

WORD COLLOCATIONS IN FL VOCABULARY INSTRUCTION*

MARIA KOROSADOWICZ-STRUŻYŃSKA

The Jagellonian University of Cracow

1. *Vocabulary in context*

There are very few real "knows" about effective vocabulary instruction. There is, however, at least one universally acknowledged sure-fire procedure, which has been utilized during the last 30 years or so, viz. that of teaching vocabulary in context. To do justice to the audio-lingual movement, it was just this approach to which we owe the popularity of this idea: according to its proponents, words ought to be presented to the learner in self-defining contexts, preferably linked with some extralinguistic situation, the latter being usually hinted at in a dialogue or text.

In fact, a context may mean a variety of things. From the point of view of lexicon, there can be found several contexts with no clear boundaries between any two of them but rather in a continuum from the broadest to the narrower. One could picture a double scale like the following:

EXTRALINGUISTIC CONTEXT

Culture	Referential clusters (human activities, situations)			REFERENT
A language	Discourse (discourse clusters)	Utterance	Phrase	WORD

LINGUISTIC CONTEXT

* I am indebted to Prof. Edward M. Anthony and Prof. Ruta Nagucka for their helpful comments and constructive criticism of this paper.

Thus, starting from the broadest context, we have the socio-cultural one within which any study of vocabulary should be conducted (cf. Anthony 1975a: 49 ff.). Inseparably linked with it is the situational context characterized by activities typical of or valued by a given culture (i.e. referential clusters) and means of speaking about them (i.e. discourse clusters) (cf. Anthony 1954; 1975 a, b). Awareness of the fact that such lexical clusters exist can be extremely helpful in teaching recognition.¹ Since a great many referential clusters are common to all cultures, transfer from the NL of the learner may often bring about a certain facilitation in the form of intuitive guessing at the meaning of a word placed in such a context (i.e. in a corresponding discourse cluster). Certain activities, however, are culture specific and unless the student is made familiar with the cultural background in which they appear, the probability of occurrence of certain lexical items will not be so very obvious to him.

As for the other side of the above scale — the linguistic context, there are the following (in descending order of inclusiveness): discourse — text, utterance — sentence, and phrase. And, again, if a sentence or utterance is placed in a clearly defined, meaningful situation, it is fairly easy to find out the meaning of a new word or phrase which occurs in it. The more so if the word or phrase is presented in several contrasting sentential contexts.

It is worth mentioning that the context of phrase is often ignored in the literature on language didactics and, consequently, in vocabulary instruction. This is due to the fact that most writers in the field consider a student's knowledge of vocabulary to be tantamount to "knowledge of the words carrying lexical meaning" (Politzer 1970:135). This would mean that his knowledge is equivalent to the number of isolated words the student has in his permanent vocabulary. As a reflexion of such a standpoint, the prevailing practice in teaching vocabulary consists in imparting the knowledge of isolated words (although presented in context) to the students. At the most, in advanced language courses, within the goal of teaching word formation, compound words are discussed and practised (cf. Norris 1970).

A similar view pervades the related field of psycholinguistics, and in studies on verbal learning and verbal behaviour investigation of how isolated words (if not nonsense syllables) are learned and stored in memory predominates.²

It should be noted, however, that although single words themselves cause trouble in the foreign-language classroom (as the learner has to know not only

¹ For further arguments along these lines and a discussion of some other possible applications of lexical clusters in language pedagogy, see Anthony (1975b:25).

² For the most part they are studies on acquiring meaning and on mnemonic devices, such as interactive-image elaboration (cf. e.g. Ott et al. 1973), used to learn words in isolation (cf. Henning 1973) or word-paired associates (cf. e.g. Bugelski et al. 1968).

their form, grammatical function and denotation(s) but also their connotations, and at all these points languages differ widely making this task even more difficult), knowing even a great number of them is not enough for the student to function in the language. And even these teachers who have jumped on the "creativity" bandwagon would admit that in any linguistic community there are certain rules for putting words together and therefore, as Bastian puts it, "the lexical items are never concatenated with complete freedom" (Bastian 1968:19). It is concatenations of lexical items that cause the greatest trouble in vocabulary acquisition and the learner's mastery of such rules, not his knowledge of single words, should be an indication of his progress.

Phrases, which are nothing but instances of the above mentioned concatenations, are made out of lexical material by means of both formal (i.e. grammatical) and semantico-conceptual syntagmatic relations. Let us leave aside the former ones as irrelevant to the point at hand, since on the production level, which will now be our primary concern, they can be assumed to have been already mastered by the learner; instead I shall concentrate on the second type of syntagmatic relations called collocations.

2. Collocations

2.1. Theoretical considerations³

The notion of collocation is connected with word distribution and its probability of occurrence in certain contexts. Generally speaking, collocations (sometimes called "conventional syntagmes")⁴ (cf. Szule 1971:67) might be regarded as situationally appropriate forms of language which are to a great extent institutionalized.

Each lexical system, though being open, has at its disposal certain closed, determined or, in other words, conventional structures. Needless to say, they vary from one language to another. These structures can be roughly divided into

- A. *idioms*, i.e. phrases the meaning of which cannot be inferred from the meaning of their components (e.g. make (a) book=arrange a series of bets on an event) or phrases used as words (e.g. son of a gun) (Anthony; personal communication);
- B. *set expressions* whose meaning can be inferred providing the learner knows the meaning of their component lexical items and, in certain specific

³ The discussion here is mainly based on the probabilistic theory of meaning by collocation developed by J. R. Firth (1957, 1968).

⁴ My translation.

cases, has some background knowledge (e.g. make place (room) for somebody, Independence Day, respectively)⁵ (CRITERION I).

What obscures the issue is the fact that sometimes both types of determined structures are classed as idioms since it is often believed that no hard and fast distinction can be made between them. For example, Firth's definition of collocation does not help to rule out idioms:

Collocation states the habitual company a key-word keeps. Words must at some level or other be taken at their face value in their common and usual verbal environment (Firth 1968:113).

It may be assumed only that from the formal point of view both phrases and compounds will fall under the category of collocations. And surely enough, in another place Firth (1968: 107—108) illustrates the concept of collocation by means of "collocational compounds" and "collocational phrases" in which "common words" appear, e.g.

safety — safety match, safety first, etc.;

words for the *months* — March hare, August Bank Holiday, May week, May day, April shower, April fool, etc.;

English — the English people, Public Schools; English literature, reserve, manner, countryside, universities, versus

British — Army, Commonwealth, Empire, way of life, Isles, Foreign Policy, etc.

He also emphasizes the fact that certain words are compatible with lots of others, certain with only a few. For instance the word *ass* in its figurative use most frequently collocates with *silly* as in *You silly ass*, or with *be*, as in *Don't be an ass* (Firth 1957:195). The word *time*, on the other hand, enters into numerous conventional phrases, as it

...can be used in collocations with or without articles, determinatives, or pronouns. And it can be collocated with *saved*, *spent*, *wasted*, *frittered away*, with *presses*, *flies*, and with a variety of particles, even with *no* (Firth:1957:195).

On the basis of these examples alone one more important difference between idioms and collocations suggests itself, viz. the former are closed while the latter more or less closed, that is, determined to a certain degree only (CRITERION II). What follows, an idiom cannot be altered, no other synonymous word can be substituted for any word in an idiomatic phrase and the arrangement of the words can rarely be modified (e.g. *he kicked the bucket* is to be treated as a whole and as such cannot be conceivably passivized).

⁵ Palmer (1932) refers to the former as "non-normal formulas" or "non-normal phrases", while some of the latter ones are called by him "normal phrases", others "pologs" (pp. 26—40).

With collocations, however, the situation is different: most of them do not lack syntactic flexibility, and, moreover, although representing syntagmatic relations they are highly predictable,⁶ some paradigmatic (i.e. individual) choice is often given to the speaker, e.g.

a speech	can be	given
		made
		delivered
		read, etc.
but it cannot be		*enunciated
		*expressed
	or	*pronounced

to say having looked up the word *wygłosić* (as a Pole is likely in a bilingual pocket dictionary). The examples with an asterisk cannot be said to be ungrammatical, yet they certainly are inappropriate in English — they are incompatible with speech (cf. McIntosh 1966). In other words, word collocation implies that having thought or said A to express a certain meaning, you also have to say B or C — you have committed yourself to say either one (or probably one out of several possibilities available, say B, C, D, or E); you must not, however, say K, L, or M because, according to the linguistic practices of the community whose language you are learning, they are not compatible with A.

And yet one cannot be completely satisfied with what has been stated so far; we are still left with a residue of phrases (more specifically, with some figurative expressions) which in the light of the above definitions and criteria lie on the borderline between idiom and collocation. Here an additional criterion (CRITERION III) is provided by Lado (1955), viz. identifying idioms by comparing two languages. The following extended citation will give an accurate picture of what he has in mind:

'Idioms' — expressions peculiar to a language — are identifiable as we compare two languages rather than within the language itself. An expression which may seem peculiar to native speakers may be quite natural to speakers of another language and would therefore not be an idiom to them. On the other hand, an expression which seems quite natural to native speakers may be strange to foreign speakers of a particular language background. If we should find on comparing the expression with a variety of languages that it is strange to all or nearly all of them, we would be justified in calling it an idiom in general, but even then the statement would be meaningless in those cases in which the other language had a parallel expression (1955:287).

According to Lado's suggestion, expressions such as *kill time* which, in terms of criterion II (impossibility of passivization), should be classed as

⁶ In this sense collocations are akin to what Lyons (1969: 422, 440) refers to as syntagmatic presuppositions, although what differs them from this class is, as stated above, that they are never completely determined.

idioms, will not be idioms for a Polish learner because his native language uses a similar turn of phrase (in this instance *zabijać czas*). Since their meaning can easily be inferred by this particular learner, they do not fulfil the condition set by criterion I.

Although none of the criteria discussed above is ironclad, all three together, I believe, provide a principled basis for determining which from among a vast repertory of set expressions are collocations proper.

Now let us consider what bearing the above discussion has on teaching vocabulary and what difference it makes, if any, whether one can or cannot distinguish idioms from collocations.

2.2. *Pedagogical implications*

Idioms are difficult on the perception level, but since they do not appear too frequently in everyday situations the student can quite effectively do without them in his performance. That is why it is often enough if they become part of his passive vocabulary only. On the other hand, if we find it necessary to activate in the student's production the most frequent ones, because of their very nature, i.e. the fact that they are completely closed and therefore belong to low-level linguistic decisions,⁷ they simply have to be learned by heart and a habit formation in the drill form should be very effective.

Collocations, however, occur practically in every utterance or sentence. Besides, as has already been hinted at, they lie on the borderline between high- and low-level decisions: their choice is to a certain degree conscious on the part of the speaker, and as such their acquisition involves some more complicated procedures in teaching terms.

Although they are relatively easy on the level of perception (and that is why they are usually lost sight of in teaching procedures), on the production level they present a hundred pitfalls to the foreign student because of a certain choice given to the speaker, and the possibility of false analogy, i.e. intra-language interference. Moreover, they provide a great opportunity for mother-tongue interference and are therefore the most difficult part of the language for a foreigner to master. This becomes evident when we listen to an advanced student's speech, even after years of exposure to the target language, or when we have to correct his written compositions. With a learner who already possesses considerable fluency of expression, errors due to his inadequate mastery of grammar are rather rare; yet those which consist in a faulty use of lexical items he has at his disposal⁸ and, more specifically, in

⁷ For a detailed discussion of high- and low-level decisions in language use, see Rivers (1969:50ff., 1976: 81) and Levelt (1977).

⁸ Cf. e.g. Grauberg (1971); Spalatin (1971); both in James (1977).

his departure from some tolerated ranges of collocability — simply crop up.⁹ Some examples of both intra- and interlanguage interference operating in this domain of lexicon have long been language teachers' classic.

Errors in the use of word collocations surely add to the foreign flavour in the learner's speech and writing and along with his faulty pronunciation they are the strongest markers of "an accent"; it is difficult to say, though, whether they barely impair communication. Generally speaking, the problem of error gravity is not a simple one. Any attempts to establish a scale of errors, be it on purely linguistic criteria or on native speakers' judgements,¹⁰ have so far failed to provide us with valid conclusions. Johansson (1977) suggests that such a scale should be related to the learner's communicative needs as

The more an error interferes with communication, the more serious it seems from the point of view of the learner (1977:41).

According to the system of evaluation presented below (Fig. 1), erroneous sentences which are fully comprehensible and cause no irritation on the part of the interlocutor belong to the lowest grade of errors. Those sentences (utterances) however, which, although comprehensible, "make the receiver tired or irritated or draw away his attention from the contents of the message" (1977 : 43) should be considered serious errors.

Errors in word collocations would belong, as it seems, to the latter group.¹¹

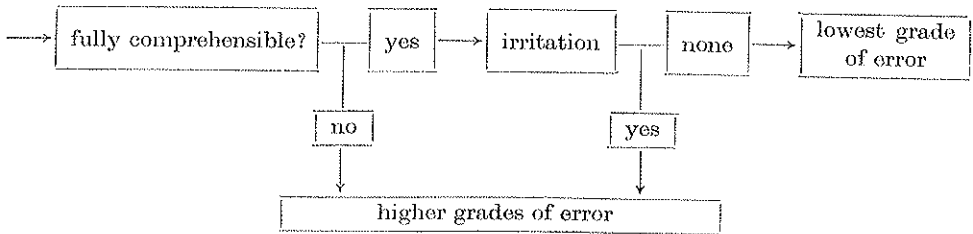


Fig. 1. The evaluation of errors (adopted from Johansson 1977: 43, fig. 3).

Though one cannot but agree with Smith's statement that "mastery of the utterance should be the culmination of learning, not the beginning"

⁹ That is why modern measures of proficiency in a FL, such as "clozentropy" — the scoring procedure based on the principles of probability developed by Darnell (1970), measure first of all student mastery of word collocations.

¹⁰ There is some evidence from studies on judgements of error gravities (as, for instance, the study carried out by James 1977) which indicates that native speakers of English consider lexical errors as the least serious and detrimental to intelligibility. For a discussion of linguistic measures of error gravity, see James (1974).

¹¹ This is of course only a speculation and as such it has to be verified by an exhaustive study of receivers' overt and covert reactions to students' errors.

(1971:42), yet there is little reason to believe that steps to promote such mastery cannot and should not be taken from the first stages of teaching. The latter seems to be a sounder procedure than arranging remedial courses afterwards, when it is usually too late because lexical errors tolerated at the beginning stages have become fossilized or, in James' phrase, "indelibly entrenched" (1977 : 118). We should make it a point at the beginning of a language course to develop in the student some ability to get his message across both in oral and written productions in the FL. It is a long process, but it pays off in the long run. Thus, one might paraphrase the above quotation from Smith by saying that mastery of the commonest collocations should be one of the essential outposts to genuine mastery of English (or any other language).

Having this goal in mind, I would concede that a widespread idea of limiting vocabulary at the beginners' level is sound providing that this limitation affects key-words only and that the words from this limited stock are shown and practised in class in their most typical collocational settings. Hence our decision of taking collocations into account will surely influence the selection of lexical items to be taught. The usual procedure consists in selecting the so-called minimum essential words. This is done on the basis of usefulness and frequency of usage¹² to which Anthony (1954 : 179, 1975b : 26) contributes the situational criterion of referential clusters, i.e. selecting certain culture-bound real-word situations around which lexical meanings of words cluster and disambiguate one another. I would add to this one more criterion, that of word-collocability. Along this line, key-words which meet the above criteria and which, additionally, have a great collocational power — in the sense that they trigger off numerous collocations and, moreover, enter into collocations with one another — should be selected for teaching purposes.

What are some possible classroom applications of these considerations? In order to reduce the problem to teachable dimensions, having selected such key-words, a pretty exhaustive collection of collocations of high frequency should be made. Then these selected key-words should be presented in a variety of most typical contexts (situational, sentential, and collocational) to show the students their meaning, both denotative and connotative.¹³

It is rather obvious that this cannot and should not be done in solid blocks — not all the selected collocations of a key-word can be presented at a time. On

¹² Cf. e.g. Fries and Traver (1960); Widdowson (1978). Here, in keeping with the "individual needs and interests" cliché, the vocabulary content will be tailored to students' demands, and in the case of beginners — the everyday register, in the case of advanced students — some restricted register will probably be chosen.

¹³ It is precisely at this point that the notion of word collocations will be exploited along the lines pointed out by Firth, i.e. in determining meaning.

the contrary, students' knowledge of the meaning of words (and, what follows, their collocations) should be enlarged step by step and therefore collocations will be introduced in sequence at different intervals until we come to the stage of a review — a word study lesson where the hitherto learned collocational contexts will be brought together.

Since the native language of the learner and his thought habits must not be ignored, in order to prevent mislearning through mistaken assumption about analogical relationships in the native and target languages, the most typical collocations of the given word can be contrasted with the collocations in which the equivalent native-language word or words appear. Needless to say, such a technique is advisable only when there are considerable differences between the two languages and only when we anticipate a negative transfer from the NL. In such cases, especially with adult learners, it seems quite effective to make them fully aware of the problem.¹⁴

This should be followed immediately by meaningful practice in using the collocations. Such practice can take the form of what I would call a "quick translation" exercise for which native language equivalent phrases are chosen at random by the teacher and the students are supposed to answer mechanically in the target language. The whole procedure is aimed at both automatization of the phrases at issue and at expressing the right meaning.¹⁵ By way of illustration, the collocational set being that of the noun *walk* (preferably put on the board for the students to see), viz.

1. go for a _____,
2. take a person for a _____,
pet
3. have a _____,
take
4. X is only an "n" minutes _____, w dolnej frakcji from here,
the cues will be respectively:
 1. iść na spacer,
 2. wziąć kogoś na spacer,
 3. iść na spacer, odbyć spacer, spacerować,
 4. "n" minut stąd.

English phrases (1), (2) and (4) will probably require some additional practice because of the danger of mother-tongue interference; for (1) and (2) the Polish student is likely to use the preposition *on* by analogy to Polish *na*, in (4) he will show a tendency to omit the word *walk* because in his NL indicating the means of covering a distance is not obligatory.

The next exercise can involve "limited creativity" — here, along the lines

¹⁴ For a detailed discussion of contrastive analysis as a classroom technique, see Marton (1973). For an opposite view, see Levelt (1977:68), who maintains that "contrastive teaching should only be practised in case the subject has made an error".

¹⁵ For an opposite view regarding the latter aim, see Anthony (1975b:29).

suggested by Anthony (1952/53:83), the cues concerning a particular teaching point are given by the instructor again in the students' NL, and the learners are to answer in the TL, using the suggested phrase in sentences of their own, e.g.

T : iść na spacer

S₁ : I often go for a walk in the evening.

S₂ : I prefer to take a walk early in the morning, etc.

Some question-and-answer practice can follow, where the students' task is to ask questions of one another using or provoking the use of one of the collocations. No cues are given by the teacher unless necessary. In order to avoid monotony several collocational sets can be practised at a time. In the case under discussion, the students might produce something like the following:

- How many times a day do you take your dog for a walk?
- Three times at least.
- Would you like to take a walk in the park?
- I'd love to.
- How far is it to the railway station?
- It's a ten minutes' walk from here.
- Let's go for a walk, shall we?
- O.K. Let's.

As a homework assignment, the students can be asked to write a simple dialogue or a story in which the indicated collocations will have to be used.

Then, the mastery of such a set (or sets) will be put to test in a cloze procedure where the learner is presented with passages from which every *n*-th word has been deleted (in the case at issue it will be a member of a collocation) and is asked to supply the missing items. In scoring procedures I would advocate accepting reasonable synonyms, not only those responses identical to the originally deleted words.

Testing exercises of the latter type should be reintroduced at intervals over a period of time, their lexical content enlarged as the course progresses, with interspersed opportunities for the students to use the already acquired collocations in a communicative, situationally bound interaction.

Much of what has been said here is in no way new. The suggestions presented are tentative and exploratory. Since it is rather obvious that no single pedagogical approach will suffice for all our problems, a caveat should be inserted at this point, namely the idea of teaching word collocations must not be treated as a panacea. Nevertheless, if our goal is to help the student to reach a mastery of a FL effectively and economically, this seems to be a promising field.

REFERENCES

- Anthony, E. M. 1952/53. "The teaching of cognates". *Language learning* 4, 2/3. 79—82.
- Anthony, E. M. 1954. "An exploratory inquiry into lexical clusters". *American speech* 29/3. 175—180.
- Anthony, E. M. 1975a. *Towards a theory of lexical meaning*. Singapore: Singapore University Press.
- Anthony, E. M. 1975b. "Lexicon and vocabulary — some theoretical and pedagogical aspects of word meaning". *RRLC journal* 6/1. 21—30.
- Bastian, J. 1968. "The biological background of man's languages". In O'Brien, R. J. (ed.). 1968. 17—24.
- Bugelski, B. R. Kidd, E. and J. Segmen. 1968. "Image as a mediator in one-trial paired-associate learning". *Journal of experimental psychology* 76. 69—73.
- Croft, K. (ed.). 1972. *Readings on English as a second language: for teachers and teacher-trainees*. Cambridge, Mass.: Wintthrop Publishers.
- Darnell, D. 1970. "Clozentropy: a procedure for testing English language proficiency of foreign students". *Speech monographs* 37/1. 36—46.
- Firth, J. R. 1957. *Papers in linguistics 1934—1954*. London: Oxford University Press (reprinted in 1969).
- Firth, J. R. 1968. *Selected papers of J. R. Firth 1952—1959*. (Ed. by F. R. Palmer) Bloomington: Indiana University Press.
- Fries, C. C. and A. A. Traver. 1960. *English word lists — a study of their adaptability for instruction*. Ann Arbor: The George Wahr Publishing Co.
- Henning, G. H. 1973. "Remembering foreign language vocabulary: acoustic and semantic parameters". *Language learning* 23/2. 185—195.
- James, C. 1974. "Linguistic measures for error gravity". *Audio-visual journal* 12. 3—9.
- James, C. 1977. "Judgements of error gravities". *BLT* 31/2. 116—124.
- Johansson, S. 1977. "Problems in studying the communicative effect of learner's errors". In Valdman, A. (ed.). 1977. 41—52.
- Lado, R. 1955. "Patterns of difficulty in vocabulary". In Croft, K. (ed.). 1972. 277—291.
- Levelt, W. J. M. 1977. "Skill theory and language teaching". In Valdman, A. (ed.). 1977. 53—70.
- Lyons, J. 1969. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Marton, W. 1973. "Contrastive analysis in the classroom". *PSiCL* 1. 15—22.
- McIntosh, A. and M. A. K. Halliday. 1966. *Patterns of language in general, descriptive and applied linguistics*. London: Longmans.
- McIntosh, A. 1966. "Patterns and ranges". In McIntosh, A. and M. A. K. Halliday. 1966. 183—199.
- Norris, W. E. 1970. "Advanced reading — goals, techniques, procedures". In Croft, K. (ed.). 1972. 188—207.
- O'Brien, R. J. (ed.). 1968. *Georgetown University Round Table — Selected papers on linguistics 1961—1965*. Washington, D. C.: Georgetown University Press.
- Ott, E. C., Butler, D. C., Blake, R. S. and J. P. Ball. 1973. "The effect of interactive-image elaboration on the acquisition of foreign language vocabulary". *Language learning* 23/2. 197—206.
- Pabner, H. E. 1932. *The grading and simplifying of literary material*. Tokyo: Institute for Research in English Teaching.
- Politzer, R. L. 1970. "Learning syntactical patterns". In Croft, K. (ed.). 1972. 133—147.

- Rivers, W. M. 1969. *The psychologist and the foreign language teacher*. Chicago: The University of Chicago Press.
- Rivers, W. M. 1976. *Speaking many tongues* (expanded 2nd edition). Rowley, Mass.: Newbury House.
- Smith, P. D., Jr. 1971. *Toward a practical theory of second language instruction*. Language and the Teacher, vol. 16. Philadelphia: The Center for Curriculum Development.
- Szulc, A. 1971. *Lingwistyczne podstawy programowania języka*. Warszawa: PZWS.
- Valdman, A. (ed.). 1977. *Studies in second language acquisition*. Vol. 1. Bloomington: The Indiana University Linguistics Club.
- Widdowson, H. G. 1978. *Teaching language as communication*. Oxford: Oxford University Press.