon-building activities, those intended to support the creation of mental links between items sharing some underlying structural feature. Research in CALL design needs to concentrate on principles for providing guidance in the use of dictionaries, thesauri, concordancers and other lexical tools to identify these features, and on means of adapting an on-going interaction to the general aim of weaning learners away from surface approaches. Some areas where such research is on-going have been identified, for example Hayet's (op.cit) work on the development of an associational dictionary for French, Johns' (1994, op.cit) development of teaching material for his Contexts programme, the work of Cumming et al. (op.cit) on learning and teaching strategies in the use of lexical tools. Other areas are suggested by the findings of this study: collocation, for example was left unexplored by most of the subjects in the test programme, and it would be useful to establish whether automatic 'background' searching through a more extensive corpus could be used to identify other items in the co-text of a selected item which were significant collocates of it, and the information presented to the learner. Word grouping also, does not need to be left to the learner's intuition, as preparation of source texts could predict, to some extent, the semantic fields which the items most likely to be selected could be assigned to (see the test findings on the similarity of selected word list for the two groups of subjects). The titles of these fields could be selected from a database and proposed to the learner at lexicon-building time. The incorporation of learner self-assessment as a means of enriching a system's representation of the state of their knowledge, is another important principle. For vocabulary it can be done reasonably simply, as the test programme showed, although the consistency and reliability of the information may vary amongst individual learners and there is research to be done on the ways

that learners develop and use their ability to judge the extent of their vocabulary-knowledge. Such a representation is necessary if a CALL system is to adapt to the requirements of individual learners and base its tutorial intervention on assumptions about the actual learning processes that have taken place, rather than just the apparent end-product of them.

In general, CALL-for-vocabulary needs to take on board its role as a contributor to the practice of instructed second language acquisition as described by Ellis (1992, p.195). This theory, although defined with classroom-learning in mind, contains two important hypotheses that are relevant to autonomous learning in a one-to-one learner-computer interaction, ie:

- Implicit and explicit knowledge constitute different kinds of knowledge; explicit knowledge, which is largely form-focused, enables learners to 'notice' L2 features in meaning-focused input which would otherwise be ignored. The computer's essence as a manipulator and purveyor of formal representations fits it particularly to the processing of explicit knowledge, and the instructional role that this enables should be part of the design of its interaction with a learner.

- Instructional input is 'filtered' by learners in accordance with their learning style; they will be predisposed to attend to linguistic features in the input if they respond positively (cognitively and affectively) to the type of instruction. The findings of this study show that the interpretation of input is also influenced by the depth of the approach to learning, which is in part determined by the design of the computer-based task itself. The computer's ability to model the actions of the learner should be used to adapt the task to favour a positive cognitive response.

Whilst it is true that the value of the 'tutor metaphor' as a principle for the design of computer-based learning has been generally discredited during the last two decades of research in intelligent CAL (Hartog 1989), the 'information-resource' paradigm which has replaced it is also deficient for learning purposes. This study has shown that the need for CALL-for-vocabulary to adopt a principled approach to

providing tutorial support for the use of lexical information resources to aid vocabulary-learning is as paramount as ever.

Section 4: Conclusion and Further Research

The work described here falls into the general category of research described by Sussex et al. as being concerned with "..environment(s) for learning, and tool(s) for investigating learning and teaching.." (op.cit, p.145). The main finding is that conscious vocabulary learning, to a threshold level of word production and at an optimum rate for L2, can be achieved without human tutorial assistance via a computer-based strategy in which support is provided for item selection, lexicon-building and productive retrieval activity. On the assumption that better understanding of the principles of the interaction will enable us to design richer and more supportive environments, which will enable learners to extend their learning rates and deepen the level of their word knowledge, it is proposed that further research with the current design could focus on the following general areas:

i) How well does this interaction work in the medium/long-term? Data could be collected on the learning rates, notes/groups-output and lexicon-sizes achieved by learners over a period corresponding to an accepted term of study in the relevant target language. This could then be used to evaluate the design in terms of conventional or institutional vocabulary-learning objectives.

ii) How does this kind of vocabulary-learning transfer into communicative contexts? Investigations could be carried out into the use that learners make of their target-items, in productive contexts such as essay-writing, oral presentation and monitored discussion. This could be used to determine whether the quality of the learning supported by the design was sufficient to support an increase in the size of the learners' vocabulary at the fully-usable end of the knowledge scale.

iii) What governs the 'optimal' use of particular texts for the purpose of vocabulary-extraction? The differential effects on learning outcome of using a few

'difficult' or many 'easy' texts could be examined, as also could the effects of using one continuous piece of discourse, as opposed to many separate ones. Results might be used to design specific 'courses' with the program, intended to expose learners to a particular vocabulary register or lexical syllabus.

iv) How do learners differ in their ability to 'self-assess' their vocabulary knowledge? Independent testing mechanisms could be used to determine whether learners' judgements of their knowledge of words was based on the kind of processing offered by the program, or whether it was a constant of the way they regarded vocabulary (eg: they always over- or under-estimate their knowledge). This would be very useful in deciding how these assessments should be incorporated into the interaction.

In addition to these specific suggestions for further research, there are the major (and related) design issues involving the nature of tutorial support for learning about lexical structure, and the development of an intelligent learner model for lexical learning. This study has indicated some the requirements and suggested some possible approaches, but the size of these questions is well beyond its present scope.

Bibliography:

AITCHISON J. (1987): Words in The Mind. Blackwell

ARNOLD P. & BEJOINT H. eds. (1992): Vocabulary and Applied Linguistics. McMillan Academic & Professional.

ATKINSON R. & RAUGH M. (1975): An application of the Mnemonic Keyword Method to the Acquisition of a Russian Vocabulary. Journal of Experimental Psychology - Human Learning and Memory, 104, pp.125-133

BIALYSTOK E. (1988): Psycholinguistic Dimensions of Second Language Proficiency. In Rutherford & Sharwood-Smith (eds), Grammar and Second-Language Teaching, Newbury House.

BLOOMFIELD L. (1933): Language. Holt, Rinehart & Winston.

BLUM S. & LEVENSTON E. (1978): Lexical Simplification in Second Language Acquisition. Studies in Second Language Acquisition, Vol.2, No.2, pp.399-415

BOGURAEV B., & BRISCOE T. eds. (1989): Introduction - Computational Lexicography for Natural Language Processing. Longman. p.15

BORCHADT F. & JOHNSON E. eds. (1993): Proceedings of the Computer Assisted Learning and Instruction Consortium 1993 Annual Symposium on 'Assessment'. Duke University.

BROWN T. & PERRY F. (1991): A Comparison of Learning Strategies for ESL Vocabulary Acquisition. TESOL Quarterly, Vol.25, No.4, pp.655-670

BRUMFIT C., PHILLIPS M., & SKEHAN P. (1985): Computers in English Language Teaching: A View from the Classroom. Pergamon Press & The British Council.

CARTER R. & McCARTHY M. eds. (1988): Vocabulary and language Teaching. Longman.

CATT C. (1992): CALL Authoring Programs and Vocabulary Development Exercises. CALL Journal, Vol 4, Part 3.

CHANNELL J. (1987): Applying Semantic Theory to Vocabulary Teaching. English Language Teaching Journal, Vol.35, pp.115-122

CHANNELL J. (1988): Psycholinguistic Considerations in the Study of L2 Vocabulary Acquisition. In Carter & McCarthy (eds), pp.83-96

CHAPELLE C. & JAMIESON J. (1989): Research Trends in Computer-Assisted Language Learning. In M. Pennington (ed): Teaching Languages with Computers: The State of the Art. Athelstan. p54

CHAPELLE C. (1990): The Discourse of Computer-Assisted Language Learning: Toward a Context for Descriptive Research. TESOL Quarterly, Vol. 24 No.2 pp.199-225

CHAPELLE C. (1993): Issues in Computer-Assisted Analysis for One-Word Test Responses. In Borchadt & Johnson (eds), pp.28-32

CHOMSKY N. (1957): Syntactic Structures. Mouton.

CLARKE M. (1992): "Vocabulary Learning With and Without Computers - Some Thoughts on a Way Forward", Computer Assisted Language Learning, Vol.5.3, pp.139-146.

COHEN A. & APHEK E. (1981): Easifying Second language Learning. Studies in Second Language Acquisition, Vol 3, No.2.

COLLINS A. & LOFTUS e. (1975): A Spreading-Activation Theory of Semantic Processing. Psychological Review, Vol.82, No.6.

COWIE A. (1988): Stable and Creative Aspects of Vocabulary Use. In Carter & McCarthy (eds), pp.126-137

COWIE A. (1992): Multiword Lexical Units and Communicative Language Teaching. In Arnaud & Bejoint (eds), pp.1-12

CRAIK F. & LOCKHART R. (1972): Levels of Processing - A Framework for Memory Research. Journal of Verbal Learning and Verbal Behaviour 11, pp.671-684

CROW J. & QUIGLEY J. (1985): A Semantic Field Approach to Passive Vocabulary for Reading Comprehension. TESOL Quarterly 19, pp.495-513

CRUSE D. (1986): Lexical Semantics. C.U.P.

CRYSTAL D. (1971): Linguistics. Penguin.

CUMMING G., SUSSEX R., CROPP S. (1993): Learning English as a Second language - Towards the 'Mayday' Intelligent Educational System. Computers in Education, Vol.20 No.1 pp.119-126, Pergamon Press.

CUMMING G., SUSSEX R., CROPP S. (1994): The Teacher-Learner-Computer Triangle in CALL - Frameworks for Interaction and Advice. Computer Assisted Language Learning, Vol.7, No.2, pp.107-123.

DE BOT K. (1992): A Bilingual Production Model - Levelt's 'Speaking' Model Adapted. Applied Linguistics, Vol.13/1, pp.1-24

DECOO W. (1993a):-Lexical Composition and Morpho-Syntactic variation in Language Textbooks - Computer-Based Approaches as Another Dimension of CALL. Computer Assisted Language Learning, Vol. 6, No.2, pp.123-144

DECOO W. (1993b): "Customer Satisfaction in CALL - Queries from the Users and Our Final Impact", CALICO Journal 10.3, pp.50-64

DELL G. (1986): A Spreading-Activation Theory of Retrieval in Sentence Production. Psychological Review, Vol.93, No.3, pp.283-321

ELLIS N. (1995 forthcoming): The Psychology of Foreign Language Vocabulary Asquisition - Implications for CALL. International Journal of Computer-Assisted Language Learning (CALL), Special Issue on Cognitive Aspects of Second Language Acquisition, edited by D.Green & P. Meara.

ELLIS R. (1992): Instructed Second Language Acquisition. Basil Blackwell.

FAY D. & CUTLER A. (1977): Malapropisms and the Structure of the Mental Lexicon. Linguist Inquiry, Vol.8,Pt.3, pp.505-520

FILLMORE C. (1968): The Case for Case. In Bach & Harms (eds), Universals in Linguistic Theory, Holt, Rinehart & Winston.

FIRTH J. (1957): The Techniques of Semantics. In Papers in Linguistics, O.U.P.

FORSTER K. (1989): Basic Issues in Lexical Processing. In W. Marslen-Wilson (ed), Lexical Representation and Process, MIT, pp.75-107

FORSTER K. (1990): Lexical Processing. In Language - An Invitation to Cognitive Science Vol.1, (eds) D. Osherson & H. Lasnick, M.I.T. pp.95-131

FOX J. (1993): EC Research in Language Learning and IT - Some Experiences with LINGUA. Journal of Computer Assisted Learning, Vol 9, pp.100-106

FOX J., MATTHEWS A., MATTHEWS C., ROPE A. (1990): Educational Technology in Modern Language Learning. Report for the Training Agency by University of East Anglia and the Bell Educational Trust.

FRIES C. (1945): Teaching and Learning English as a Foreign Language. University of Michigan Press, Ann Arbor.

GAIRNS R. & REDMAN S. (1986): Working with Words. C.U.P.

GARRETT N. (1991): Where Do we Go from Here - And Who is leading the Way? ReCALL Journal, No.5.

GARRETT N. (1993): Integrating Different Approaches to Assessment. In Borchadt & Johnson (eds),p.62

GILLESPIE J. & GRAY G. (1992): "Hypercard and the Development of Translation and Vocabulary Skills", Computer Assisted Language Learning, Vol.5, 1 & 2, pp.3-11.

GOODFELLOW R. (1993): CALL for Vocabulary: Requirements, Theory and Design. Computer Assisted Language Learning, Vol.6,2, pp.99-122

GOULDEN R., NATION P., READ J. (1990): How Large can a Receptive Vocabulary be? Applied Linguistics 11, pp.341-363

GUILLOT M-N, KENNING M-M. (1994): Electronic Monolingual Dictionaries as Language Learning Aids - A case Study. In Thompson & Chesters (eds). pp.63-73

HALLIDAY M. (1966): Lexis as a Linguistic Level. In Bazell, Catford, Halliday & Robins (eds), In Memory of J.R. Firth, Longman, London.

HALLIDAY M. (1978): Language as Social Semiotic: The Social Interpretation of Language and Meaning. Edward Arnold, London.

HALLIDAY M., McINTOSH A. & STREVENS P. (1964): The Linguistic Sciences and Language Teaching. Longman.

HARLAND M. (1990): Hypercard - From Computer Dictionaries to On-Line Courseware Tools. Literary and Linguistic Computing, Vol.5, No.1, pp.89-93

HARTOG R. (1989): Computer-Assisted Learning - from Process Control Paradigm to Information Resource Paradigm. Journal of Microcomputer Applications, No.12, pp.15-31

HAYET M. (1994): A French Learner's Workbench. In Thompson & Chesters (eds).pp.97-106.

HENNING G. (1973): Remembering Foreign Language Vocabulary - Acoustic and Semantic Parameters. Language Learning, Vol.23, No.2, pp.185-196

HIGGINS J. & JOHNS T. (1984): Computers in Language Learning. Collins Educational.

HOEY M. (1991): Patterns of Lexis in Text. OUP.

HORTON E., ELLIS D. & BLACK P. (1990): The Design and Development of an Intelligent Tutoring System for Adult Literacy Students. Computer Assisted Language Learning, Vol.2, pp.69-81

HUDSON R. (1990): English Word Grammar. Blackwell, Oxford.

HUDSON W. (1989): Semantic Theory and L2 Lexical Development. In Gass & Schachter (eds), Linguistic Perspectives on Second Language Acquisition. C.U.P. pp.222-238

HULSTIJN J. (1989): Experiments with Semi-Artificial Input in Second-Language Acquisition Research. In Hammarberg (ed), Language Learning and Learner Language, Vol.8, pp.20-40, University of Stockholm.

HULSTIJN J. (1992): Retention of Inferred and Given Word Meanings - Experiments in Incidental Vocabulary Learning. In Arnaud & Bejoint (eds), pp.113-125

JOHNS T. (1986): Micro-Concord - A Language-Learner's Research Tool. System Vol.14, Pt.2, pp.151-162

JOHNS T. (1988): Whence and Whither Classroom Concordancing? In Bongaerts et al. (eds), Computer Applications in Language Learning. Foris.

JOHNS T. (1991a): Should you be Persuaded - Two Examples of Data-Driven Learning. ELR Journal Vol.4, 1-13

JOHNS T. (1991b): From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-Driven Learning. ELR Journal Vol.4, 27-45

JOHNS T. (1992): Personal Communciation.

JOHNS T. (1994): Contexts - The Background, Development and Trialling of a Concordance-Based CALL Program. TACT Conference, Lancaster University.

JUDD E. (1978): Vocabulary Teaching and TESOL - A Need for Reevaluation of Fixed Assumptions. TESOL Quarterly,Vol 12, No 1, pp 71-76

JUNG U. (1987): Evaluating Microcomputer Software - The State of the Art. In L. Legenhausen & D. Wolfe (eds): CALL and Innovative EFL Methodology, University of Augsberg, p42

JUNG U. (1988): Introduction, or a 'Petit Tour' of the World of CALL. In Jung (ed), Computers in Applied Linguistics and Language Teaching, Verlag Peter Lang, pp.1-20

KENNING M-M. (1990): Computer Assisted Language Learning. Language Teaching, April 1990, C.U.P., pp.67-76

KENNING M-M. & GUILLOT M-N. (1993): Electronic Monolingual Dictionaries as Language learning Aids. EuroCall 93, University of Hull.

KENNING M-M., & KENNING M. (1990): Computers and Language Learning - Current Theory and Practice. Ellis Horwood.

KNOWLES F. (1990): "Language and IT - Rivals or Partners?", Literary and Linguistic Computing, Vol.5.1, pp.40-44.

KNOWLES F. (forthcoming): Collocability in Languages for Special Purposes - Some Preliminaries. In A. Cowie (ed), Proceedings of the International Symposium on Phraseology, Leeds 1994, OUP

KRASHEN S. & TERRELL T. (1983): The Natural Approach. Alemany Press & Pergamon Press.

KUKULSKA-HULME A. (1988): A Computerized Interactive Vocabulary Development System for Advanced Learners. System, Vol.16, No.2, pp.163-170

LAST R. (1989): Artificial Intelligence Techniques in Language Learning. Ellis Horwood.

LAUFER B. & OSIMO H. (1991): Facilitating the Retention of Vocabulary - The Second-Hand Cloze. System Vol.19, No.3, pp.217-224

LAUFER B. (1991): The Development of L2 Lexis in the Expression of the Advanced Learner. Modern Language Journal 75.

LAUFER B. (1992): How Much Lexis is Necessary for Reading Comprehension? In Vocabulary and Applied Linguistics, Arnold P. & Bejoint H. (eds), 126-132, McMillan Academic & Professional.

LAURILLARD D. & MANNING P. (1993): Categorisation of Learner Performance for CALL Design. ReCALL No.9, pp.8-13

LAURILLARD D. (1979): The Processes of Student Learning. Higher Education, 8, pp.395-409.

LAURILLARD D. (1990): "Unit 2: How Computers can Assist Learning", in EH232 Computers and Learning, 2nd Level Course, Open University.

LAURILLARD D. (1991): "Principles for Computer-based Software Design for Language Learning", Computer Assisted Language Learning, Vol.4, Pt.3, pp.141-151

LAURILLARD D. (1993): Rethinking University Teaching - A Framework for the Effective use of Educational Technology. Routledge, London.

LEVELT W. (1989): Speaking. From Intention to Articulation. M.I.T.

LEVENSTON E. (1979): Second Language Acquisition - Issues and Problems. Interlanguage Studies Bulletin, Vol.2.No.4. p.154

LEVINE A. & REVES T. (1990): Does the Method of Vocabulary Presentation Make a Difference? TESL Canada Journal, Vol.8, No.1, pp.37-51

LEWIS M. (1993): The Lexical Approach. Language Teaching Publications.

LYMAN-HAGER M.A., DAVIS J., BURNETT J., CHENNAULT R. (1993): Une Vie de Boy: Interactive Reading in French. In Borchadt & Johnson (eds), p.93

LYONS J. (1970): Chomsky. Fontana/Collins.

MARTON F. & SALJO R. (1976): On Qualitative Differences in Learning. II - Outcome as a Function of the Learner's Conception of the Task. British Journal of Educational Psychology, Vol.46, pp.115-127

MARTON F. (1986): Describing and Improving Learning. In Learning Strategies and Learning Styles, (ed) R. Schmeck, Plenum, NY.

MATTHEWS C. (1992): Going AI - Foundations of ICALL. Computer Assisted Llanguage Learning, Vol.5, Pt.1&2, pp.13-31

McCARTHY M. (1984): A New Look at Vocabulary in EFL. Applied Linguistics, 55, pp.11-22

McCARTHY M. (1990): Vocabulary. O.U.P

McELLIGOTT A. & O'NEILL G. (1993): A Dictionary for CALL. EuroCall 93, University of Hull.

McKEE J. (1993): Computer-Assisted Vocabulary Acquisition - An Evaluation of the Text Analysis Program. ReCALL, No.8, pp.9-15

MEARA P. (1980): Vocabulary Acquisition - A Neglected Aspect of Language Teaching.Survey article, Language Teaching & Linguistics Abstracts 13, CUP.

MEARA P. (1984): The Study of lexis in Interlanguage. In Davies, Criper & Howatt (eds), Interlanguage, Edinburg University Press, pp.225-235

MEARA P. (1985): "Lexical Skills and CALL", in Computers in English Language Teaching, J. Brumfit (ed), Pergamon Press, pp.83-89

MILLER G. & FELLBAUM C. (1991): Semantic Networks of English. Cognition, Vol.41, pp.197-229

MILLER G. (1986): Dictionaries in the Mind. Language and Cognitive Processes, Vol.1, No.3, pp.171-185

MILLER G. (1990): WORDNET: An On-Line Lexical Database. International Journal of Lexicography 3.

NAGY W. & HERMAN P. (1987): Breadth & Depth Of Vocabulary Knowledge: Implications For Acquisition And Instruction. In McKeown & Curtis (eds), The nature of Vocabulary Acquisition. Lawrence Erlbaum.

NATION P. & COADY J. (1988): Vocabulary and Reading. In Carter & McCarthy (eds), pp.97-110

NATION P. (1983): Learning and Teaching Vocabulary. Victoria University, Wellington, New Zealand.

NATTINGER J. & DE CARRICO J. (1992): Lexical Phrases and Language Teaching. OUP.

NATTINGER J. (1988): Some Current Trends In Vocabulary Teaching. In Carter & McCarthy (eds), p.70

O'MALLEY J. & CHARNOT A. (1990): Learning Strategies in Second Language Acquisition. C.U.P.

OTT C., BLAKE R., BUTLER D. (1976): Implications of Mental Elaboration for the Acquisition of Foreign Language Vocabulary. International Review of Applied LInguistics in Language Teaching, Vol.14.No.1, pp.37-48

OWEN C. (1993): Corpus-Based Grammar and the Heineken Effect: Lexico-Grammatical Description for Language Learners. Applied Linguistics, Vol.14, No.2, pp.167-187

OXFORD R. & CROOKALL D. (1990): Vocabulary Learning: A Critical Analysis of Techniques. TESOL Canada Journal, Vol.7,No.2, p.14

PALMBERG R. (1987): Patterns of Vocabulary Development in Foreign Language Learners. Studies in Second Language Acquisition, Vol.9, pp.201-220

PALMBERG R. (1988): On Vocabulary-Knowledge Continua and Foreign-Language Learners' Mental Lexicons. Finnish Journal of Language Learning and Language Teaching, Vol 7, pp.93-99

PASK G. (1976): Styles and Strategies of Learning. British Journal of Educational Psychology 46, pp.128-148

PAWLEY A. & SYDER F. (1983): Two Puzzles for Linguistic Theory - Nativelike Selection and nativelike Fluency. In Richards & Schmidt (eds), Language and Communication, Longman.

PENNINGTON M. (1989): Teaching Languages with Computers. Athelstan.

PHILLIPS M. (1985): Logical Possibilities and Classroom Scenarios for the Development of CALL. in Brumfit et al. (eds), pp.25-46

PHINNEY M. (1989): Computers, Composition and Second Language Teaching. In Pennington (ed) pp.81-96

QUILLIAN M. (1967): Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities. Behavioural Science 12.

RAMSDEN R. (1992): Learning to teach in Higher Education. Routledge, p.44

RICHARDS J. (1986): The Role Of Vocabulary Teaching. TESOL Quarterly 10

RIVERS W. (1968): Teaching Foreign Language Skills. University of Chicago Press, Chicago.

RUDZKA B., CHANNELL J., OSTYN P., PUTSEYS Y. (1981): The Words You Need. Macmillan.

SCHOUTEN-VAN PARREREN C. (1992): Individual Differences in Vocabulary Acquisition - A Qualitative Experiment in the First Phase of Secondary Education. In Arnaud & Bejoint (eds), pp.94-101

SIEGEL M., & MISSELT A. (1984): Adaptive Feedback and Review Paradigm for Computer-Based Drills. Journal of Educational Psychology, 18, pp.294-309

SINCLAIR J. (1966): Beginning the Study of Lexis. In Bazell, Catford, Halliday & Robins (eds), In Memory of J.R. Firth, Longman, London.

SINCLAIR J. (1990): Corpus, Concordance, Collocation. OUP.

SMITH N. & WILSON D. (1979): Modern Linguistics - The Results of Chomsky's Revolution. Penguin.

STEVENS V. (1991): Concordance-Based Vocabulary Exercises - A Viable Alternative to Gap-Fillers. English Language Research Journal, Vol.4, pp.47-61

SUSSEX R. & CUMMING G. (1990): Intelligent Tools and Lexical Support for the Language Learner - Providing Advice at Task and Discussion Levels. In McDougall & Dowling (eds), Computers in Education, Elsevier Science Publications, pp459-464

SUSSEX R. (1991): Research in Practice and Practice on Research in CALL. CALL Journal, Vol 4 Part 3. pp.191-206

SUSSEX R., CUMMING G., CROPP S. (1994): A Tools-Based Environment for Discovery-Oriented CALL - Cognitive, Pedagogical and Ergonomic Issues for Interactive Learning. Computer Assisted Language Learning Vol.7, No.2, pp.133-149

SWARTZ M., KOSTYLA S., HANFLING S, HOLLAND M. (1991): Preliminary Assessment Of Foreign Language Learning Environment. CALL , Vol 1. pp.51-64

THOMPSON J, CHESTERS G. eds. (1994): Emancipation through Learning Technology - Selected Papers from the EUROCALL '93 Conference. Computers & Education, Vol.23, No.1/2

TIEBALT T. (1993): Assessing Hypertext for Languages for Specific Purposes. In Borchadt & Johnson (eds), p.166

TWADDELL F. (1973): Vocabulary Expansion in the ESOL Classroom. TESOL Quarterly 7, pp.61-78

UNDERWOOD J. (1984): Linguistics, Computers and the Language Teacher. Newbury House.

VAN EK J. (1977): The Threshold Level for Modern Language learning in Schools. Longman.

WEST M. (1953): A General Service List of English Words. Longman.

WEST M. (1960): Minimum Adequate Vocabulary. Longman.

WIDDIG W. & ESSER R. (1988): New Possibilities In Computer-Assisted Learning of Foreign Language Vocabulary Using Apple Macintosh. In Jung (ed), pp.265-278

WILKINS D. A. (1974): Second-Language Learning and Teaching. Edward Arnold, London.

ZAHNER C. (1991): "Word Grammars in ICALL", proceedings of Eurocall 1991, (eds) H. Savolainen & J. Telenius, Helsinki School of Economics & Business Administration, Dept. of Languages, pp.291-295

ZAHNER C, GUPTA G, OLOHAN M. (1994): Lexical Resources in CALL. In Thompson & Chesters (eds).pp.75-80.

ZECHMEISTER E., D'ANNA C., PAUS C., SMITH J. (1993): Meta-cognitive and Other Knowledge about the Mental Lexicon. Do we Know How Many Words we Know? Applied Linguistics, Vol 14, No.2. 188-205.

Appendix Ia (main text for pilot study 2):
(From Spotlight on Spanish, edited by K.Hall & D.Utley, Pan 1981)

ADOLESCENCIA: AÑOS CRUCIALES

Los adolescentes pagan un alto tributo por las ventajas que aporta la moderna tecnología en diversas partes del mundo y la mayoría de los esfuerzos realizados para cambiar esta situación y favorecer a los jóvenes que han fracasado en la presente sociedad postindustrial, según se desprende de un estudio sobre los «años cruciales» de la adolescencia publicado por la «Organización Mundial de la Salud» (O.M.S.) y dado a conocer en Ginebra.

«La transformación radical de la sociedad ha complicado las relaciones humanas y perturbado la seguridad personal de los adolescentes, enfrentando así las sociedades modernas a un grave problema» agrega el estudio.

«Por otra parte -sigue-, la cronología de la evolución mundial era medida al principio en millones de años y después del desarrollo de la agricultura en miles de años. Pero la revolución industrial sólo requirió doscientos años y en la sociedad postindustrial de hoy en día las motivaciones profundas acontecen de una generación a otra.». «La adolescencia - sostiene el estudio de la O.M.S.-es un período crucial para la cristalización de un cierto número de hábitos que tienen influencia directa sobre la salud, como el uso del tabaco, el alcohol, las drogas y el comportamiento al volante, además de la alimentación y el ejercicio físico.»

Finalmente señala que «el abuso de drogas de todo tipo, convertido actualmente en una verdadera epidemia, juega un papel primordial en las decisiones que adoptan las autoridades sanitarias de muchos países para intenter una eficaz solución a los problemas de salud que padecen los adolescentes».

¡Juventud, divino tesoro,
ya te vas para no volver!
Cuando quiero llorar, no lloro...
y a veces lloro sin querer...

Appendix Ib (main text for SFL group - chapter five):

MUERTE POR FUSILAMIENTO
José María MENDIOLA

El Ministro de Información levantó una mano.
Por favor, caballeros - dijo, con voz profunda - no puedo contestar varias preguntas al mismo tiempo. La salita estaba atestada. Como consecuencia de la larga espera que habían padecido, los corresponsales de Prensa habían fumado con exceso. Apenas se podía respirar. En el suelo había colillas, restos de papeles que tal vez habían envuelto emparedados, ceniza. El ministro estaba desagradado. Le molestaba que aquella gente no guardara las formas, que los nudos de las corbatas estuvieran flojos y se divisara, al fondo, un periodista en mangas de camisa ....
¡ Usted! - señaló, acusador - ¿Quién es usted, si me hace el favor?
Las miradas de los periodistas buscaron el objetivo del Ministro.
 - Usted - repitió éste, con voz seca - . El señor que no tiene la chaqueta puesta.
El aludido se azoró. Mascaba chiclé, naturalmente.
- Jaime Ardilla, de "La Hora" . . .
- ¿Quiere hacer el favor de ponerse la chaqueta?
 - Perdone . . . - El periodista se la puso, y hasta se peinó, con los dedos, su cabello revuelto - . Lo lamento.
- Señores. - El Ministro odiaba la desenvoltura en los demás -. Quiero advertirles que cumpliré mi cometido, facilitando una simple nota offcial, si el comportamiento de ustedes ....
Hubo protestas. Alguien, también en el fondo, levantó los brazos, agitándolos. Se seguía fumando de una manera desordenada.
- ¡ Una nota oficial ! - repitió, amenazador.
- ¡Por favor! - pidió un hombrecillo de la primera fila. Se volvió a sus compañeros y gritó -: ¡Dejadme a mí!
El Ministro aguardaba. El hombrecillo se enfrentó con sus compañeros, levantó los brazos y gritó: "¡ Yo preguntaré !"
Lentamente, el vocerío fue cesando. Un fotógrafo se acercó, y el Ministro dijo:
- ¡ Nada de cámaras !
El hombrecillo se adelantaba ahora. Había conseguido un silencio discreto, aunque no total.
- Señor Ministro - empezó -. Soy Zelada, de "El Tiempo". Permita que sea yo quien . . .
- Sí, empiece. Y no toleraré ningún desorden.
- Sí, señor. ¿Es cierto lo que. . .?
- Si - dijo el Ministro -. El Presidente ha sido asesinado.

Appendix Ic (main text for ESL group - chapter five):

The other night I was cycling home from my evening class when I suddenly became aware that the bike was bumping along as if I was riding over a ploughed field. I stopped and had a look, and sure enough it was a puncture. I didn't have anything to fix it with, but with slow punctures you can sometimes just pump it up again and maybe it'll last until you get home, so I got out the pump and was just attaching it to the tyre when this guy came up and asked me for a light. Well, I don't smoke, but on this occasion I did happen to have some matches in my pocket, because I was going to have a barbecue later, so I offered them to him. He struck a match and at that moment a cat came running out of a nearby house. It must have been startled by the light because it swerved and ran into the road, right in front of a car. The car screeched to a halt just in time and the cat escaped, but the guy who was lighting his cigarette looked up and said "Mary! I don't believe it!". The woman who was driving the car looked over and her jaw dropped. "Michael?". It turned out they used to be married, about 10 years ago, but it hadn't worked out and they'd separated and not seen each other since. They'd missed each other, though, and were really pleased to meet again. There was a pub just over the road so we all went and had a glass of champagne to celebrate. Someone came in selling flowers and Michael bought her a single red carnation, it was very romantic. Then we got the barman to take a photograph of us all together, and we exchanged addresses and they would said they send me a copy of the picture. They went off really happy to start a new life together. And all because of a puncture and a cat! Life is pretty surprising sometimes. When I got outside someone had stolen my bike.

Appendix IIa (subjects' selections, notes and groups - main test chapter five):

S16 (ESL)

Selected Words	Notes	Groups
aware bumping ploughed puncture riding pump it up struck startled swerved screeched jaw dropped	knowledge politically connaissances en politique bump to hurt somebody heurter quelqu'un attache d'un cheval ou d'un tracteur . perforation d'une pressure l'art et la practiaue d'une activites de quelque chose ex: d'une voiture, d'un velo augmenter la pression allumer une allunette to be surprise or frightened son ou un cri de pleure to be surprise fell down .	BYCYCLE pump puncture riding bumping ING WORDS riding bumping ED WORDS dropped ploughed swerved startled SOUNDS screeched OUGH ploughed ED WORD dropped

S19 (ESL)

Selected Words	Notes	Groups
aware bumping ploughed puncture struck swerved screeched riding startled suddenly halt	agonizante/ aterrorizado bater /trombar carreta,arrado furo,buraco passado de greve/ batida/ riscar /'acao' balancar,chacoulhar cantar do pneu freando assento de cavalo, ou bicicleta surpreso/ impressionado precipitado / rapido parado/interropcao .	FEELINGS startled ACTIONS bumping halt swerved startled ING bumping riding LY suddenly ED startled ploughed

S18 (ESL)

Selected Words	Notes	Groups
aware	Informed notificar informacion general	ED WORDS screeched
bumping	Dice que la zona transitada no esta lisa,y tiene tropiezos , incovenientes, obstruciones, no esta plana	startled swerved BICYCLE WORDS puncture riding
ploughed	arado,	
puncture	a hole in the tyre	
attaching	to join ,fasten, connect to	
struck	lighted a match	
startled	to be soprised or fraightened diverso ,	
swerved	cruzar, desviar , inclenarse,	
screeched	high pitched sound or cry.	
jaw	dark rum. especie de animal marino se llama tiburon.es criminal, carnivor, reprodutivo, comestible. flower of this	
carnation	plant is pink ,o reddish pink color.	
I suddenly	de repente , sorpresivo, inprevisto, inesperado,	
riding		
stopped		

S10 (SFL)

Selected Words	Notes	Groups
atestada	packed	PP
padecido	suffer, endure, put up with	envuelto
larga	long	padecido
colillas	butt	
envuelto		CIGARETTE
emparedados	sandwich	colillas
ceniza	ashes	ceniza
guardara las formas	maintain order	
mangas de camisa		TIME
señaló	to mark, point out, fix	larga
se azoró	to get alarmed	
mascaba chiclé	chew	DISCONTENT
cabello		padecido
odiaba	odiar to hate	se azoro
desenvoltura	free and easy manner	
cometido	task, commitment	PROXIMITY
amenazador		se acerco
enfrentó	to confront	se adelantaba
vocerío	shouting	
se adelantaba		
se acercó		

S17 (ESL)

Selected Words	Notes	Groups
aware	Informed , knowing knowlegdeable Informado, conociendo, conocimiento	REPAIRING attaching puncture
bumping	pegar,	ROAD
ploughed		bumping
puncture		ploughed
attaching	juntar	
struck	Golpear.aporrear.	ED WORDS
startled	Sorprendido, austado.	screeched
swerved	turning away	startled
screeched	Chillido, sonido fuerte.	swerved
jaw	Quijada.	ploughed
stress	emphasis	
relevant	pertinent	COUNTRY
stresses		ploughed
challenging		
tend	tendencia	BODY
earnings		jaw
thrilled	excited	
nip	pinch - sting sensation	ACCIDENT
muddy	fined-grained soft wet	screeched
bound	deposit	struck
race riots	sure -	swerved
fair share of trouble	racism fighting	
crawls	visit a lot pubs	OUGH=OW ploughed
		FEELINGS startled aware

S15 (ESL)

Selected Words	Notes	Groups
aware	surprise	FEELINGS
bump	to knock shake	aware
ploughed	like a imperfect field not	
puncture	plain	BODY'S PART
attaching	hole in the tyre fix,conect	jaw
swerved		U
jaw	part of the mouth	bump
struck	light by friction	puncture
thrilled	suspect	
regret	sorry, guilt	BICYCLE
although		attaching
nip	small bite	
muddy	coverd or not clear	
bound	certain	
rate	money currency	
riots	public fait, demostration	
muggers	some body who robs	
mutual	experienced or expressed by more tham two people	
crawls	to walk like a dog	
looking forward	directed or moving ahead	
seems	looks like	

Selected Words	Notes	Groups
aware	knowledge conhecimento sobre alguma coisa	FEELINGS startled
bumping	hit strike slam against some thing	thrilled regards
ploughed	acao de bater em alguma coida como se tivesse arando a terra [a maneira dele estar andando de	ACTION ploughed screeched
puncture	bicileta]	startled
attaching	a ismall hole [furo]	swerved
struck	conectado	bumping
startled	acao de riscar o fosforo a acao de causar surpresa	ING
swerved	ou medo acao de movimentar ou virar	bumping attaching
screeched	is the noise from the tyer agains the road or the sound the bird [som do passaro]	ED puncture screeched startled
jaw	queixo	swerved
thrilled	excitante , assustador emocionatne	FAST MOVEMENT
nip	severe frost or cold sensacao de frio	swerved screeched
muddy	not clear or bright	SENSATION
bound	certain to	nip
bother about		startled
race riots	racial tumult	
im looking forward to seeing you	expecting happy	IDIOMATIC EXPRESSION
regards	goods wishes	I'm looking forward to seeing you
		FAITH race riots
		FORM struck

Selected Words	Notes	Groups
caballeros	gentlemen	ROOM
la salita	room	el suelo
padecido	suffer	el techo
el suelo	floor	la salita
le molestaba	imperfect, was annoying	
flojos	loose , undone	BAD FEELINGS
mangas de camisa		le molestaba
los nudos	the knots, bends ties.	padecido
colillas	cigarette end fumador.	
envuelto	wrapped their sandwiches	VOLVER
emparedados		
las miradas de los	to look, the views of the	desenvoltura en los
periodistas	journalists.	demas
		cabello revuelto
seca		envuelto emparedados
el aludido se azoró	the indicated one was	
	alarmed.	HUMAN BODY
mascaba chiclé	mascar to chew	los dedos
se la puso	pret. poner. chaqueta	corpus
se peinó	peinar to comb	el mano
los dedos	finger toe el maño.	
cabello revuelto	revolver to scramble	HAIR
	messy. hair.	se peino
odiaba	to hate se gusta.	cabello revuelto
desenvoltura en los	free and easy of the rest	
demás	of them volver	ADVANCES
cometido	cometido task	se adelantaba
comportamiento	behaviour	se acerco
hubo	preterite haber	
se seguía	imp to continue	
	pursue	
amenazador		
gritó	to shout	
se acercó	preterite acercar to	
	bring	
se adelantaba	to advance	

Selected Words	Notes	Groups
salita	room	PAST PARTICIPLES
atestada	packed	atestada
padecido	padecer-to suffer	desagradado
apenas	hardly	
colillas	cigarette butt	SMOKING
emparedados	sandwich	ceniza
ceniza	ash	colillas
flojos	slack,loose,limp,lazy,idle	fumar
se divisara	to sight, to spot-imperfect subjunctive	
		TYING
desagradado	dissatisfied	nudos
molestaba	bothered-imperfect	flojos
formas	forms	
nudos	knots	SUFFERING/
mangas	sleeve	ANNOYANCE
seca		molestaba
aludido	to allude to,to refer to-past participle	padecido desagradado
azoró	to get alarmed past	
caballeros	gentlemen ·	IMPERFECTOS
larga	lengthy	molestaba
espera		
acusador	accusing	CHEWING
miradas	glances,looks	mascaba chiclé
mascaba	to chew-3rd person imperfect	
periodista	journalist	
puesta	pp de poner	
se peinó	se peinar-to comb one's hair hair	
cabello		
odiaba	odiar-to hate	
desenvoltura	sloppiness	
demás	the rest	
advertirles	advertir-to warn	
cumpliré	cumplir-to fulfill-future	
cometido	task	
se seguía	he continued-imp	
amenazador	threatening	
un hombrecillo	hombre	
dejadme	leave it to me	

Selected Words	Notes	Groups
una mano		COWBOYS
levantó	to raise	colillas
salita	room	fusilamiento
padecido	to endure	
prensa	the press	PERFECT
apenas	hardly	padecido
colillas	cigarette butt	atestada
fusilamiento	shooting	cometido
atestada	packed	aludido
el suelo	floor	
envuelto	to cover	NEWSPAPER
ceniza	ash	un periodista
desagradado	displease	prensa
nudos	knot	periodico
flojos	loose	
mangas	sleeve	SMOKING
un periodista	journalist	ceniza
se divisara	to distinguish	colillas
miradas	glance look	
chaqueta puesta	wearing a jacket	ROOM
aludido	alluded to	salita
se azoró	to get alarmed	el suelo
mascaba	to chew	
chiclé	chewing gum	CHEWING
se peinó	to comb	mascaba
revuelto	messy	chiclé
dedos	finger	
odiaba	to hate	COMBING
la desenvoltura	free and easy manner	revuelto
los demás	the others	se peinó
cumpliré	carry out	dedos
cometido	task	
facilitando	providing	HATE
el fondo		desagradado
se seguía	to continue	revuelto#padecido
		odiaba
		FLOPPY
		flojos
		la desenvoltura
		nudos

Selected Words	Notes	Groups
caballeros	Gentlemen	SMOKING
salita	room	ceniza
atestada	Packed	colillas
padecido	To suffer To bear	
apenas	Hardly	WRAPPED
colillas	cigarette end	revuelto
emparedados	Sandwich	envuelto
ceniza	Ashes	desenvoltura
mangas	Sleeve	
miradas	under the gaze of	RUNNING HIS
puesta	dressed	FINGERS THROUGH
se azoró	to be alarmed	HIS HAIR
mascaba chiclé	was chewing gum very	se peinó
naturalmente	naturally	los dedos
se peinó	combed	cabello
los dedos	fingers	
revuelto	Messed up	REFLEXIVE
desenvoltura	In a free and easy manner	PRETERITES
envuelto		se peinó
demás	the rest	se azoró
advertirles	to warn	se acercó
cometido	assignment, task	
comportamiento	behavior, conduct	HALL
desordenada	disorder	atestada
gritó	to yell	emparedados
dejadme	leave it to me	
aguardaba	to wait	HATE OF FREE AND
se acercó	come near	EASY MANNER
		demás
		desenvoltura

Appendix IIb (log files - main test ch.5, subject 14's 1st session):

LOGGED ON william 11-04-1993
CORPUS START 14:18:01
View Text bstory3
Add Text bstory3
CORPUS END 14:19:05
SELECT START 14:19:10
View Text bstory3
Select Word caballeros MEANING
Select Word la salita MEANING
Select Word la salita CONTEXT
Select Word padecido MEANING
Select Word padecido FORM
Select Word padecido CONTEXT
Select Word el suelo MEANING
Select Word le molestaba FORM
Select Word flojos MEANING
SELECT END 14:27:50
GROUPING START 14:27:55
Click Word caballeros MEANING
Dictionary caballeros
Save Notes caballeros
Click Word el suelo MEANING
Save Notes el suelo
New Group room MEANING
Item to Group el suelo room MEANING
Extra Item el techo room MEANING
Click Word flojos MEANING
Dictionary flojos
Dictionary flojos
Concordance flojo
Save Notes flojos
Click Word la salita MEANING
Click Word la salita CONTEXT
Dictionary la salita
Save Notes la salita
Item to Group la salita room MEANING
Click Word padecido MEANING
Click Word padecido FORM
Click Word padecido CONTEXT
Concordance padec
Dictionary padecido
Save Notes padecido
Click Word le molestaba FORM
Dictionary le molestaba
Save Notes le molestaba
Click Word padecido FORM
GROUPING END 14:48:21
SELECT START 14:48:23
View Word le molestaba bstory3
Select Word le molestaba MEANING
SELECT END 14:48:40
SELECT START 14:48:42

SELECT END 14:48:44
GROUPING START 14:48:45
Click Group room MEANING
New Group bad feelings MEANING
Click Word le molestaba MEANING
Click Word le molestaba FORM
Item to Group le molestaba bad feelings MEANING
Click Word padecido MEANING
Click Word padecido FORM
Click Word padecido CONTEXT
Item to Group padecido bad feelings MEANING
Click Word caballeros MEANING
Click Word el suelo MEANING
Click Word flojos MEANING
Click Word la salita MEANING
Click Word la salita CONTEXT
Click Word le molestaba MEANING
Click Word le molestaba FORM
Click Word padecido MEANING
Click Word padecido FORM
Click Word padecido CONTEXT
GROUPING END 14:51:21
TEST START 14:51:29 padecido
Items 6
 Score 000.0%
Response padecido parecido
Feedback TARGET WORD ALMOST RECALLED Items 6
 Score 016.7%
Response Click parecido
Response padecido padecido
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 016.7%
Next Item (forward) le molestaba
Response le molestaba la suela
Feedback OCCASIONAL LETTER(S) CORRECT Items 6
 Score 016.7%
Group Clue le molestaba yes
Response le molestaba le molestaba
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 033.3%
Next Item (forward) caballeros
Response caballeros caballeros
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 050.0%
Next Item (forward) la salita
Response la salita la suela
Feedback SOME TARGET WORD ELEMENTS CORRECT

Items 6
 Score 050.0%
Response la salita la salita
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 066.7%
Next Item (forward) el suelo
Group Clue el suelo yes
Response el suelo la suela
Feedback SOME TARGET WORD ELEMENTS CORRECT
Items 6
 Score 066.7%
Response el suelo el suelo
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 083.3%
Next Item (forward) flojos
Response flojos flojos
Feedback TARGET WORD CORRECTLY RECALLED
Items 6
 Score 100.0%
TEST END 14:59:57
LOGGED OFF 15:03:11
LOGGED ON william 11-05-1993
SELECT START 11:06:28
View Word caballeros bstory3
Select Word mangas de camisa MEANING
Select Word los nudos MEANING
Select Word los nudos CONTEXT
Select Word colillas MEANING
Select Word colillas CONTEXT
Select Word envuelto emparedados MEANING
Select Word envuelto emparedados FORM
Select Word las miradas de los periodistas MEANING
Select Word las miradas de los periodistas FORM
Select Word seca MEANING
Select Word el aludido se azoró MEANING
Select Word el aludido se azoró FORM
SELECT END 11:28:56
GROUPING START 11:28:59
Click Group bad feelings MEANING
Click Group room MEANING
Click Word colillas MEANING
Click Word colillas CONTEXT
Dictionary colillas
Dictionary colillas
Save Notes colillas
Click Word el aludido se azoró MEANING

Click Word el aludido se azoró
FORM
Dictionary el aludido se azoró
Dictionary el aludido se azoró
Concordance azor
Save Notes el aludido se azoró
Click Word envuelto
emparedados MEANING
Click Word envuelto
emparedados FORM
Dictionary envuelto emparedados
Dictionary envuelto emparedados
Concordance emparedado
Save Notes envuelto
emparedados
Click Word las miradas de los
periodistas MEANING
Click Word las miradas de los
periodistas FORM
Dictionary las miradas de los
periodistas
Dictionary las miradas de los
periodistas
Concordance periodistas
Save Notes las miradas de los
periodistas
Click Word los nudos
MEANING
Click Word los nudos
CONTEXT
Dictionary los nudos
Save Notes los nudos
Concordance los nudos
Save Notes los nudos
GROUPING END 11:52:55
TEST START 11:53:05 le
molestaba Items 13
  Score 046.2%
Group Clue le molestaba yes
Response le molestaba el
molestaba
Feedback TARGET WORD
ALMOST RECALLED Items
13
  Score 046.2%
Response le molestaba le
molestaba
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 046.2%
Next Item (forward) caballeros
Response caballeros caballeros
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 046.2%
Next Item (forward) los nudos
Response los nudos los nudos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 053.8%
Next Item (forward) la salita
Response la salita la salita
Feedback TARGET WORD
CORRECTLY RECALLED

Items 13
  Score 053.8%
Next Item (forward) colillas
Response colillas colillas
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 061.5%
Next Item (forward) las miradas
de los periodistas
Response las miradas de los
periodistas las miradas de los
periodistas
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 069.2%
Next Item (forward) envuelto
emparedados
Group Clue envuelto
emparedados no
Notes Clue envuelto
emparedados yes
Response envuelto emparedados
envuelto emparedados
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 076.9%
Next Item (forward) el aludido se
azoró
Response el aludido se azoró el
aludido se azoró
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 084.6%
Next Item (forward) el suelo
Response el suelo el suelo
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 084.6%
Next Item (forward) flojos
Response flojos flojos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 084.6%
Next Item (forward) seca
Response seca seca
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 092.3%
Next Item (forward) mangas de
camisa
Response mangas de camisa
mangas de camisa
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 100.0%
TEST END 12:03:20
LOGGED OFF 12:03:23
LOGGED ON william 11-12-
1993
SELECT START 11:00:32

SELECT END 11:00:39
GROUPING START 11:00:42
GROUPING END 11:00:47
TEST START 11:00:53 los
nudos Items 13
  Score 100.0%
TEST END 11:00:58
LOGGED OFF 11:01:01
LOGGED ON william 11-12-
1993
TEST START 11:06:39 los
nudos Items 13
  Score 100.0%
Group Clue los nudos no
Response los nudos los nudos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 100.0%
Next Item (forward) colillas
Group Clue colillas no
Notes Clue colillas yes
WordShape Clue colillas
Grammar Clue colillas yes
Next Item (forward) las miradas
de los periodistas
Grammar Clue las miradas de los
periodistas yes
Notes Clue las miradas de los
periodistas yes
Response las miradas de los
periodistas las miridas de los
periodistas
Feedback TARGET WORD
ALMOST RECALLED Items
13
  Score 100.0%
Response Click las miridas de
los periodistas
Response las miradas de los
periodistas las miradas de los
periodistas
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 100.0%
Next Item (forward) envuelto
emparedados
Group Clue envuelto
emparedados no
Notes Clue envuelto
emparedados yes
Response envuelto emparedados
envuelto empreadores
Feedback TARGET WORD
PARTIALLY RECALLED Items
13
  Score 092.3%
Response Click envuelto
empreadores
Response envuelto emparedados
envuelto emparedados
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
  Score 100.0%
Next Item (forward) el aludido se
azoró

Notes Clue el aludido se azoró
yes
Grammar Clue el aludido se
azoró yes
Group Clue el aludido se azoró
no
WordShape Clue el aludido se
azoró
Response el aludido se azoró el
apelido se azuró
Feedback TARGET WORD
PARTIALLY RECALLED Items
13
 Score 092.3%
Next Item (forward) mangas de
camisa
Notes Clue mangas de camisa no
Group Clue mangas de camisa no
Grammar Clue mangas de
camisa yes
WordShape Clue mangas de
camisa
Next Item (forward) seca
Response seca sec
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 13
 Score 084.6%
WordShape Clue seca
Response seca sece
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 13
 Score 084.6%
Response seca seca
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
Next Item (forward) padecido
Notes Clue padecido yes
Grammar Clue padecido yes
Group Clue padecido yes
WordShape Clue padecido
Next Item (forward) caballeros
Response caballeros cabollores
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 13
 Score 092.3%
WordShape Clue caballeros
Response Click cabollores
Response caballeros caballeros
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
Next Item (forward) la salita
Group Clue la salita yes
Response la salita la salita
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
Next Item (forward) el suelo
Group Clue el suelo yes
Response el suelo el suelo

Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
Next Item (forward) flojos
Response flojos flojos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
Next Item (forward) le molestaba
Group Clue le molestaba yes
Response le molestaba le
molestaba
Feedback TARGET WORD
CORRECTLY RECALLED
Items 13
 Score 092.3%
TEST END 11:31:29
SELECT START 11:31:34
View Text bstory3
Select Word mascaba chiclé
MEANING
Select Word mascaba chiclé
CONTEXT
View Word mascaba chiclé
bstory3
View Word mascaba chiclé
bstory3
Delete Word seca (Finished)
MEANING
Delete Word flojos (Finished)
MEANING
Delete Word le molestaba
MEANING
Delete Word el suelo (Finished)
MEANING
Select Word se la puso
MEANING
Select Word se peinó MEANING
Select Word se peinó FORM
Select Word se peinó CONTEXT
Delete Word caballeros
(Finished) MEANING
Delete Word la salita
CONTEXT
Delete Word los nudos
CONTEXT
Delete Word colillas CONTEXT
Select Word los dedos
MEANING
Select Word cabello revuelto
MEANING
Select Word cabello revuelto
FORM
Select Word odiaba MEANING
Delete Word colillas (Finished)
MEANING
Select Word desenvoltura en los
demás MEANING
SELECT END 11:42:21
GROUPING START 11:42:25
Click Word cabello revuelto
MEANING
Click Word cabello revuelto
FORM
Dictionary cabello revuelto
Save Notes cabello revuelto

Click Word desenvoltura en los
demás MEANING
New Group volver MEANING
Item to Group desenvoltura en
los demás volver MEANING
Click Word cabello revuelto
MEANING
Item to Group cabello revuelto
volver MEANING
Click Word envuelto
emparedados MEANING
Click Word envuelto
emparedados FORM
Item to Group envuelto
emparedados volver MEANING
Click Word desenvoltura en los
demás MEANING
Dictionary desenvoltura en los
demás
Save Notes desenvoltura en los
demás
Click Word los dedos
MEANING
Dictionary los dedos
Save Notes los dedos
New Group human body
MEANING
Click Word los dedos
MEANING
Item to Group los dedos human
body MEANING
Extra Item corpus human body
MEANING
Extra Item el mano human body
MEANING
Click Word mascaba chiclé
MEANING
Click Word mascaba chiclé
CONTEXT
Dictionary mascaba chiclé
Click Word odiaba MEANING
Dictionary odiaba
Save Notes odiaba
Click Word se la puso
MEANING
Dictionary se la puso
Save Notes se la puso
Concordance se la puso
Save Notes se la puso
GROUPING END 12:04:40
LOGGED OFF 12:04:41
LOGGED ON william 11-18-
1993
SELECT START 12:50:10
SELECT END 12:50:13
GROUPING START 12:50:16
GROUPING END 12:50:22
TEST START 12:50:28 los
nudos Items 20
 Score 060.0%
TEST END 12:50:30
LOGGED OFF 12:50:32
LOGGED ON william 11-18-
1993
GROUPING START 13:57:48
Click Word cabello revuelto
MEANING

Click Word cabello revuelto
FORM
Click Word desenvoltura en los
demás MEANING
Click Group volver MEANING
Click Word los dedos
MEANING
Click Word mascaba chiclé
MEANING
Click Word mascaba chiclé
CONTEXT
Dictionary mascaba chiclé
Save Notes mascaba chiclé
Click Word padecido MEANING
Click Word padecido FORM
Click Word padecido CONTEXT
Click Word padecido MEANING
Click Word se la puso
MEANING
Concordance chaqueta
Click Word se peinó MEANING
Click Word se peinó FORM
Click Word se peinó CONTEXT
Click Word se peinó MEANING
Dictionary se peinó
Save Notes se peinó
New Group hair MEANING
Click Word se peinó MEANING
Item to Group se peinó hair
MEANING
Click Word cabello revuelto
MEANING
Click Word cabello revuelto
FORM
Item to Group cabello revuelto
hair MEANING
GROUPING END 14:08:17
SELECT START 14:08:20
View Text bstory3
View Word padecido bstory3
Delete Word padecido
MEANING
Delete Word envuelto
emparedados MEANING
Delete Word la salita (Finished)
MEANING
Delete Word los nudos
(Finished) MEANING
Delete Word el aludido se azoró
MEANING
Delete Word se peinó
MEANING
Select Word cometido
MEANING
Select Word comportamiento
MEANING
Select Word hubo MEANING
Select Word se seguía
MEANING
Select Word se seguía FORM
Select Word amenazador
MEANING
Delete Word los dedos
(Finished) MEANING
Delete Word las miradas de los
periodistas MEANING
Delete Word se peinó
CONTEXT

Delete Word mascaba chiclé
CONTEXT
Delete Word padecido
CONTEXT
Select Word gritó MEANING
Select Word se acercó
MEANING
Select Word se adelantaba
MEANING
SELECT END 14:19:58
GROUPING START 14:20:01
Click Word amenazador
MEANING
Dictionary amenazador
Click Word cometido MEANING
Dictionary cometido
Save Notes cometido
Click Word comportamiento
MEANING
Dictionary comportamiento
Save Notes comportamiento
Click Word gritó MEANING
Dictionary gritó
Save Notes gritó
Click Word hubo MEANING
Dictionary hubo
Save Notes hubo
Concordance hubo
Click Word se acercó MEANING
Dictionary se acercó
Save Notes se acercó
Click Word se adelantaba
MEANING
Dictionary se adelantaba
Save Notes se adelantaba
New Group advances MEANING
Click Word se adelantaba
MEANING
Item to Group se adelantaba
advances MEANING
Click Word se acercó MEANING
Item to Group se acercó advances
MEANING
Click Word se seguía MEANING
Click Word se seguía FORM
Dictionary se seguía
Concordance segu
Save Notes se seguía
Click Word se adelantaba
MEANING
Click Word se acercó MEANING
Click Word hubo MEANING
Click Word gritó MEANING
Click Word comportamiento
MEANING
Click Word cometido MEANING
GROUPING END 14:35:08
SELECT START 14:35:15
SELECT END 14:38:52
GROUPING START 14:38:54
Click Word comportamiento
MEANING
GROUPING END 14:39:03
SELECT START 14:39:05
SELECT END 14:44:05
TEST START 14:44:13 padecido
Items 28
Score 042.9%

Notes Clue padecido yes
Group Clue padecido yes
Response padecido padecido
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 042.9%
Next Item (forward) los nudos
Response los nudos los nudos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 042.9%
Next Item (forward) mascaba
chiclé
Notes Clue mascaba chiclé yes
Response mascaba chiclé
mascaba chiclé
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 046.4%
Next Item (forward) se la puso
Notes Clue se la puso yes
Response se la puso se le puso
Feedback TARGET WORD
ALMOST RECALLED Items
28
Score 050.0%
Response se la puso se lo puso
Feedback TARGET WORD
ALMOST RECALLED Items
28
Score 050.0%
Response se la puso se lo puso
Feedback TARGET WORD
ALMOST RECALLED Items
28
Score 050.0%
Grammar Clue se la puso yes
Response se la puso se la puso
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 050.0%
Next Item (forward) caballeros
Response caballeros cabellores
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 28
Score 050.0%
Response caballeros caballeros
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 050.0%
Next Item (forward) colillas
Response colillas colillas
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
Score 050.0%
Next Item (forward) el aludido se
azoró
Response el aludido se azoró se
alludido se azuró
Feedback TARGET WORD
PARTIALLY RECALLED Items

28
  Score 050.0%
Response Click se alludido se
azuró
Response el aludido se azoró el
aludido se azoró
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 053.6%
Next Item (forward) envuelto
emparedados
Notes Clue envuelto
emparedados yes
Response envuelto emparedados
envuelto emparedades
Feedback TARGET WORD
ALMOST RECALLED Items
28
  Score 053.6%
Response Click envuelto
emparedades
Response envuelto emparedados
envuelto emparedados
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 053.6%
Next Item (forward) se peinó
Response se peinó se peinó
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 057.1%
Next Item (forward) hubo
Response hubo hubo
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 060.7%
Next Item (forward) cabello
revuelto
Response cabello revuelto
envoltura cabello
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 28
  Score 060.7%
Response cabello revuelto
cabello envuelto
Feedback TARGET WORD
PARTIALLY RECALLED Items
28
  Score 060.7%
Group Clue cabello revuelto yes
WordShape Clue cabello
revuelto
Response cabello revuelto
cabello revuelto

Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 064.3%
Next Item (forward) se seguía
Response se seguía se seguía
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 067.9%
Next Item (forward) mangas de
camisa
Notes Clue mangas de camisa no
Group Clue mangas de camisa no
Grammar Clue mangas de
camisa yes
Next Item (forward) seca
Response seca seca
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 067.9%
Next Item (forward) los dedos
Response los dedos los dedos
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 071.4%
Next Item (forward) odiaba
Notes Clue odiaba yes
Response odiaba odiaba
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 075.0%
Next Item (forward) cometido
Response cometido cometido
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 078.6%
Next Item (forward) desenvoltura
en los demás
Notes Clue desenvoltura en los
demás yes
Response desenvoltura en los
demás envoltura y de los démas
Feedback TARGET WORD
PARTIALLY RECALLED Items
28
  Score 078.6%
Response Click envoltura y de
los démas
Grammar Clue desenvoltura en
los demás yes
Response desenvoltura en los
demás desenvoltura en los demás
Feedback TARGET WORD
CORRECTLY RECALLED

Items 28
  Score 082.1%
Next Item (forward)
comportamiento
Notes Clue comportamiento yes
Response comportamiento
comportamiento
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 085.7%
Next Item (forward) amenazador
Notes Clue amenazador no
Response amenazador
amenazador
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 089.3%
Next Item (forward) gritó
Notes Clue gritó yes
Response gritó gritó
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 092.9%
Next Item (forward) se acercó
Group Clue se acercó yes
Response se acercó se acerdido
Feedback SOME TARGET
WORD ELEMENTS CORRECT
Items 28
  Score 092.9%
Response se acercó se acerco
Feedback TARGET WORD
PARTIALLY RECALLED Items
28
  Score 092.9%
Response se acercó se acercó
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 096.4%
Next Item (forward) se
adelantaba
Response se adelantaba se
adelantaba
Feedback TARGET WORD
CORRECTLY RECALLED
Items 28
  Score 100.0%
Next Item (forward) la salita
Next Item (forward) el suelo
Next Item (forward) flojos
Next Item (forward) las miradas
de los periodistas
Next Item (forward) le molestaba
Notes Clue le molestaba yes
TEST END 15:10:07
LOGGED OFF 15:10:11

Appendix IIc (self-assessment data - pilot study 2 chapter four, and main test chapter five):

S1 (SFL -Pilot Study 2)

Item	Remember the Context	Know the Meaning	Use in other Context
aporta	√	√	√
esfuerzos	√	√	√
favorecer		√	√
fracasado	?	√	√
desprende	√	?	√
perturbado	√	√	√
enfrentando		√	√
agrega el estudio	√	√	√
de hoy en día	?	√	√
motivaciones	√	√	√
profundas		√	√
acontecen			
comportamiento	√	√	√
sueño	√	√	√
palmeras	√	√	√
ventanas	√	√	√
enlozado	√	√	√
húmedo		√	
rodeado			
arcada	√	?	
estrecha			
juzgar		√	√

S16 (ESL)

Item	Remember theContext	Know the Meaning	Use in other Context
aware			
bumping	√	√	
ploughed			
puncture	√	√	√
riding			
pump it up	√	√	√
struck			
startled			
swerved			
screeched	?	?	
jaw	?	√	√
dropped		√	?

S18 (ESL)

Item	Remember theContext	Know the Meaning	Use in other Context
aware	√	√	√
bumping	√	√	
ploughed			
puncture	√	√	√
attaching		√	√
struck			
startled			
swerved		-	
screeched			
jaw			

carnation			
i suddenly	√	√	√
riding	√	√	√
stopped	√	√	√

S17 (ESL)

Item	Remember theContext	Know the Meaning	Use in other Context
aware	√	√	
bumping	√		√
ploughed	√	?	√
puncture	√	√	√
attaching	√	√	√
struck	√	√	√
startled	√	√	√
swerved	√	√	√
screeched	√	√	√
jaw	√	√	√
stress	√	√	√
relevant			
stresses		?	
challenging		?	
tend			
earnings	√	√	√
thrilled	?	√	√
nip	√	√	√
muddy	√	√	√
bound	√	?	
race riots	√	√	√
fair share of trouble	√	√	√
crawls	√		?

S20 (ESL)

Item	Remember the Context	Know the Meaning	Use in other Context
aware	√	√	
bumping	√	√	
ploughed			
puncture	√	√	√
attaching		√	√
struck	√	√	
startled			
swerved	√	√	
screeched	√	√	√
jaw		√	√
thrilled	√	√	√
nip	√	√	
muddy	√	√	√
bound			
bother about	√	√	√
race riots	√	√	√
im looking forward to seeing you	√	√	√
regards			

S19 (ESL)

Item	Remember the Context	Know the Meaning	Use in other Context
aware			
bumping	√		
ploughed			

	Remember the Context	Know the Meaning	Use in other Context
puncture	√	√	√
struck			
swerved			
screeched	√		
riding			
startled			
suddenly	√	?	?
halt			

S15 (ESL)

Item	Remember the Context	Know the Meaning	Use in other Context
aware	?	√	√
bump			
ploughed	√		
puncture	√	√	√
attaching	√	√	√
swerved			
jaw		√	
struck			
thrilled			
regret			
although			
nip	√		
muddy			
bound			
rate	√	√	√
riots			
muggers	√		
mutual		√	
crawls			
looking forward			
seems			

S10 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
atestada	√	√	√
padecido	√	√	√
larga	√	√	√
colillas	√	√	√
envuelto	√	√	?
emparedados	√	√	√
ceniza	√	√	√
guardara las formas	√	√	√
mangas de camisa	√	√	√
señaló	√	√	√
se azoró	√	√	√
mascaba chiclé	√	√	√
cabello		√	√
odiaba	√	√	√
desenvoltura	√	√	√
cometido	√	√	?
amenazador	√	√	√
enfrentó	√	√	√
vocerío	√	√	√
se adelantaba	√	√	√
se acercó	√	√	√

S12 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
salita	√	√	√
atestada	√	√	√
padecido	√	√	?
apenas	?	√	√
colillas	√	√	√
emparedados	√	√	√
ceniza	√	√	√
flojos	√	√	√
se divisara			
desagradado	√	√	√
molestaba	√	√	√
formas	√	√	
nudos	√	√	√
mangas	√	√	√
seca	√	√	√
aludido	√	√	√
azoró			
caballeros	√	√	√
larga	√	√	√
espera	√	√	√
acusador	√	√	?
miradas		?	
mascaba	√	√	?
periodista	?	√	√
puesta			
se peinó	√	√	√
cabello	√	√	√
odiaba		√	√
desenvoltura			
demás			
advertirles		√	√
cumpliré		?	
cometido			
se seguía		√	
amenazador			
un hombrecillo		√	
dejadme			

S11 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
una mano	√	√	√
levantó	√	√	√
salita	√	√	√
padecido			
prensa	√	√	√
apenas		√	√
colillas	√	√	√
fusilamiento	√	√	√
atestada			
el suelo	√	√	√
envuelto	?	?	
ceniza	√	√	√
desagradado			
nudos	√	√	√
flojos	√	√	√
mangas	?	√	√

	Remember the Context	Know the Meaning	Use in other Context
un periodista	√	√	√
se divisara	?		
miradas	√	√	
chaqueta puesta	√	?	
aludido	?	√	√
se azoró			
mascaba	√	√	√
chiclé	√	√	√
se peinó	√	√	√.
revuelto			
dedos	√	√	√
odiaba	√	√	√
la desenvoltura	√	√	√
los demás	√	√	√
cumpliré	?		
cometido			
facilitando			
el fondo	√	√	√
se seguía			

S13 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
caballeros	√	√	√
la salita		√	√
atestada	√	√	√
padecido			
los corresponsales		√	
prensa			
colillas	√	?	√
envuelto	√	ɷ	?
emparedados			
ceniza	√	?	√
molestaba	√	√	√
las formas	√	√	
flojos			
aludido			
se azoró			
se peinó		√	?
lamento		√	√
desenvoltura			
los demás		√	√
cometido		√	√
advertirles			
dejadme	?	√	√
vocerío	?	?	
guardara las formas	√	√	
aguardaba		?	

S9 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
caballeros	√	√	√
salita	√	√	√
atestada	√	√	√
padecido		√	√
apenas		√	√
colillas	√	√	√
emparedados	√	√	√
ceniza	√	√	√
mangas	√	√	√

miradas	?	√	
puesta	√		
se azoró	?	√	√
mascaba chiclé naturalmente	√	√	√
se peinó	√	√	√
los dedos	√	√	√
revuelto	?		
desenvoltura		√	√
envuelto	√	√	√
demás		√	√
advertirles	?	√	?
cometido			
comportamiento			
desordenada	?.	√	√
gritó	√	√	√
dejadme			
aguardaba		√	
se acercó	√	√	√

S14 (SFL)

Item	Remember the Context	Know the Meaning	Use in other Context
caballeros	√	√	√
la salita	√	√	√
padecido		√	√
el suelo	√	√	√
le molestaba	?	√	√
flojos	√	√	√
mangas de camisa	?	?	
los nudos	√	√	√
colillas	√	√	√
envuelto emparedados	?	√	√
las miradas de los periodistas	√	?	√
seca	√	√	√
el aludido se azoró			
mascaba chiclé	√	√	√
se la puso			
se peinó	√	√	
los dedos	√	√	√
cabello revuelto	√	√	√
odiaba	√	√	√
desenvoltura en los demás			
cometido			√
comportamiento			
hubo			
se seguía	?	√	√
amenazador	√	√	√
gritó	√	√	√
se acercó			
se adelantaba			

Appendix III (validation of qualitative data):

D= assessed as DEEP
S= assessed as SURFACE
?= not sure

General Approach to Vocabulary Learning

Subject L1 & L2	Subject reference & explanatory comments	Subject's Comment	Pilot Judge	My Assessment	Independent Judge
S17 L1=Spa L2=Eng	(referring to some booklets on a rack)	..vocabulary is difficult when I'm talking..for example: these are 'leaflets' but some are 'magazines'..I try to use a general word..	D	D	D
..		..I try to repeat (write) new words on paper 5 times..it works	S	S	S
S18 L1=Spa L2=Eng		..If I don't write it (straight away) I don't remember...if I write it down 10 times in a notebook..it works!	S	S	S
		..I never use words which I don't know what they mean..with colleagues it doesn't have to be academic..for someone with more culture I am more careful..	D	D	D
S20 L1=Por L2=Eng		..I create words from Portuguese..but the pronunciation is difficult..it takes longer to use an (L2 dictionary) though I think it's better..one words pulls another..I try to use words in conversation..	?	D	D
S19 L1=Por L2=Eng		..I write them down then I put them in my mind..I'm quite good at remembering..	S	S	S
		..when I go to the toilet I take the English dictionary..it's not good to use a Portuguese dictionary..it makes confusion..you need to think in English.. you need to read books.. newspapers every day	D	D	D
S15 L1=Por L2=Eng		..no system for learning words.. sometimes I look in the (L2-L1) dictionary..I'm lazy	S	S	S
S16 L1=Fre L2=Eng		I have a note book for English and French..I check words and try to speak with those words...I (used to) (re-wite) them once a week..it is very useful to repeat all the time..after 6 months people were very surprised because I have improved..	?	S	S
		I used to know every irregular verb..I cover the columns and test myself..I never use auxilliary verbs..I know them and when to use but not how to use.	S	S	S
S10 L1=Eng & Fre L2=Spa		..I underline words then use the dictionary and write down Eng or French equivalent..if ..meeting word again..I remember writing it down I refuse to look up in the dictionary again.	D	S	S

		I don't like to jump into the dictionary.. I look up eventually.. always like to get a general meaning then go for the words..the next time I see a word - if I've struggled to get it I can use it confidently	D	D	D
..		..You can do a summary without knowing the meaning of 10% of all words..you can pick up a dictionary & look up - then you don't remember..	D	D	D
S12 L1=Eng L2=Spa		..lots of different texts..underline words..look them up..going back 3 or 4 times..hoping it will have' sunk in..word lists not very useful..	D	D	D
S14 L1=Eng L2=Spa		..I don't always rely on a dictionary.. there are lots of colloquial meanings.. get the general meaning..	D	D	D
S13 L1=Eng L2=Spa		..I try to get the root of the word..break it up..look for the infinitive..then you can know what the rest of it means..	S	D	D
		..It doesn't matter what individual words mean..I must read the whole text to get an idea of what this part means..go back to the words when I've established what the rest of the text is about..	D	D	D
S9 L1=Eng L2=Spa		..I tend to look for the gist first...then look up more words rather than fewer..	D	D	D
S11 L1=Eng L2=Spa		..the answers (vocabulary items) I just look up slip out of my mind the fastest..because I'm older than the others it takes longer for the forms to stick..	?	S	S

Approach to tasks in the interaction

a)=selection
b)=categorisation
c)=notes
d)=retrieval

Subject L1 & L2	Item context referred to, sub-task & explanatory comments	Subject's Comment	Pilot Judge	My Assessment	Independent Judge
S17 L1=Spa L2=Eng	'startled' b) classifying for form	..it's because it's a past tense verb	S	S	S
	'jaw' a) '..her jaw dropped..'	..it's not a verb..something on the face..the same for the animal in the film?	S	D	D
	'drive' a) '..the advert stresses initiative and drive..'	'drive'..is different from drive a car..	D	D	D
	'..race riots..' a)	..it's about foreign/ black people - I have to put it together..	D	D	D

xxiv

S18 L1=Spa L2=Eng	'attaching' a) '..attaching it to the tyre..'	..like attack?	D	S	S
	'startled' b) classifying for form	..it is past (tense)..	S	S	S
S20 L1=Por L2=Eng	'..fit into..' a)	'fit' means suitable..it has other meanings..it's a phrasal verb	D	D	D
	'ploughed' d) trying to guess	..I remember the tractor in the field. ..we have no way to say this in Portuguese..	D	D	D
S19 L1=Por L2=Eng	'startled' b) classifying for form	.. I put it in form because it's like 'start'	D	S	S
	'..find the work interesting and challenging..' a)	'challenge' is like 'discover'?	D	D	?
S15 L1=Por L2=Eng	'..there is a nip in the air..' c) writing notes	..can I use 'there is a nip from smoking'?	D	D	D
S16 L1=Fre L2=Eng	'struck' c) Looking up ('he struck a match') in the dictionary	..to 'hurt' a match?	?	S	S
	'riding' b) classifying for form	..because it is a small word..	S	S	S
S10 L1=Eng L2=Spa	'envuelto' (wrapped up) b) classifying for form	..part of 'volver'..shall I put it in form?	S	D	S
	'molestaba' (bothered) a)	..obviously not a pleasant word	S	D	D
	'colillas' (cigarette ends) a)	..not sure if its a conjugation or a noun..it's a noun ending..	S	D	D
	'señaló' (indicated/announced...also fix/settle) a)	..I knew I had seen it before..there's another meaning.. to fix..I remember the context of the other meaning but not this meaning here..	D	D	D
S12 L1=Eng L2=Spa	'acusador' (adjectival form accusingly) c) looking up in dictionary	..what part of language is 'acusador'?	S	D	D
	'el aludidio se azoró' (the person referred to was alarmed) a)	..he aluded to got alarmed..	S	S	?
S14 L1=Eng L2=Spa	'levantó'/'molestaba' a) `	..We've just done imperfect & preterite..levantó and molestaba catch my eye...	S	D	D
	'envuelto' a)	..a stem-changing verb with perfect ending	S	D	D
	'se sequia' (went on) c) writing notes	..I'll put 'pursue' because the verb is like the French.	D	D	D

			S	D	D
	'padecido' (suffered - from infinitive 'padecer') a)	..from 'decir'..looking at the root..pp. of 'decir'..	S	D	D
	'desagradado' (displeased) a)	..agradable means 'pleasant'..desagradado is the opposite..	D	D	D
	'fila' (row) a)	..I don't know it but it doesn't seem important..I think I can get away with it..only if it's going to help me with the text..	D	S	?
	'dejadme' (leave it to me) d) guessing it correctly	..I knew the English..and 'me'..& fiddled around with what the verb could be..things associated..guessed it began with 'd'	D	S	D
	'la larga espera que habian padecido' (the long wait they had suffered) a)	..the sentence is in my head..if it comes up in the test I'll remember it - without knowing what it means..		S	S
S9 L1=Eng L2=Spa	'caballeros' (gentlemen) d) trying to guess	..it was a formal speech and he had raised his hand..	D	D	D
	'emparedados' (sandwiches) in the test sentence for 'ceniza' (ash) d) trying to guess 'ceniza'	..is emparedados 'bury'	S	S	S
	'corbatas ...flojos' (loose ties) a)	..their ties were floral?	S	S	S
	'el ministro odiaba la desenvoltura' (the minister hated the causalness/ sloppiness) a)	..minister dared the desenvoltura.. ..minister hated the unwrapping..	S	S	?
	'dejadme' (leave it to me) a)	..that's an imperative..he's obviously quite irate..	D	D	D
	'sea' (subj. verb 'to be') a)	..that's a subjunctive..he's trying to influence him? ..all the time they've been trying to get a question at him..	D	D	D
S11 L1=Eng L2=Spa	'salita' (room) c) writing notes	..it sounds South American	S	S	S
	'habian padecido' (had endured) c) reviewing notes	..it used the perfect form with 'haber'	S	D	D
	'envuelto' (wrapped) a)	..it's part of last year's vocab..'muy basica'	S	S	S
	'constestar' (to answer) a)	..I can't contest the questions..	S	S	S
	'puso' (put - preterite of 'poner') c) writing notes	..I get muddled with poner.. pedir..podir	D	D	D

	'desenvoltura' (sloppiness) c) reviewing notes	..because he was picking them up on their behaviour..	S	D	D
	'revuelto' (untidy) 'cumplire' (I will fulfill) 'cometido' (task) d) trying to guess them	..my flatmates! ..the birthday word! .the dinner word!	S	S	D
L1= Eng & Ita L2=Spa	'guardara las formas' (observe the niceties) a)	..look after?..I'm thinking in Italian..	D	S	S

Self-assessment comments

Subject's L1 & L2	Item referred to & explanatory comments	Subject's Comment	Pilot Judge	My Assessment	Independent Judge
S17 L1=Spa L2=Eng	'pub-crawl'	can you say 'shop crawl'?	D	D	D
S20 L1=Por L2=Eng	'aware'	.I tried once but did it in a wrong way..I try to use this as 'knowledge'..in a composition	?	D	D
	'bumping' (from '..he became aware that the bicycle was bumping..')	..he aware the bicycle was bumping	D	D	D
S19 L1=Por L2=Eng	'aware'	..sometimes I can aware my English..?	D	S	S
	'jaw'	..I can use this but I say my mouth..	D	S	S
S16 L1=Fre L2=Eng	'pump'	I was pumping my bicycle when the phone ring..	D	S	D
	'jaw' (from '..her jaw dropped..')	.this (face)..the woman saw an accident, she was afraid, she screamed	D	D	D
S10 L1=Eng L2=Spa	'guardara las formas' (from '..le molestaba que....no guardara las formas..')	..would come in handy in an essay..it was subjunctive.. must have been a 'que'..	?	D	D
	'señaló'	I've been coming across it a lot in the news..it's press language..	D	D	D
	'desenvoltura'	..yesterday, in an article about drugs..	D	D	D
S12 L1=Eng L2=Spa	'emparedados' (from '..envuelto emparedados..')	'..revueltos..'	D	D	D
	'flojos' (from '..los nudos de las corbatas estaban flojos' - 'their ties were undone')	'..nudos..'	D	D	D

	'molestaba' (from 'le molestaba que no guaradara las formas' - 'it bothered him when the formalities weren't observed')	' no guardara las formas '	D	D	D
	'seca' (from '..con voz seca..' 'with a dry voice')	'..con voz seca..'	D	D	D
	'espera' (from: 'la larga espera que habian padecido' - the long wait they had endured')	..long wait they had suffered - padecer!	D	D	D
	'cabello' (hair) (from: 'la cabello revuelto' - 'untidy hair')	'..hair..greasy..'desenvoltura'?	D	D	D
S12 L1=Eng L2=Spa	'colillas' (cigarette ends)	..'el suelo' (floor)...cigarette ends	D	D	D
	'las miradas' (from '..las miradas de los periodistas..' the glances/gazes of the journalists)	..of the journalists.. 'periodistas' was one of the first words I learned in Spanish	D	D	D
	'cabello' (from 'se peino su cabello revuelto' - he combed his untidy hair')	..scruffy hair..'se peino'?	D	D	D
	'odiaba' (he hated)	..opposite 'gustar'	S	D	D
S13 L1=Eng & Ita L2=Spa	'las formas' (from 'le molestaba que no guaradara las formas' - 'it bothered him when the formalities weren't observed')	'..no guardaba las formas..'	S	D	D
	'lamentó' (he lamented)	.I knew what it meant so I didn't pay much attention to the context..	S	S	S
	'prensa' (the Press)	..similar to Italian ..'presa' to be in a hurry..means something different..	S	S	S
	'se peinó' (he combed)	'..pettinarse..'(Italian - to comb)	?	S	S
S9 L1=Eng L2=Spa	'emparedados' (sandwiches)	..I'd use 'bocadillos'..	S	S	?
	'puesta' (from 'la chaqueta puesta..' with a jacket on..)	.sus propios puestas..does it mean 'own'?	S	S	D