

## TESTING AND MASTERY LEARNING OF ENGLISH VOCABULARY AT UNIVERSITY LEVEL

LEOPOLD K. ENGELS

*University of Leuven*

The Belgian education system allows all students who have finished their secondary school curricula successfully to enter higher education and start studying in any faculty or college of the human sciences. Consequently all these institutions of higher education apply a thorough selection process at the end of the first year, so that more than fifty per cent of the students do not succeed, become repeaters or shift to another faculty, or to another university or college of higher education.

This has been going on since the authorities decided to 'democratize' higher education, providing a fair chance to any intelligent student from any social background. The high number of failures has alarmed educational authorities, parents and finally also the Universities themselves, so much so that the investigation of the causes of the numerous failures has become a serious concern of the University Staff.

But one can express one's anxiety in different ways; the easiest solution is to shuffle all responsibility onto the shoulders of secondary education; another evasive argument claims the incapacity, irresponsibility and laziness of the students of today; to provide a closer argument others claim that the study-burden is getting heavier every year and that professors do not take into account the fact that there are limits to the student's capabilities and endeavour. Finally there are some groups which have become interested in the *learning problems* of the student and which have started some research in order to provide efficient help when, where and to whom it is needed.<sup>1</sup>

All the Universities of the Flemish part of Belgium (K.U. Leuven, U. I.

---

<sup>1</sup> I.P.E.K. — project (Interuniversitair Project voor Efficiënter Kandidatuuronderwijs). Sponsored by the Belgian Board of Education 1978–81..

Antwerp, R. U. Gent, V. U. Brussel, L. U. C. Hasselt) have started a multi-disciplinary research programme on the ways in which computer-terminals can guide and improve the students' study-behaviour. We do not intend to develop C.A.I. (computer aided instruction), i.e. we do not try to find a substitute for their *instruction*, which is being provided in different forms by the teaching staff ('Vorlesungen', seminars, work-shops, etc.). We try to give the students a study-method by presenting problem-solving activities in several disciplines, so that they are enabled to pass their tentamina and exams more successfully. We also let them do exercises, which contain a *transfer* on learnt material, or an *increase* of their insight and skills when they want this, according to the results of *pre-tests* at the beginning of their university curriculum.

At the K. U. Leuven there are several teams in the project: physics, chemistry, pedagogy, also Dutch and English-language learning. We shall describe the part we play in this project as applied linguists.

We intend very humbly to improve *discrete points* (vocabulary and grammar), but presented in *wholistic entities*, i.e. a number of texts<sup>2</sup> (Allen and Davies 1978). In this paper we give the results obtained until now by means of a *pre-test* on English vocabulary, and an attempt to improve this vocabulary by means of group-work round computer-terminals, which evaluate the *exercices* and the *tests*, after every mastery-learning period with six texts.

### 1. The Pre-Test: English vocabulary (productive mastery = recalling) at the beginning of university studies

We should be able to define the vocabulary that we test. The students come from different schools all over the Flemish country. The only reliable criterion is the objective frequency of the English words in a descending ranklist. We do not know how far we can descend in that list, but we think that the *common core* of words that can be encountered in a written and written-to-be-spoken corpus of the English language will be known by Flemish students at the end of their secondary school courses, because in all schools written texts (literary and/or journalistic) prevail as the basis of a non-intensive teaching period of 6, 5 or 4 years. This common core consists of the 2000 first *lemmata* of an objective frequency-list, which was taken from a merger of written American

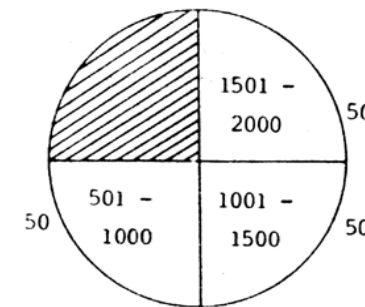
<sup>2</sup> We do not object in any way against *wholistic testing*, or against more creative exercises by means of role-play, discussions, games, etc. As a matter of fact our students do get language training of that kind. An evaluation of them is, however, not yet possible on computer terminals, whereas the more individual *Mastery Learning* of discrete points of language can be measured and evaluated objectively. We are, accordingly, convinced that future language specialists and teachers need a large receptive and productive vocabulary, and a fair command of the syntactic rules of the language combined with their numerous functions.

English, written British English and written to be spoken British theatrical language (the three corpora are of equal weight; they contain about 1,035,000 running words *each*).<sup>3</sup>

Until now we have tested the most frequent '*word-forms*' of the lemma (i.e. it may happen that one of the marked word-forms happens to be more frequent than its unmarked *lemma*) with their most frequent meaning (West 1977).

We have not tested the tokens surrounding each *head-word*, although we are convinced (Van Parreren, Detering, Denninghaus 1979) that our students also master the inflected and conjugated forms of the word, including also its common derivatives (prefixes & suffixes), when the meaning of the *head-word* remains constant. We do not test the 500 first *head-words*, because they consist mostly of *structural words*. Later on we shall include many of these words when we have detected the objective frequency of phrasal verbs, adjectives or nouns. We shall then add the collocations to the lemmata of the *head-words* as we have already done for other tokens (derivations, inflected forms).

We have divided the 1500 remaining lemmata into 3 regions of 500 in descending frequency order.



The students get one or more sentences (context) for each of the 150 chosen items. There is a blank for the word to be filled in (CLOZE-procedure). But in order to get the word we want, the translation of it is given underneath as a prompt. Acceptable synonyms have been taken into account in the feedback.

The results of the pre-test measuring the vocabulary-control of our students has been further examined:

1. It suffices to let the computer choose 50 randomized items from each of

<sup>3</sup> Formula of the difference coefficient according to Yule (1944)

$$\frac{F_{\text{Brown}} - F_{\text{Theatre}}}{F_{\text{Brown}} + F_{\text{Theatre}}}$$

Cf. also Geens (1878) and Johansson (1980).

the three blocks:    501—1000  
                           1001—1500  
                           1501—2000

to get a reliable test-corpus.

In 1978 we found (Van Beckhoven and Morris 1978) that a sample of 50 randomized items in a block of 500 objective F-words in descending order suffices to get a test with the same degree of difficulty for every *group of subjects* taking the test. We applied an intrinsic criterion nl. a *t*-test on 2 samples; the value for *t* should be smaller than 2.101, if the samples do not differ significantly; our hypothesis was confirmed because *t* was as small as 1.40. We also asked judgement of 15 students and 5 teachers to get an extrinsic criterion: these results correlated closely with the intrinsic criterion.

2. We have submitted about 800 subjects to the test each time we took it<sup>4</sup> (i.e. K. U. Leuven, KULAK, UFSAL, VUB, etc.). The mass of the test subjects also guaranteed the number of times each item of the test was randomly chosen by the computer.

The analysis of the answers allows us to detect the words, which can be excluded from our future tests (transparent words e.g.) without changing therefore the objective ranklist, because it must also fulfill other purposes (analyser of texts). If we left those out of our analysing machine, they would appear as 'outsiders'. Pedagogical goals, however, differ from objective measuring and parsing, used in mechanical analysis.

3. The test is also reliable as to the constant degree of difficulty for the same subjects. About 80 subjects who took part in the test at the beginning of October 1979 were given the test again quite unexpectedly (with a different randomized choice of items) at the beginning of November 1979. To the results we applied Pearson's correlation test (Eerdeken 1981). The outcome is: if we know the results of the first test, we can predict 85.71% of the second test as well.

$$R = \frac{\Sigma(x_1 - \bar{x})(y_2 - \bar{y})}{[\Sigma(x_1 - \bar{x})^2][\Sigma(y_2 - \bar{y})^2]^{\frac{1}{2}}}$$

R=0.8571 (strong relationship)  
 P=0.0000 (less than 1/10<sup>th</sup> per cent)

4. For most students the degree of active mastery of the words decreases gradually from block 501—1000 to 1001—1500 and further to 1501—2000; from the test of October 1980 we took 75 randomized subjects. The comparison of the results obtained on the block 501—1000 was a 40/50-mean: whereas the rest of the two other blocks together showed a mean of 33/50.

The standard deviation was 3.455 and the correlation in this subtest was 0.661.

The difference between block 501—1000 and 1001—1500 showed also a decrease in mastery:

part I m=40  
 part II m=35  
 part III m=32

We induce from these examinations that:

a) The blocks of 500 in descending rank order of objective frequency form a constant *entity* with a *discriminating power* for measuring the VOC-control (active mastery) of the students at the end of their secondary studies, or at the start of the university.

A more fundamental conclusion might be that, accordingly, we should never leave out *objective frequency* from discrete-point testing strategies in general, and we should encourage manual writers to take it into account again in the future.

b) As good students get about 45/50 on the block 501—1000 and the bad students about 38/50; and as the general mean obtained by all the students reaches 40/50, we might induce that this block has sufficiently been mastered by the students starting with *specialized philological studies* (this is a sensible restriction because it would not apply to students who do not choose English for a future profession).

We could leave this block out of the test in future tests and only check the descending rankorder 1001—2000 of the objective ranklist. This would allow us to include a test on *subjective selections* in so far as they *do not* overlap already with our objective ranklist of 2000 headwords.

We have used two kinds of subjective selections; **COVERAGE** or defining vocabulary used in dictionaries like M. West's (1493 words) and Longman's *Dictionary of Contemporary English* (2134 words).

**FAMILIARITY**, which was measured for 4500 concrete words selected from 2 dictionaries (chosen by Richards 1970) by 1000 students of Laval University.

We calculated the overlapping with the L.E.T.-list for *coverage*:

1— 500=overlapping with TL=82%  
 501—1000=    "    "    " =64%  
 1001—2000=    "    "    " =48%

The overlapping of Richard's *familiarity*-words with the L.E.T.-list is not so frequent, because Richard's choice was restricted to *concrete nouns* only and they do not occur frequently in the written texts on which Brown and LOB are based.

<sup>4</sup> 1978 and 1979.

In the future tests the non-overlapping rest of 2100 COV-Lo and of 1500 FAM-Ri will become part of our PRE-TEST.

## 2. *Mastery-Learning Programs E. VOC.*

Students of English as a foreign language should increase their active control of English vocabulary. What are the assumptions to be taken into account if we want our students to KNOW more words? (Richards 1976)

1. A native adult speaker continues to expand his vocabulary (but with little development of syntax).
- \*2. Knowing a word = knowing degree of PROBABILITY of encountering that word in (speech or) print.
3. Knowing LIMITS of use of a word (according to variations of function and situation).
- \*4. SYNTACTIC behaviour of a word  
(Verb + Prep.; Adj. + Prep.; tr./intr. verbs; etc.)
- \*5. Underlying form of word + its DERIVATIONS
6. NETWORK of ASSOCIATIONS
7. SEMANTIC value
- \*8. POLYSEMY (diff. meanings)<sup>5</sup>

Mastery-learning by means of computer-terminals implies substantial preliminary computer programming on vocabulary.

### 1. Automated analysis of TEXTS

#### a) *Construction of a VOC — Analyser*

Components (Headwords, Morphology + Derivatives):

#### *Objective Frequency*

Merger of Written Corpora

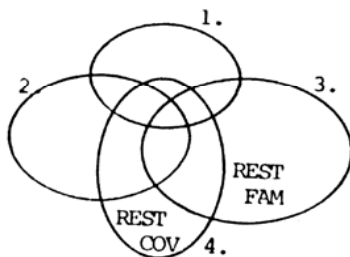
(Brown, Univ., American English,

LOB-corpus, British English)

Written to be spoken Corpus

(Theatre, K. U. Leuven, British English)

2000 L.E.T.-ranklist



1. 2000 L.E.T.-list
2. 1493 WEST
3. 1500 RICHARDS
4. 2100 LONGMAN

<sup>5</sup> The (\*) means that the computer can detect these conditions automatically and that it can be programmed to compose exercises on these assumptions.

### *Subjective Selection*

COVERAGE Lists — WEST 1493

LONGMAN Dict. of Cont. E. 2134

FAMILIARITY List: 1500 from descending FAM-ranklist

(overlapping with Obj. Freq. in L.E.T.-list)

+ REST : COV & FAM

#### b) *Automated ordering of degree of difficulty of texts*

Criteria applied to each text followed by an automated ordering:

1. n words of text
2. n different words
3. mean length of sentences
4. n words between 1—2000 of L.E.T.-list
5. n words FAM ( 1—1500)
6. „ „ „ (1501—4500)
7. „ „ within REST (Lo + We + Ri; to 1500)
8. Outsiders beyond: L.E.T. 2000+REST (Lo + We + Ri; to 1500)

Post-editing is applied to this automated ordering and to the shifting of the outsiders.

Afterwards there is a second ordering, according to the number of words from 501 down to the rank 2000, present in every text. When we started ordering our 100 texts in this way we found out that when we reached the 50th text we had encountered more than 90% of these words. In this way we can assure our students that the 1500 words of the pre-test will get at least one brushing-up during the mastery-learning sessions of the texts. In this way we do not only focus on *new* words (outsiders), but we also try to fix the meaning of the words they are supposed to have studied earlier.

### 2. Automated EXERCISES

- a) The students get a number of texts in book-form. The computer looks up the *outsiders* (for the first 25 texts till 1500 down the objective frequency rank + 750 FAM words; for the next 25 texts till 2000 down the objective frequency rank + 1500 FAM words). All the *outsiders* are defined according to the dictionary. The students have to read the texts, look up and mark the definitions that suit the context. This preparatory work can be done in their own study.
- b) When the texts have been prepared, the students can go to the computer terminals and do the exercises which the computer has automatically composed for them. The texts must not be consulted any more.
  - Vocabulary-exercises (filling-in) on the *outsiders* (new words)
  - Exercise on the *prepositions* and *particles* (phrasal verbs) used in the text (filling-in exercise)

- Exercises on the *deictic* elements of the texts (filling-in exercise)
- Exercises on the *conjunctions* used in the text (filling-in exercise)
- CLOZE-procedure: every seventh word has been left out. The students have to fill in the blanks.
- Exercises on words in the text which occur *between 501—2000* in the descending ranklist.

1. *New context-exercises* for these words
2. *Guessing-exercise*: three or four sentences with a *blank*. The same word has to be filled in, although its meaning may differ (polysemy)
3. *Semantic exercise on polysemy*: Definitions and sample sentences have been separated from each other, and both have been given a random order. The student has to bring the matching definitions and examples together again.

- c) After this mastery-learning of every set of five texts there is a *test* on the learned vocabulary.
- d) A computer programme will read the solutions of the exercises typed by the students into the terminal and compares them with the solutions already stored in the computer memory at the moment of the input of the exercises. The students get immediate feedback when wanted, but the results are also recorded in several *result-memories*. When the exercises have been completed, the results are gathered in one result-memory bank, then sorted under the names of the students; the mean obtained in each exercise, and in the several sets of exercises is calculated and afterwards compared with the results of the *test* after the exercises on a series of five texts. The students are also exposed to the words which caused difficulties during the whole mastery-learning process.
- e) The *PRE-TEST* is taken again in order to calculate the improvement of the mastery of the F words between position 501 down to 2000 in the ranklist.

## REFERENCES

- Block, J. H. (ed.). 1971. *Mastery Learning, Theory and Practice*. New York: Holt.
- Carroll, J. B. 1971. "Problems of Measurement related to the concept of learning for mastery." In J. H. Block (ed.), 1971.
- Carryn, F. 1979. *Critical Investigation of Subjective and Objective Approaches to Vocabulary Selection*. M. A. Diss. K. U. Leuven, Section of Applied Linguistics. Mimeographed.
- De Groot, A. D. 1973. "Ieder kan op de duur alles leren; gaat er ook nooit iets boven uw pet?." *RITP — memorandum 027*. Amsterdam.
- Deterin, K. and R. Högel (eds). 1979. *English auf der Secundärstufe*, Hannover: Wolters-Noordhoff.

- Eerdeken, H. 1981. *Prevoco; Productive English Vocabulary Command. Social Diversities between first year university students in the Flemish provinces of Belgium*. M. A. Diss. K. U. Leuven, Section of Applied Linguistics. Mimeographed.
- Engels, L. K. 1979. "Wortschatzunterricht für Fortgeschrittene". In W. Kühlwein and A. Raasch, (eds). 1979. Band II: 218—19.
- Engels, L. K. e.a. 1981. *A basic Vocabulary for Learners of English, combining objective frequency and subjective word-selection*. Leuven. Section of Applied Linguistics.
- Engels, L. K. e.a. 1981. *Beheersingleren van Engelse woordenschat (met de hulp van computerterminalen)*. Teksten en oefeningen. Reeksen I en II, Leuven, ACCO (voorlopige uitgave).
- Everts, R. 1980. *Towards an Amalgamation of Objective Frequency and Subjective Word-Selection in the Construction of a Basic English Vocabulary for Learning and Teaching Purposes*. M. A. Diss. K. U. Leuven; Section of Applied Linguistics. Mimeographed.
- Gallison, R. 1979. *Lexicologie et enseignement des langues*. Paris: Hachette.
- Geens, D., L. K. Engels and Martin, W. 1975. *Leuven Theatre Corpus and Frequency List*. PAL. Institute of Applied Linguistics. K. U. Leuven.
- Johansson, St. 1980. "The LOB Corpus of British English Texts: Presentation and Comments." *ALLC Journal*, Vol. 1, n° 1. 25—36.
- Kucera, H. and W. N. Francis, 1967. *Computational Analysis of Present-Day American English*. Providence, Rhode Island: Brown University Press.
- Kühlwein, W. and A. Raasch (eds). 1979. *Sprache: Sprache und Verstehen*. Tübingen: Günter Narr.
- LDOCE, 1978. *Longman Dictionary of Contemporary English* London: Longman.
- Levenston, E. A. 1979. "Second Language Acquisition: Issues and Problems." *Interlanguage Studies Bulletin Vol. 4, n° 2*: 147—160, Utrecht.
- Meara, P. 1980. "Vocabulary Acquisition: a Neglected Aspect of Language Learning: Survey Article." *Language Teaching & Linguistics: Abstracts*. Cambridge. 221—246.
- Pools, P., K. Robberechts, and H. Van de Putte. 1978. *A 2000 Word Frequency List Based on the Brown University Corpus and the Theatre Corpus*. K. U. Leuven. M. A. Diss. K. U. Leuven. Section of Applied Linguistics. Mimeographed.
- Richards, J. C. 1971. *Word Familiarity as an Index of Vocabulary Selection With Indices for 4495 words*. Doct. Diss. Université de Laval, Canada. Mimeographed.
- Richards, J. C. 1974. "Word List; Problems and Prospects." *RELC*, Vol. V, n° 2: 69—84.
- Richards, J. C. 1976. "The Role of Vocabulary Teaching." *TESOL Quarterly* Vol. 10: 77—89.
- Van Beckhoven, and B. L. Morris, 1978. *English Vocabulary Command. Testing the Initial Status of First Year University Students in the Flemish Provinces of Belgium. A Pilot Study*. M. A. Diss. K. U. Leuven; Section of Applied Linguistics; Mimeographed.
- Van Ek, J. A. 1977. *The Threshold Level for modern language learning in schools*. London: Longman.
- Van Lóock, and R. G. Vanderloock. 1981. *A detailed Analysis of the K. U. Leuven English Vocabulary Test*. M. A. Diss. K. U. Leuven; Section of Applied Linguistics, Mimeographed.
- Van Parreren, C. 1979. "Erwerb eines fremdsprachlichen Wortschatzes." In K. Deterin and R. Högel (eds). 1979: 22—36.
- Warries, E. (ed.) 1979. *Beheersingleren, een strategie*. Groningen: Wolters-Noordhoff.
- West, M. 1965—1977. *An International Reader's Dictionary*. London: Longman.
- Yule, G. U. 1944. *The statistical Study of Literary Vocabulary*. Cambridge: Cambridge University Press.

*DATA-banks on magnetic tape*

- *LODCE* - Longman
- *Brown-corpus* & word-frequency
- *Lewen Drama Corpus* & word-frequency
- *LOB-corpus* & word-frequency
- Databank of the *Pre-Test*
- *Textanalyser* (2000 L.E.T.-list; 2000 LDOCE, defining vocabulary; 1493 West, defining vocabulary; 1500 Richards; Familiarity)

*Computer*

IBM 3033

*Computer-language*

PLI