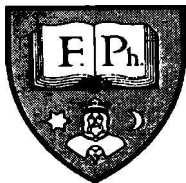


# PUBLICATIONES UNIVERSITATIS MISKOLCINENSIS

SECTIO PHILOSOPHICA  
TOMUS X. – FASCICULUS 2.



E TYPOGRAPHEO UNIVERSITATIS  
MISKOLC 2005



## TARTALOMJEGYZÉK

Mária Csernoch – Ildikó Korponayné-Nagy: A Comparative Analysis of the Word Types of the New Headway Coursebooks	3
Sándor Czeglédy: The Remnants of Bilingual Education in the „Golden State”, 6 Years after Proposition	11
Tamás Demeter: The Empiricists’ Trouble with Private Language	31
Mihály Dobrovits: Who were the Sultans of the Anatolian Seljukids?	43
Éva Forintos: Hungarian-English Language Contact in Australia – a Corpus-Driven Study from a Phonological Aspect	51
Rafał Majewski: Environment of Polish Democratisation Changes	69
Otto Gerhard Oexle: „Facts and Fiction”. On the Current Fundamental Crisis in History	77

# A COMPARATIVE ANALYSIS OF THE WORD TYPES OF THE NEW HEADWAY COURSEBOOKS

MÁRIA CSERNOCH – ILDIKÓ KORPONAYNÉ-NAGY

## ***Introduction***

Surprisingly, until recently the selection of the vocabulary of a given coursebook received very little attention in teaching foreign languages. Earlier, neither the writers of these coursebooks nor the teachers using them attributed much importance to the vocabulary used. As a consequence, when selecting a coursebook testing of the vocabulary was either completely dismissed or, if looked at, it was a minor point as compared to the grammar, to the exercises and sometimes even to the pictures and colors used (Carter and McCarthy 1991; Grant 1994; Oxford and Scarcella 1994; Scurfield 2003).

Recently a number of criteria have emerged that could be used when planning the vocabulary of a coursebook. Here we would like to call the attention to two of these criteria, namely, the facts that a coursebook should include the most frequent and the most useful words (Cunningsworth 1995; Scurfield 2003). The list of the most frequent words can be accessed from several different sources, therefore the writers of nowadays coursebooks have the opportunity to use these frequency lists when planning their coursebook. The second criterion given above for the selection of the vocabulary, the most useful words, is not so easily comprehensible and not so easily tested as the frequency. The problem arises from the simple fact that usefulness has more of a subjective component than frequency. An additional criterion mentioned by Cunningsworth (1995) and by Yongqi (2003), too, that a coursebook should teach at least one thousand words, or as many as possible, during the course. While the given number seems to give a rather objective criterion, it can be very misleading since the „magical” one thousand words can not be completed at all levels and with all age groups. Further uncertainty comes from the term „teach as many words as possible”. It is hardly justifiable to expect from students of a second language to know more words than that used by a native speaker.

## ***Methods***

In these studies the volumes of the *New Headway* series were analyzed. The reason for selecting this particular series was the comment found on the cover of these books stating that the *New Headway* series is the „most trusted English course” (Soars and Soars 2003), which gives a coursebook for every level of studies.

For the analyses the full material of each and every volume of the series – *beginner, elementary, pre-intermediate, intermediate, upper-intermediate,*

*advanced* – was digitized. The volumes were then analyzed individually and, after concatenating the volumes as they follow one another in the series, the complete long texts were also examined.

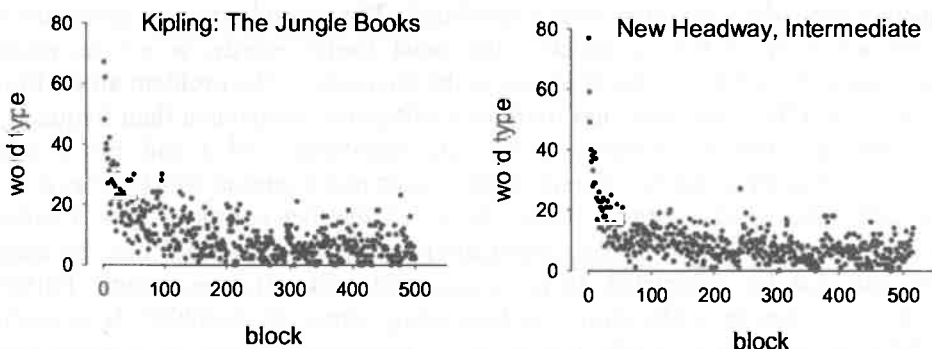
For the analyses the full digitized material was split into intervals containing one hundred words (blocks) and the texts in these small blocks were examined further in detail. The number of word types was counted in each and every block and then stored which, consequently, allowed us to explore the strategy used by the authors for the introduction of new word types. These results were then compared to literary works of approximately equal lengths to establish to what extent does the vocabulary and the tendency of the introduction of word types in a carefully planned (or should have planned) coursebook differ from the same parameters of a novel or a collection of short stories.

### Results

Our results clearly demonstrated that the vocabulary of a coursebook as well as the introduction of the word types in these books greatly resemble the vocabulary of a collection of short stories (*Fig. 1*).

**Figure 1.**

The introduction of word types  
in Kipling, *The Jungle Books* and in the *New Headway Intermediate* coursebook



We compared a collection of short stories, Kipling's *The Jungle Books*, with a coursebook of approximately equal length from the *New Headway* series. The data show that both the number of word types (Table 1) and their distribution (*Fig. 1*) are essentially identical in the two texts. In other words, there is no significant difference between the vocabulary in the two texts. This, furthermore, means that a literary work can just as efficiently be used for the developing of the vocabulary, learning new words and repeating them, of a second language learner as a coursebook written for this particular purpose. Earlier studies (Yongqi 2003) have

also concluded that the reading of literary works can be a more efficient way of developing one's vocabulary than the studying of coursebooks which have, in most cases, independent and randomly selected themes (Oxford and Scarcella 1994).

**Table 1.**

The comparison of the vocabulary of Kipling's *The Jungle Books* and the *New Headway Intermediate* (lemma = block\*100)

Title	Block	Type	Hapax legomena
The Jungle Books	516	4688	2064
Headway Intermediate	500	4803	2072

There is another clear result from counting the word types in the *New Headway* series. The number of word types by far exceeds the expected *one thousand types/volume* in each and every volume of the series (Tables 2 and 3). This observation is noteworthy even though we have not lemmatized our texts.

When a computer aided analysis of texts is carried out the question of lemmatization, whether the lemmatization is needed for the accurate evaluation, always reappears. The same is true for coursebooks, one should consider if lemmatization should precede the actual analysis. However, if we take into account that during the teaching/learning of a second language the different forms of the given word usually all appear in the students' vocabulary books, we can dismiss the lemmatization of the text and, probably, the result will be closer to real life and thus be more reliable. This is especially true for beginners who not only learn the lexical entry (lexeme) of a word but, due to the lack of knowledge of appropriate grammatical rules, the inflected forms as well. There are a number of examples when the student does not yet know the lexical term, but knows, from films, from the lyrics of a song, from readings, or uses in conversations other forms of the given word, e.g. *born, satisfied, data*. Taken all these into account lemmatization did not precede the actual analysis words were counted in the form they appeared in the text.

**Table 2.**

The number of words (lemma), word types and hapax legomena in the volumes of the *New Headway* series

Title	Block	Type	Hapax legomena
Beginner	163	1539	501
Elementary	239	2452	864
Pre-Intermediate	317	3309	1373
Intermediate	500	4803	2072
Upper-Intermediate	511	5646	2430
Advanced	513	6724	3274

The number of words in the given volume increased steadily until the volume *intermediate* while the last three volumes had approximately equal number of words. There is no clear reason for setting an upper limit for the number of words. A more reasonable approach would have been to increase the number of words in the oncoming volumes together with setting an upper limit on the number of word types. This would have increased the number of repetition of a given word and, consequently decreased the number of hapax legomena.

On the other hand, every volume considerably exceeds the one thousand words that are expected to be taught on a course, and there are cases when several thousand word types are found in a volume. It should also be stressed that the number of hapax legomena (Tables 2 and 3) and the number of words occurring less than ten times in the whole series is extremely high. This should be viewed in the light of a number of studies pointing out students fail to remember half of the words that they encounter only once or twice. A word should appear at least 7-8 times (extreme opinions require 5-16 occurrences) for students to remember them (Yongqi 2003; Scurfield 2003; Oxford and Scarcella 1994). By the end of the series the students encounter 224 500 words, 11 648 word types from which 4 636 occurs only once in the series. Both the number of word types and the number of hapax legomena are, in our opinion, too high. As a result, teachers using this series should be prepared that the series can only be used successfully from the point of vocabulary teaching if they use substantial amount of additional material which is systematically designed from the point of lexical content. This is in accordance with the opinion of Scurfield (2003) who, based on earlier studies (Oxford and Scarcella 1994; Schmitt 2000), suggested that most second language teachers, logically, assume that writers of these coursebooks have a general idea based on which they build up their books. Therefore, the vocabulary of the coursebooks should make a whole. Unfortunately, and unforgivably sometimes this is not the case.

**Table 3.**

The number of words in the concatenated volumes of the <i>New Headway</i> series			
Title	Block	Type	Hapax legomena
Beginner	163	1539	501
Beginner→Elementary	402	2943	962
Beginner→Pre-Int.	719	4550	1628
Beginner→Intermediate	1220	6760	2607
Beginner→Upper-Int.	1731	8989	3456
Beginner→Advanced	2245	11648	4636

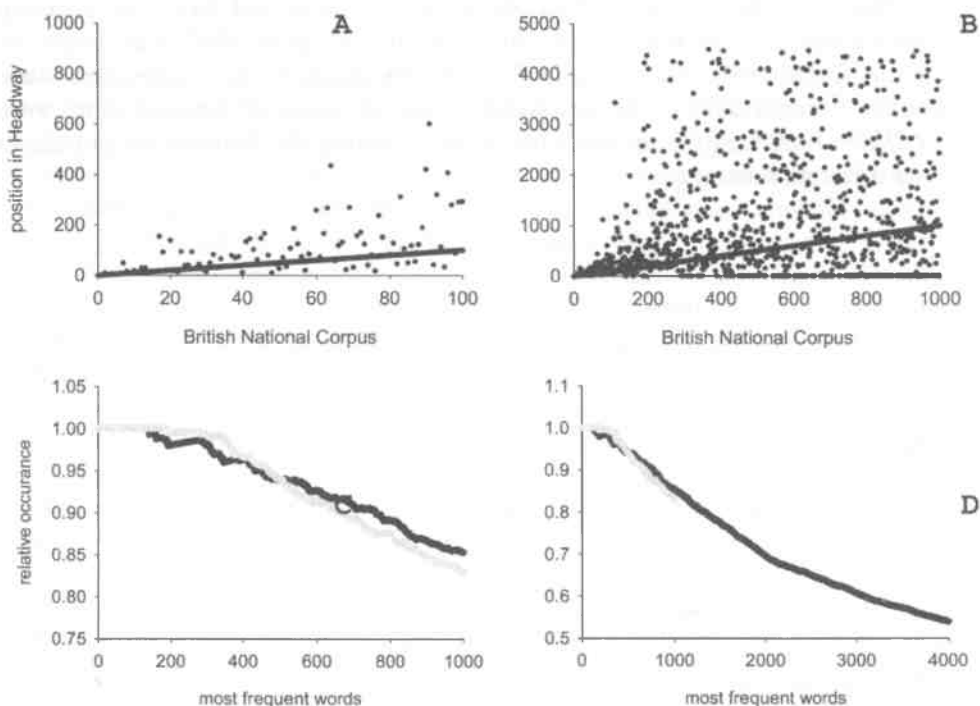
When writing a coursebook it is of vital importance to decide which words should, or must be included in the vocabulary. Although there are several aspects to the

question (Oxford and Scarcella 1994; Carter and McCarthy 1991; Scurfield 2003; Yongqi 2003), almost every work in the field agrees that the teaching of the most frequently used words should be the primary goal. Therefore, we used the frequency list of the *British National Corpus (BNC)* (Aston and Burnard 1998) (<http://www.itri.bton.ac.uk/~Adam.Kilgarriff/bnc-readme.html>) and the 1000 most frequently used words from the *Reading Teacher's Book of Lists* (Fry et al. 2000) (<http://www.duboislc.org/EducationWatch/First100Words.html>) to examine to what extent do the volumes of the *New Headway* series fulfill the requirement of teaching the most frequently used words.

In order to enable the comparison of the word list from *Headway* with the other frequency lists they first had to be formed to have the same structure. Since the structures of the frequency lists from the *Headway* and from the *Reading Teacher's Book of Lists* were almost identical, the list from *BNC* was altered to match the former two. To this end first the compounds and expressions were separated into individual words then the different meanings of the same word were counted. The final frequencies were obtained by adding the frequencies associated with the different meanings.

**Figure 2.**

Comparison of the frequency list from the *BNC* with the first three concatenated volumes of the *New Headway* series. The continuous lines in **A** and **B** show the ideal case, in which the two lists would agree. In contrast, the number of words in the series which have different frequencies than in the *BNC*, is rather high. The graphs in **C** and **D** show how many of the first  $n$  words in the frequency list of the *BNC* (black trace) and of *the Reading Teacher's Book of Lists* (gray trace) are found in the *New Headway* series. As long as the ratio is 1, the two values are identical, if it is less than 1, then there are words which, albeit present in the lists of large corpora are not in the *New Headway* series.



Comparing the frequency list of the *New Headway* series to those of the two large corpora we found that there is no large difference between the frequencies of words from closed classes. In contrast, the coursebooks contain a number of words the knowledge of which is essential for solving the exercises in these books. These words have a much lower frequency in the *BNC* and are not in the first 1000 words of the list from the *Reading Teacher's Book of Lists*. This means that, even at the beginning of second language teaching words that students are taught words that



they use exclusively on language classes and, therefore, their practical value is negligible. Apart from technical terms associated with the exercises and the grammar, coursebooks tend to have larger frequency for words associated with eating, with working or with family relations. On the other hand, there is a large number of words that appear in the front of the frequency lists from large corpora but have a very low frequency in the *New Headway* series.

### **Summary**

In our studies we found that the vocabulary and the way how word types are introduced in the analyzed coursebook series does not differ significantly from the same parameters of a collection of short stories. Furthermore, we found that the number of word types and the number of words that are mentioned only once in the whole series are too large. This observation, together with earlier studies suggests that the students will, most likely, not remember the words that have such low frequencies. The question thus arises how useful for students is and how effectively can a teacher use a coursebook with such parameters. The numbers show that the books provide an extensive, albeit not necessarily the most frequent and useful vocabulary, therefore, their efficient use in the classroom requires that the teachers prepare large additional material for the students.

### **References**

- Aston, G. – Burnard, L. 1998: *The BNC Handbook: Exploring the British National Corpus with SARA*, Edinburgh University Press
- Carter, R. – McCarthy, M. 1991: *Vocabulary and Language Teaching*, London: Longman Group
- Cunningsworth, A. 1995: *Choosing your Coursebook*, Oxford: Heinemann
- Fry, E. B. – Kress, J. E. – Fountoukidis, D. L. 2000: *The Reading Teacher's Book of Lists*, New Jersey: Jossey-Bass Hoboken
- Grant, N. J. H. 1994: *Making the most of your Textbook*, London: Longman
- Oxford, R. L. – Scarcella, R. C. 1994: „Second language vocabulary learning among adults: State of the art in vocabulary instruction”, *System* 22, No. 2.
- Schmitt, N. 2000: *Vocabulary in Language Teaching*, Cambridge University Press
- Scurfield, S. 2003: „Give vocabulary learning a chance”, *TESOL-Spain Newsletter* (<http://www.tesol-spain.org/newsletter/scurfield.html>)
- Soars, J. – Soars, L. 2003: *New Headway Advanced*, Oxford University Press
- Yongqi, P. G. 2003: „Vocabulary learning in a second language: Person, task, context and strategies”. *Teaching English as a Second or Foreign Language*, No.2