DEVELOPING LEXICAL COMPETENCE: A COMPUTER-ASSISTED, TEXT-BASED APPROACH

J. Lachlan Mackenzie* and Mike Hannay**
Vrije Universiteit Amsterdam, Netherlands

Abstract

The *Lexvaardig* project, developed at the Vrije Universiteit Amsterdam, aims to help advanced students of foreign languages to develop lexical competence by providing them with a body of annotated texts from which they can choose freely, an informal introduction to lexicology and a personal database. After an initial training in reception, production and resourcing skills, students improve their awareness of lexical issues by identifying and noting phenomena in texts of their own choosing: they explore different kinds of lexical unit, discover items with interesting phonological, orthographic and morphological characteristics, investigate the syntactic and semantic properties of words, and inquire into their various social, cultural and historical aspects. The project seeks to put many of the insights in contemporary lexicological research into practice, hoping thereby to give students positive motivation towards vocabulary learning. By instilling respect for the immeasurable complexity of the lexicon, *Lexvaardig* creates a constructive attitude to the task of acquiring a rich and flexible command of the target lexicon, while also giving the students an introduction to hands-on linguistic and textual analysis in an attractive computerized learning environment.

Key words

Advance organizing, applied lexicology, autonomous learning, computer-based workbench, Electronic Card Index, inferencing, lexical phrase, pragmatic formula, resourcing, word families.

* Lachlan Mackenzie is Professor of English Language at the Vrije Universiteit, Amsterdam. He teaches a range of subjects within his discipline, both practical and theoretical. His research is concerned with developing Functional Grammar as a theory of verbal communication and with the implications of functionalist linguistics for advanced language teaching.

** Mike Hannay is senior lecturer in English language at the Vrije Universiteit, Amsterdam. He teaches and does research in functional linguistics, translation and writing skills, and in addition coordinates faculty developments in ICT and language teaching. His publications include, as editor-in-chief, the concise Van Dale Dutch-English and English-Dutch dictionaries.

RESUMEN

El proyecto 'Lexvaardig', desarrollado en la Universidad Libre de Amsterdam (Vrije Universiteit Amsterdam), tiene como objetivo ayudar al estudiante de lenguas extranjeras de nivel avanzado a desarrollar la competencia léxica proporcionándole un corpus de textos anotados de entre los que puede elegir libremente, una introducción informal a la lexicología y una base de datos personal. Tras una preparación inicial en las destrezas de recepción, producción y la del uso de obras de referencia, el estudiante mejora su conciencia sobre las cuestiones léxicas mediante la identificación y la observación de fenómenos en textos de su propia elección: explora tipos diferentes de unidad léxica, descubre ejemplos con características fonológicas, ortográficas y morfológicas interesantes, investiga las propiedades sintácticas y semánticas de las palabras, e indaga en sus diversos aspectos sociales, culturales e históricos. El proyecto busca poner en práctica muchas de las ideas en la investigación lexicológica contemporánea, y con ello espera dar al estudiante motivación positiva para el aprendizaje del vocabulario. Al inculcar respeto por la inconmensurable complejidad del léxico, Lexvaardig crea una actitud constructiva hacia la tarea de adquirir un dominio rico y flexible del lexicón meta, al mismo tiempo que proporciona al estudiante una introducción al análisis lingüístico y textual práctico en un atractivo entorno de aprendizaje informatizado.

PALABRAS CLAVE

Aprendizaje autónomo, escritorio informatizado, familias de palabras, fichero electrónico, fórmula pragmática, frase léxica, fuentes de referencia, inferencia, lexicología aplicada, organización previa.

RÉSUMÉ

Le projet 'Lexvaardig', élaboré à l'Université Libre d'Amsterdam (Vrije Universiteit Amsterdam), a pour objectif d'aider l'étudiant de langues étrangères de niveau avancé à développer sa compétence lexicale en lui fournissant un choix de textes annotés, une introduction informelle à la lexicologie et une banque de données personnelle. Grâce à une réflexion préalable sur la réception et la production lexicales et une initiation à l'utilisation d'ouvrages de référence, l'étudiant se sensibilise aux questions lexicales en identifiant et observant des phénomènes pertinents dans des textes qu'il a lui-même choisis. Il explore différents types d'unité lexicale, découvre des exemples qui manifestent des caractéristiques phonologiques, orthographiques et morphologiques intéressantes, est amené à examiner les propriétés syntaxiques et sémantiques des mots, et à se pencher sur la dimension socioculturelle et historique de ces mots. Le projet met en pratique diverses idées et démarches qui proviennent des recherches lexicologiques contemporaines afin de créer chez l'étudiant une motivation positive à l'égard de l'apprentissage du vocabulaire. En inculquant le respect de la complexité incomparable du lexique,

Lexvaardig crée une attitude constructive qui permet l'acquisition d'une maîtrise riche et flexible de la langue cible. En même temps, l'étudiant est initié à la pratique de l'analyse linguistique et textuelle dans un environnement attrayant d'apprentissage informatisé.

Mots-clé

Apprentissage autonome, atélier informatisé, familles de mots, fichier électronique, formule pragmatique, inférence, lexicologie appliquée, locutions figées, organisation préalable, ouvrages de référence.

1. Lexical Competence

The purpose of this article is to report on a project that is being carried out in the Faculty of Arts at the Vrije Universiteit Amsterdam, The Netherlands. We will focus on the Department of English, but the project applies faculty-wide, specifically to the Departments of English, French and German; our report is thus, *mutatis mutandis*, equally relevant to our colleagues' work in French and German studies[1].

First-year students embarking on a four-year course of studies towards a degree in English Language and Culture already possess a basic command of English vocabulary. They generally have sufficient words and expressions at their active disposal to get by in everyday situations. Most of these beginners display considerable fluency and self-confidence, based on long-term exposure to English, which is a language that young people value highly and encounter very frequently in the Netherlands. Nevertheless, these students tend to operate under an invisible ceiling: their very linguistic aplomb, bolstered by compliments received on foreign holidays, blinds them to the fact that they are in fact continually recycling a rather restricted repertoire and thereby denying themselves expressive possibilities more typical of the competence of educated native speakers. A major goal of university tuition in vocabulary is to make the students aware of their self-imposed limitation on the use of English, and thereby to help them break through to enhanced lexical competence.

Skills are of course only one component of university training in a foreign language: beginning students not only make their first acquaintance with English literature(s), they are also introduced in their first

---

[1]   The other members of the team were: Tine Greidanus, Onno Huber, Marieke Krajenbrink, Freya Martin, Ria van Minnen, Floor Oosting and Maurice Vliegen.

year to the linguistic analysis of English. Traditionally, proficiency training and linguistics have been regarded as quite separate aspects of the curriculum, typically not even taught by the same tutors. Whereas it may be stated with some justification that there is little connection between phonological or syntactic theory and students' mastery of pronunciation or the production of effective sentences, the lexicon, we will be arguing, is an area *par excellence* where theory and practice come together. If students can be sensitized to the categories that have been applied in linguists' attempts to bring order to the lexicon, they will bring an informed mind to the seemingly monstrous task of acquiring thousands of new words and expressions within a few years. This is the thinking behind the project to be described.

The project (which is called *Lexvaardig*, from the Dutch words *lexicale vaardigheid* 'lexical skill') takes two crucial features of student experience with the acquiring the vocabulary of the target language as its starting point. Firstly, students have most of their contact with English outside that part of the curricular environment which is programmed for vocabulary learning, i.e. either in the reading or listening they have to do for other courses or in extramural activity (television, internet, English-speaking friends, etc.). As a result, much of their vocabulary learning is in point of fact incidental and autonomous. Secondly, the new words and expressions they are motivated to learn occur above all in texts which they themselves choose to read, ranging from literary works to magazines and websites. What the project does is to accept these facts as given, and to guide and structure the students' learning process by providing them with a computer-based package, consisting of three components (see §7 below): a body of annotated texts from which they can choose freely; an introduction to lexicology; a personal data base.

The project thus differs radically from attempts to impose word lists upon students, starting from the most frequent and general vocabulary and gradually working outwards to rarer or more specialized words. Rather, the students are trained to develop their own lexical competence, to become wordwise. They are given the wherewithal to think about and improve their own understanding of the English lexicon, and above all to become aware of what they encounter in their reading and listening: rather than guessing wildly what is meant, or simply disregarding unfamiliar lexical items, they now have the tools to analyse their difficulties and to learn from the experience for future occasions. And a second advantage that appeals to the students of today, each of

whom has her[2] own personal background experience of the English language, is the individualization of vocabulary learning: whereas the learner is subtly steered through the whole process, it is essentially the individual student who decides what she finds useful to learn. From a didactic viewpoint, this is no problem, because *what* is learned is less important than *how* it is learned. If the skills are there, the knowledge will come.

In devising the materials to be presented here, we have attempted to translate the major results achieved in the vocabulary acquisition research of recent decades, research into such matters as lexical knowledge, inferencing strategies and lexical phrases, into pedagogical practice that is appropriate to the context of advanced learners in the Dutch context. Accordingly, the following sections will concentrate on the practical implementation of the materials. We begin by sketching how the students' task is organized in advance (§2), going on in §3 to deal with the different kinds of lexical unit to which students must attend. The following sections detail the various types of knowledge about lexical units which are pointed out to students, covering phonological, orthographical and morphological matters (§4), the syntax-semantics interface (§5) and finally social, cultural and historical aspects (§6). §7 presents the 3 components of the computerized package that is placed at the disposal of all students, and leads to the conclusions (§8).

2. RECEPTION, PRODUCTION AND RESOURCING

Students who set out on the path to greater lexical competence must have clear insight into how their understanding of words is going to change. This 'advance organizing' (cf. Ellis, 1994, p. 537) orients them to the three activities of:

(a)   reception
(b)   production
(c)   resourcing

In the very first weeks of their university education, the students are given training in how best to organize these three activities, which will now be briefly dealt with in turn.

---

[2]   Since the great majority of our students are female, we will use the feminine form for generic reference to students.

Reception is presented as being in this context first and foremost a matter of identifying lexical units (see §3 below) and establishing their meaning in the clause in which they occur, i.e. as a function of the interaction with the other meaning-bearing words in that clause. But it is equally important to stress that lexical items are chosen in virtue of general communicative strategies and thus also make a contribution to the coherence and purpose of larger textual units such as sentences or paragraphs, and students need to be sensitive to this wider context of interpretation (cf. Hoey, 1991). In many poetic contexts, for example, all the potential meanings of a word are potentially in play, whereas in a legal document ideally only one sense will apply. These factors need to be considered even with vocabulary that is familiar to the student. However, where the student is ignorant of a word in a text, or notices something peculiar about an item she has already acquired, more skills are called upon.

In the context of *Lexvaardig*, students are encouraged not to grab the dictionary as soon as they encounter an item they do not know. Rather, they should go through a number of inferencing strategies. They are encouraged not only to identify the word class(es) involved and to understand words and expressions in their clausal environment, but also to study the preceding and above all the following context, determining for example whether the item returns, or whether the context might contain a synonym or some other clue. Only after all these stages have been traversed and they have made an intelligent guess about the form and its meaning should the students look up the item in a dictionary, to see if their guess is right; if it is not right, they should try and work out where and why their reasoning failed. Students who work in this way find that the extra effort put into 'researching' the item aids the memorability of items and strengthens their ability to employ inference in the understanding of text.

Production activities involve skills in selecting lexical items which will effectively convey the speaker/writer's intended meaning. At the same time, the items should be used correctly as to spelling or pronunciation. Alongside these rather obvious requirements, students also need to become sensitive to the particular combinatorial properties of lexical items within the clause, as well as the ways in which the use of words contributes to textual cohesion or is determined by genre conventions. These matters will return at greater length in §5 and §6.

Just as with reception, production often confronts the student with gaps in her knowledge. *Lexvaardig* draws her attention to various com-

pensation strategies that allow her to retain her fluency in speech or in writing by using lexical dummies (words like *thingummy* in speech or, say, *phenomenon* in writing), superordinate terms (*piece of furniture* to cover ignorance of the word *couch*), or paraphrases (*advise someone not to* replacing the unknown *discourage*). And by paying attention to coinages in the texts that they read, students will gain a feeling for the circumstances under which they, too, can dare to create a new word.

The third activity, which –following Chanot (1987)– we call 'resourcing', relates to the use of the many resources that are available to today's learner. Freshmen generally have limited experience in the use of dictionaries, and have at best incidental knowledge of what Internet has to offer. As we will see in greater detail in §7, *Lexvaardig* makes use of a computerized environment which provides students with all the tools they need to structure their active learning: annotated texts, with structured assignments, an electronic handbook on applied lexicology, and an 'electronic card index' in which they can note all their discoveries. They also have various electronically accessible monolingual and bilingual dictionaries at their disposal.

Although dictionaries play an indispensable role in language learning, students have generally received little guidance at school on their use. Part of the freshmen's preparatory training is therefore oriented to helping them take maximum advantage of the various types that are available (cf. Scholfield, 1997). Particularly with highly polysemous items, students need a lot of explicit help with deriving information from dictionary entries. Electronic dictionaries also have extremely useful search functions but the use of these, too, needs to be taught and practised.

Armed with this background understanding of the three activities they are about to carry out (reception, production and resourcing), the students can now turn to the task of identifying the phenomena they will be studying. It is this task that is described in the following section.

3. LEXICAL UNITS

Students have a strong tendency, which is unfortunately not undermined by most dictionaries, to equate the lexical unit with the word. Any course in lexicology must start by drawing students' attention to the fact that the individual word is only one of the forms in which a

meaning unit may be packaged. What they will be learning are lexical units (cf. Cruse, 1986). These can be of any length or complexity, and may involve 'lexical phrases' in the sense of Nattinger and DeCarrico (1992):

black box	adjective + noun
on a …. basis	prepositional phrase
spick and span	binomial expression
Beats me.	clausal fragment
The more … the more …	clause-linking formula
Like father, like son.	proverbial sentence
Our Father, which …	
for ever and ever. Amen.	entire text (The Lord's Prayer)

Due attention is given to idiomatic expressions, and in particular to how there are limitations on how these can be built into the clauses that contain them. Thus students do not only learn that *take one's hat off to* means 'praise, express one's respect for', but are also encouraged to discover how the expression is used. Thus they may find out that it is usually used in the present tense with a first-person subject to perform an act of praise, as in *I take off my hat to him*, and that it is generally not found in the past tense and never in the passive voice: ?*I took off my hat to him*; **My hat is/was taken off to him*. They cannot make these discoveries by consulting a dictionary: for this, they need to consult a corpus or to try out different forms on a native speaker. Similarly for the discovery that *barking up the wrong tree* cannot occur in the simple aspect. From a very early stage, students feel like language researchers: they realize that a great deal of what can be known about the language they are learning is not codified in existing textbooks and works of reference, and that studying at an academic institution involves them from the outset in a process of formulating and testing hypotheses.

Students need to extend their vocabulary for two purposes: to enhance their understanding of texts, and to improve their own active use of the language. It is in particular with the latter purpose in mind that their attention needs to be drawn to the structure-building devices that are so important for both the management of conversations and the production of written texts. Since students will tend to "read over the top" of these items, they are encouraged to observe and note such opening gambits as:

Note that …
It cannot be denied that …
A further point is that …

as well as more typically spoken-language formulae like:

I hope you don't mind me saying so, but …
What do you mean (my cooking is awful)?
That reminds me, (I haven't done my shopping yet).

Another category to which their attention is drawn is the 'pragmatic formula'. These are rather more characteristic of the spoken language, and are fixed expressions which have a certain recognized illocutionary status in social interaction: exclaiming *Bless you!* when someone sneezes is an obvious example. Social interaction is of course heavily influenced by cultural conventions: it may well be that even *Bless you!* is not appropriate in all situations where in Spanish-speaking cultures one might exclaim *¡Jesús!*. Students are, for example, asked to think about the various expressions used during a transaction in a Dutch shop, and thereby identify the pragmatic formulae used in their own culture. The challenge is then to discover how a comparable transaction would be carried out in an English-speaking environment. One component of a Dutch transaction is the optional formula *Het is een cadeau* (literally, 'It's a present'). In the Dutch context, this is said by a customer as a request that the salesperson giftwrap the purchase (for free). In an English-speaking context, there is no such comparable formula: one could say *Could you perhaps wrap it up for me, I'm giving it as a present*, but this is not a standard formula, and it is unlikely that the shop will offer the service free of charge anyway.

Each lexical unit noted by the student will be classified according to its status as a simple or complex lexical unit, a proverb, a structure-building device or a pragmatic formula.

4. PHONOLOGY, ORTHOGRAPHY AND MORPHOLOGY

Since the students receive their training in lexicology at the very beginning of their studies, they have not yet had any tuition in phonetics and phonology. Nevertheless, particularly with a language as orthographically whimsical as English, students will often need to note the pronunciation of items. At this stage, we tolerate the use of informal

representations, e.g. showing the two pronunciations of *garage* as 'GAr-ridge' and 'garAHZH', using capitals for the stressed syllable.

Students tend to have some intuitive understanding of variation in English, and are certainly vaguely aware of the British and American norms. They are advised to decide at an early stage in their career which of these two norms they wish to adopt (although other forms of English are also permitted), and thereafter to be consistent in that choice. This accordingly applies to every aspect of the individual student's language use, including one that is particularly relevant for *Lexvaardig*, spelling. Where an item has more than one orthography, students must be aware of this, and of which alternative is appropriate for their chosen norm.

Of particular value for students' developing lexical competence is a relatively informal introduction to morphology. They come to see that unfamiliar words can often be broken down into smaller units, the meaning of which can be ascertained more easily: thus a student encountering *tensimeter* for the first time may recognize the morpheme *meter* and then suppose (correctly) that the rest of the word (*tensi-*) is connection to the familiar word *tension*. This experience will also help her with more complicated instances like *tachometer*. Compounding is a morphological process well-known to the Dutch native-speaker, to whom a word such as *onderzoekszwaartepuntbijeenkomst* 'meeting of the research focus group' gives no problems of understanding. This word is composed of the elements *onderzoek* 'research', *zwaartepunt* 'focus' and *bijeenkomst* 'meeting', each of which is itself internally complex, involving affixation and compounding. As the example shows, complex lexical items are written in Dutch as one word; the learner's problem with English is that there are three possibilities, exemplified by the spellings *oil well, oil-well* and *oilwell*, all of which can be encountered in texts. Students tend not to notice morphological structure until they are confronted with unfamiliar instances. They are therefore introduced at an early stage to such surprising phenomena as back-formation (*televise* from *television*; *disinform* from *disinformation*) and blending (*smog* from *smoke* and *fog*; *stagflation* from *stagnation* and *inflation*). The effect is to increase their awareness that linguistic processes can take place within the word.

One major problem in the acquisition of vocabulary items is the sheer immensity of the task. In the case of morphologically complex words, that very complexity can be used as a basis for recognizing 'word families' (Bauer and Nation 1993), thereby easing the learning

task. Thus, departing from *nation*, the student can explore its family tree, identifying and grouping such kindred forms as *national, nationalize, nationalization, nationalist, nationalistic, nationalism, nationality,* but also, on other branches of the family, *nationhood* and *nation state*. And she can go on to compare the family structure around, say, *capital*. Additionally, she can select one affix, for example *–ism*, and discover how it applies in such forms as denominal *hooliganism < hooligan, Buddhism < Buddh(a), Platonism < Plato(*n)*, or de-adjectival *realism < real, pragmatism < pragmat(ic)*, and the neologisms *ageism* and *disabledism*. As ever, the purpose of the project is to stimulate the student's curiosity, and thereby to motivate her to set about the acquisition of vocabulary as a voyage of discovery.

Thus, she may develop an interest in clippings (which are very popular with the young and trendy in the Netherlands) and put together her personal collection of English ones. She will start with the well-known *lab, exam, vet, demo, pub, telly*, etc. and then use the resources available to her to discover new ones, attempting to establish which are generally established in the language and which have been created as nonce-forms for some short-term effect.

This informal training in the pronunciation, spelling and internal structure of lexical items is enough to get the students started on their research. As they encounter more and more aspects of the phonological and morphological structure of English, so their appetite will be whetted for the more precise, formal treatment they will receive in later courses in linguistics. The same applies to the initial discussion of the syntactic and semantic properties of words, to which we now turn.

5. SYNTACTIC AND SEMANTIC ASPECTS

After years of learning vocabulary at school by memorizing rough and ready equivalents in their native language (for that unfortunately continues to be the reality), students have a natural tendency to approach the meaning side of form-meaning correspondences through translation into the native tongue. Their training in lexicology therefore contains a treatment of the notion of equivalence. Where there is complete equivalence, the translation method can do no harm: consider the names of cities (German *Genf*, French and Dutch *Genève*, English *Geneva*, Italian *Ginevra*, Spanish *Ginebra*, etc.) or numerals. In other cases,

translation is less useful or even misleading. Thus in kinship systems, we often find partial equivalence: Dutch *neef* means 'nephew', but also 'male cousin'; Spanish *padres* can mean either 'fathers' or 'parents'. And languages can also divide up the human body in different ways: Russian *ruka* refers to one part of the body, from the shoulder to the fingertips, for which neither 'hand' nor 'arm' is a perfect equivalent in English; Dutch has separate words (*tand* and *kies* respectively) for incisors and molars, but no word for 'tooth'. Finally, in cases where there is no equivalence, seeking a translation is pointless. Incommensurable legal systems, for example, have terms for which there is no counterpart in other systems. Thus there is no translational equivalent for English *barrister* in other languages; indeed, even within the English-speaking world, legal terms often lack counterparts. Other examples of items for which other languages typically lack an equivalent are English *bedside manner* and German *Fernweh* (approx. 'nostalgia for faraway places').

Given the limited validity of the equivalence approach, students are encouraged to think of English lexical items operating within the semantic fields of the language itself. This entails introducing them to such lexicological notions as hyponymy, synonymy and antonymy. Armed with these notions, they are ready to tackle such issues as the relations among such sets as {*pay, remuneration, income, salary, wage(s), fee, honorarium, …*} or {*gaze, stare, gape, behold, eyeball*, …} or {*accept, acknowledge, admit, concede, confess, …*}.

Where verbs are grouped together into a semantic field, as with the last-mentioned set in the previous paragraph, questions of complementation often arise. The students' attention is therefore drawn to the interconnection between syntax and semantics, and they are encouraged to check in corpora, and/or in corpora-based grammars which of the complementation types (*that*-clause, *to*-infinitive, -*ing*-clause, etc.) is available for each of the verbs:

accept	*that*-clause
acknowledge	*that*-clause or –*ing*-clause
admit	*that*-clause or –*ing*-clause, or prepositional complement *(to* V-*ing)*
concede	*that*-clause
confess	*that*-clause or prepositional complement *(to* V-*ing)*

More generally, it is important that students understand the implications of the fact that they are learning lexical items for use in clauses. Especially but not exclusively in the case of verbs, part of know-

ing a verb is knowing whether it has particular complementation properties, but also whether it requires a specific preposition (*approve of, adhere to, react to, discriminate between,* etc.) or has unexpected selection restrictions. Students are given an informal introduction to the concept of argument structure: for each new lexical item they learn, they will be expected to ascertain the number of arguments associated with it; this also applies to familiar items such as *bet*, for which up to four arguments can be found: [*I*] *bet* [*you*] [*twenty pounds*] [*my horse will win*]. Even with little more than a vague inkling of the syntactic structure of English clauses, students will in this way gain a feel for the intricate, but very real relations between syntax and semantics.

Another recurrent connection between form and meaning, but one that is impossible to systematize, is the matter of collocation. This is the area *par excellence* where the kind of awareness training that our students receive comes into its own. This phenomenon abounds in those expressions in which words, typically adjectives or verbs, are used in a mildly metaphorical manner. English *weak coffee* is lacking in strength, Dutch *slappe koffie* 'limp coffee' lacks rigidity, German *dünner Kaffee* lacks density, French and Italian *café léger* and *caffè leggero* lacks weight, while Spanish *café claro* possesses clarity. And whereas English and French *take/prendre* a photograph, Dutch, German and Italian 'make' one (*maken, machen, fare*), and Spanish 'takes one out' (*sacar*). The student needs to become aware of the kind of context that gives rise to collocational constraints and to be constantly on the look-out for their occurrence.

By sensitizing the student to the various ways in which the meanings of individual lexical units interacts with the construction of clauses, we persuade them of the inadequacy of the vocabulary leaning technique to which their school education has accustomed them, that of the single translational equivalent. At the same time, we lay the foundation for a deeper understanding of the role of the syntax-semantics interface in everyday linguistic practice.

6. SOCIAL, CULTURAL AND HISTORICAL ASPECTS: THE SHAPE OF ENGLISH

Students have an almost intuitive understanding of the fact that social and cultural norms impinge upon their use of their native language. They may be sensitive about or indeed particularly proud of their own accent, and will certainly have opinions about other people's

use of the language. They will need little persuasion that it is in general possible to distinguish various 'levels' in their vocabulary: they will have no difficulty, for example, in recognizing neutral, formal, informal and vulgar words for, say, 'stupid' in their native tongue; similarly, they will agree that foreign learners of their language often commit errors in this regard, using either a ridiculously formal word or a scandalously informal word in a neutral context.

On this basis, students can be readily persuaded that the lexicon of the language they are learning is similarly stratified. As learners of a foreign language, they will want at least to master the neutral stylistic level in their own language production, but may also choose to affect certain characteristics of particular groups to which they feel attracted, for example vocabulary items associated with (certain) students, such as *rag, prog* or *frat*. But, in their interpretation of what they read and hear, they will also need to become aware of how lexical choices reflect the social embedding of the language user in terms of socioeconomic grouping, educational background, age, gender, ethnicity, etc. Dictionaries often give useful information about this, which students will do well not to ignore; here, too, guidance is essential.

Another component of most university degrees in English remains the study of the older forms of the language, from Old English onwards. Here, too, the practical and theoretical study of the lexicon provides vital sensitization to the mutability of language. Rather than giving up on a word encountered in a text but not found in the dictionary, the student must consider the possibility that she has stumbled across a neologism; alternatively, the expression in question may be an archaism – by researching either possibility, she will learn something of how linguists are tracking current developments and have registered the history of the language (cf. Bauer, 1994).

The student also comes to see how the past lives on in the present through an introduction to the notion of etymology. Many of the strange facts about language become more comprehensible if understood in terms of their origins. Thus an apparently obscure expression like *run the gauntlet* 'to suffer public humiliation' becomes 'motivated', i.e. understandable and memorable as soon as its origins are clear: it derives from a military or naval form of punishment from the seventeenth century in which the victim had to run between two lines of soldiers or sailors armed with canes. This may help the student remember *run*; but to remember *gauntlet*, she will also need an understanding of folk etymology, which involves a reshaping of foreign words to

approximate (or in this case become identical) to an existent form: *gauntlet* is an adaptation of Swedish *gata* ('way, passage'; cf. dialectal *gate* 'street') and *lopp* 'run'. More generally, students can be encouraged to do what comes naturally (and what goes wrong in the case of 'false friends'): to seek analogies between unfamiliar forms in the language they are learning and forms in languages which they already know.

The effect of attending to all the various parameters discussed above (in practice, even more are submitted to the students for consideration) is to increase consciousness of what we might call 'the shape of English', the various recurrent properties of the language that provide it with its own distinctive contour. Students are asked to be on the *qui vive* for these phenomena as they put their growing lexicological awareness into practice. The following is a random listing of 'typically English' structural characteristics to which students are encouraged to pay heed:

(a)	Latinate vs Germanic vocabulary	*tolerate/put up with; postpone/put off*
(b)	phrasal and phrasal-prep. verbs	*put up, see off, sound out; go in for, do away with*
(c)	complex participial forms	*tin-roofed house, boulder-strewn stream, Mozambique-based insurgents*
(d)	incorporation	*bear-hunting, shoulder-charge* (noun and verb), *bad-mouth*
(e)	compound premodifiers	*40-member committee, student-government president*
(f)	binomials	*spick and span, part and parcel, goods and chattels, sink or swim*
(g)	noun > verb conversion	*to critique, to impact, to inference*
(h)	Germanic verb > noun conversion	*(take) a walk, (have) a lie-down, (give someone) a push*
(i)	acronyms	*quango, laser, modem, radar*

| (j) | disjuncts | *surprisingly, importantly, crucially* |

The purpose of the preparatory tuition in what we might call 'applied lexicology' is to prepare students to set about the task of enhancing their lexical skills in the broadest sense. They are given a set of concepts which will enable them to think analytically about the texts they encounter. These texts will now appear as data, the various dictionaries and works of reference as equipment, and –perhaps most importantly– they will feel like researchers, a feeling that will be strengthened by the inevitable fact that they will from time to time be making discoveries that are new not only to them, but also to their teachers and to the various reference books. The work they do will thus not only be of practical benefit: from the very beginnings of their university education, they will know they are doing academic work.

Let us now turn to the *Lexvaardig* package, the computer-based workbench which is an integral part of the students' equipment.

7. *Lexvaardig*: The Package

The package consists of three separate but integrated products, all of which have been developed at the Vrije Universiteit Amsterdam and are accessible to the student on and through the Faculty computer network. In addition, other materials needed for assignments, such as monolingual and bilingual dictionaries, a concordancer and various corpora, are also available on the network or through Internet, and students are encouraged to have these open while working with the Lexvaardig products.

The first is an **Electronic Text Bank (ETB)**, currently consisting of some 120 short texts in English (and comparable numbers of texts in French and German). These are drawn from some seven genres, which have been selected to appeal to student users: (a) mass media, (b) governmental and corporate texts, (c) scientific and scholarly texts, (d) literary texts, (e) private communication, (f) legal texts, (g) texts set to music. The texts have been chosen to be attractive to the users' age group, readable, and rich in relevant lexical phenomena. The user can access and select the texts in various ways: (a) by genre, (b) by alphabetically ordered title, or (c) by subject matter (a text about 'love' could for example fall under genres a, d, e or g).

The texts have been annotated in the sense that a number of words and expressions that exemplify one or more of the lexicological issues discussed above are provided with a hypertext commentary; most annotations also contain an assignment, which directs and activates the students' interest. However, only some of the relevant material in the text has been marked in this way: it is equally important that students should learn to identify phenomena without explicit guidance.

The student also has access to a number of commercial products which she can use in doing her assignments: a concordancer, which applies to various corpora available on the faculty network and to high-quality bilingual electronic dictionaries (English-Dutch and Dutch-English); the student can also access the web, and is expected to note her findings in the ECI (see below), which is also linked to this screen. The student can be asked to (a) find examples of a particular phenomenon in a text; (b) search for more examples of a phenomenon in the electronic dictionary (which allows various types of search); or (c) discover collocations, frequencies, etc. by using the concordancer to extract data from corpora.

The second product is an **Electronically Accessible Handbook (EAH)**, which in printed form contains some 100 pages of background information on the phenomena sketched in sections 4 to 6 above. The first part of the handbook contains a systematized overview of the various dimensions that must be brought to bear on the understanding of a lexical unit. The second part is a practical guide on the use of dictionaries, both monolingual and bilingual. The whole of the EAH can be accessed directly from the ETB, and there are specific links to the relevant parts of the EAH from each annotation, so that students puzzled by a reference in one of these annotations to, say, 'clipping' or 'polysemy' can immediately check the meaning and use of these terms.

The third product is perhaps the most distinctive. This is the **Electronic Card Index (ECI)**, a facility which allows students to compile their very own, personalized database. This, rather than the dog-eared jotter of yesteryear, is where they store their word-wisdom. The ECI is a computer program structured in such a way as to integrate seamlessly with the various categories set out in the EAH: there is a place in the system for anything a student wishes to note about a lexical unit (pronunciation, spelling, complementation, socio-cultural constraints, etc.). The student's ECI is a totally individual record of lexical phenomena which she has found worth noting. She continues to add to it throughout her studies. The ECI program can be copied onto her computer at

home (if she has one), and she can keep her personal ECI on a diskette, adding to it either in the faculty or in her private quarters.

It will be clear that *Lexvaardig* fits into the philosophy of autonomous learning: the student is set free to make her own discoveries and to put together her own report on those discoveries. This process stimulates her lexicological curiosity and thereby increases her ability to acquire a more flexible lexical competence through a lasting process of conscious reading and listening. The framework for the learning is of course pre-structured in various ways: the developers have selected the texts in the ETB, they have written the EAH and they have created the structure of the ECI; what is more, the individual tutors establish certain minimum requirements in terms of the number and variety of assignments, for example that for every 50 ECI entries, at least 10 must mention collocation, at least 10 (not necessarily an entirely different 10) must pay attention to lexicosyntactic issues, and at least 10 must contain a description of the sociocultural conditions on use.

The tutor's role in this learning process becomes that of a mentor. The student reports regularly to her tutor, who takes a sample from her ECI, testing whether she has satisfied the minimum requirements but above all suggesting new parameters which she might consider worth noting: an unsuspected etymological connection, or a peculiarity of complementation that has escaped her attention. During these tutorial sessions, which are one-to-one, it becomes apparent that students tend each to have their own learning style: some have become almost obsessed with one particular phenomenon, e.g. adjectives with fixed prepositions (*fond of, similar to, …*); some draw all their examples from one kind of text (one student took all his from rock lyrics); some scour texts for words of great rarity, as if to challenge their tutor to know them all. But these are extremes, which gentle persuasion can usually cure. Generally, the ECI works well, as students come to see the intellectual reward that repays the extra effort they must put in, learning to operate the program, selecting the relevant parameters per word and then researching each lexical item.

8. Conclusions

Despite being central to the acquisition of a foreign language, vocabulary learning has had a bad press with students. Most beginners start out with an ingrained dislike of learning words, not least because

many schools have not progressed beyond demanding the memorization of unstructured lists. The purpose of the *Lexvaardig* project outlined above is to instil respect for the immeasurable complexity of the lexicon and concurrently to provide them with the wherewithal to set out on their own personal voyage of discovery through that tangled jungle. Vocabulary acquisition is a slow, gradual business, which continues throughout one's lifetime. If the four years of university study are also years of regular, informed research into how words and phrases are used, those four years will have created a constructive attitude to the task of acquiring a rich and flexible command of the target lexicon.

APPENDICES

Figure 1.   The screen through which students choose which features of a lexical unit they wish to record. This screen is in Dutch, since it is designed for students of all foreign languages.
The screen preselects four categories: lexical unit, core words, language and type of unit
The student may further choose from:
Meaning: definition; translation
Meaning extra: example of use, translation of example, collocations, synonym, antonym, hyponym, superordimate, connotation, cultural specificity
Lexical grammar: grammatical peculiarities, selection restrictions
Conditions on use: geographical, language level, social, jargon, chronological
Formal characteristics: spelling, pronunciation, formal variants, morphology
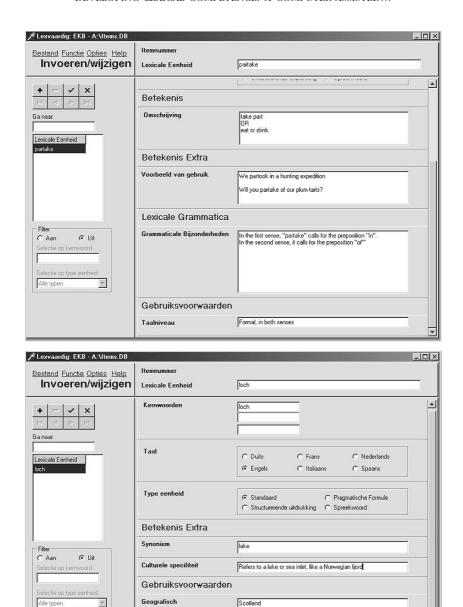Other: personal remarks, etymology, frequency, to be learned

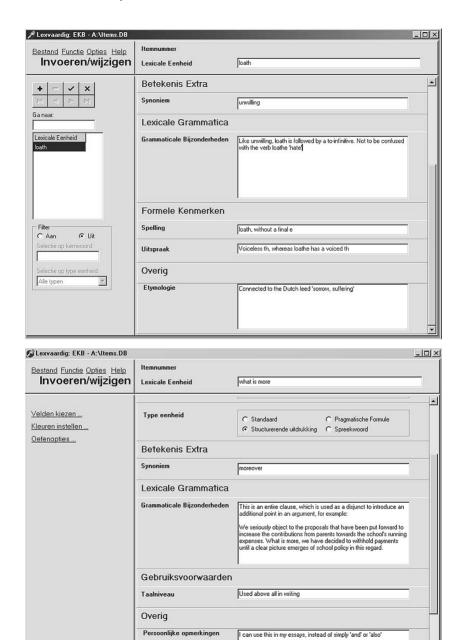Figure 2-6.   Examples of pages from a student's ECI.

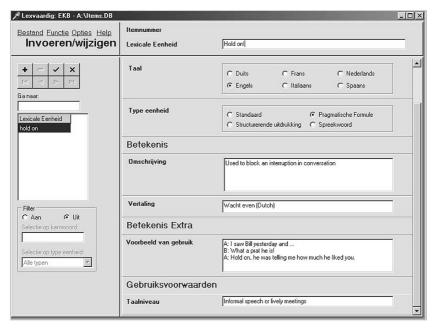Figure 2-6. Examples of pages from a student's ECI.

Figure 2-6.   Examples of pages from a student's ECI.

REFERENCES

BAUER, L. (1994), *Watching English change,* London & New York, Longman.

—, NATION, P. (1993), "Word families", *International Journal of Lexicography,* 6, pp. 253-279.

CHAMOT, A. U. (1987), "The learning strategies of ESL students", in A. Wenden and J. Rubin (eds.), *Learning strategies in language learning,* Engelwood Cliffs NJ, Prentice-Hall, pp. 71-84.

CRUSE, D. A. (1986), *Lexical semantics,* Cambridge, CUP.

ELLIS, R. (1994), *The study of second language acquisition,* Oxford, OUP.

HOEY, M. (1991), *Patterns of lexis in text,* Oxford, OUP.

NATTINGER, J. R.; DECARRICO, J. S. (1992), *Lexical phrases and language teaching,* Oxford, OUP.

SCHOLFIELD, P. (1997), "Vocabulary reference works in foreign language learning", in N. Schmitt and M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy,* Cambridge etc., CUP, pp. 279-302.