The Role of Syntactic Class, Frequency, and Word Order in Looking up English Multi-Word Expressions

Robert Lew, *Department of Lexicography and Lexicology, Faculty of English, Adam Mickiewicz University in Poznań, Poland (rlew@amu.edu.pl)*

Abstract: Multi-word lexical units, such as compounds and idioms, are often problematic for lexicographers. Dictionaries are traditionally organized around single orthographic words, and so the question arises of where to place such complex lexical units. The user-friendly answer would be to include them primarily under the word which users are most likely to look up. But how do we know which words are likely to be looked up? The present study addresses this question by examining the roles of part of speech, word frequency, and word position in guiding the decisions of Polish learners of English as to which component word of a multi-word expression to look up in the dictionary. The degree of word frequency is found to be the strongest predictor, with less frequent words having a significantly greater chance of being selected for consultation. Then there is an independent part of speech-related preference for nouns, with adjectives being second, followed by verbs in third place. Words belonging to the remaining syntactic categories (adverbs, prepositions, conjunctions, determiners, and pronouns) are hardly looked up at all. However, word placement within the multi-word expression does not seem to matter much. This study has implications for dictionary makers in considering how to list multi-word-expressions.

Opsomming: Die rol wat sintaktiese kategorie, frekwensie en woordorde speel in die naslaan van Engelse meerwoordige uitdrukkings. Meerwoordige leksikale items, soos samestellings en idiome, is dikwels problematies vir leksikograwe. Woordeboeke word tradisioneel gerangskik om enkele ortografiese woorde, en dus ontstaan die vraag waar sulke komplekse leksikale items geplaas moet word. Die gebruikersvriendelike antwoord sou wees om hulle primêr in te sluit onder die woord wat gebruikers gewoonlik eerste sou naslaan. Maar hoe weet ons watter woorde gewoonlik nageslaan word? Die huidige studie spreek hierdie vraag aan deur die rol te ondersoek wat die woordsoort, woordfrekwensie, en woordposisie speel om Poolse aanleerders van Engels se besluite te rig oor watter woord om binne 'n meerwoordige uitdrukking in die woordeboek na te slaan. Daar is gevind dat die woordfrekwensie die sterkste voorspeller is, met minder gebruiklike woorde wat 'n beduidend groter kans het om vir naslaan gekies te word. Dan is daar 'n onafhanklike voorkeur wat verband hou met die woordsoort vir selfstandige naamwoorde, met byvoeglike naamwoorde in die tweede plek, gevolg deur werkwoorde in die derde

plek. Woorde wat behoort tot die oorblywende sintaktiese kategorieë (bywoorde, voorsetsels, voegwoorde, bepalers en voornaamwoorde) word skaars opgesoek. Woordplasing binne die meerwoordige uitdrukking skyn egter nie baie saak te maak nie. Hierdie studie het implikasies vir woordeboekmakers wanneer hulle oorweeg hoe om meerwoordige uitdrukkings te lys.

Sleutelwoorde: MEERWOORDIGE UITDRUKKING, MWE, MEERWOORDIGE EENHEID, MEERWOORDIGE ITEM, WOORDEBOEK, WOORDEBOEKRAADPLEGING, WOORDEBOEK-TOEGANG, LEKSIKOGRAFIE, ENGELSE TAAL, POOLSE AANLEERDERS

1.   Introduction

At first sight it seems that dictionaries treat headwords as if users' look-up strategy is based only on single words. This implies a simplistic view of lexical items as single words, which agrees with the naive view of language, and also with the Chomskyan 'slot-and-filler' model of language[1], which itself may owe much to the impact of the structure of the (printed) dictionary on the linguist as a naive dictionary user, cf. Nowakowski (1990).

However, for describing lexical phenomena, a Sinclairian view of language may be more fitting. It emphasizes the *idiom principle*, whereby words tend to cluster into more or less fixed chunks, and such chunks often express relatively unitary meanings. On this view, the lexicographer would owe it to the user to offer a fair treatment of such *multi-word expressions* (MWEs; also *multi-word items*, *units*, or just *multi-words*) in a dictionary, giving such complex lexical items the same status as has so far been the privilege of items lexicalized in orthographically simplex words. In English, common formations of this type include noun compounds, phrasal verbs and 'idioms' in a narrower sense (the broader sense including all of the above). Multi-word sequences of the less fixed type are usually classified as collocation, which (when defined more narrowly) differs from the previously given types of multi-words in terms of (1) semantics, in that it does not typically denote a unitary concept, but rather a complex one; and (2) structure, in that it tends to be less deterministic and more flexible. A broader, distributional view of collocation might encompass all of the preceding types of word chunks, and so this view is not as helpful for lexicographers, who usually prefer working with finer categories.

The topic of the present study is multi-word expressions with more or less unitary meaning. Specifically, the question is where to place (the lexicographer's perspective) and find (the user's perspective) such items in the dictionary. If we accept that the prevailing lexicographic tradition for languages with alphabetic writing systems is to arrange mostly single-word headwords alphabetically, then we need to decide under which orthographic word one should place multi-words, assuming that the full treatment cannot be given under every single constituent word. A related question is under which lemmas a restricted (brief) treatment, usually in the form of a cross-reference, should be offered.

Dictionary compilers may choose to adopt a variety of approaches, taking into account word order (under the first content word is a frequent solution), word class (usually prioritizing nouns, sometimes verbs) and word frequency (listing multi-words under the less frequent components). Often, though, no uniform strategy is declared in the preface and none can be generalized from a mere inspection of the entries.

1.1   Previous studies on how users look up multi-words

Before user studies became mainstream, metalexicographers offered guidelines based on intuition. Careful attention is given to the issue of MWE placement in Zgusta et al. (1971: 269-270) in the form of four principles. First, multi-words should not be included in the entries for articles, prepositions and *be* as a copulative verb. Further, Zgusta et al. claim that preference should be given to component words which are semantically least clear in the context of the MWE. This principle may be hard to apply in practice, being rather subjective. The third guideline warns against prioritizing attributive elements. The final recommendation is to use the (linearly) first content word within the multi-word expression. Of these guidelines, the first and fourth have become rather popular in English lexicographic practice. A combination of the two results in a decision, sometimes mentioned in the front matter, to list multi-word expressions under the first content word.[2]

Actual user preferences in looking up multi-words have been studied by Béjoint (1981), Tono (1987), Bogaards (1990, 1991, 1992), Atkins and Knowles (1990), and Atkins and Varantola (1998). These studies will be summarized briefly below.

Béjoint (1981) investigated user look-up preferences of French students of English using a list of eight English multi-word expressions (*artificial insemination*, *boil down to*, *false alarm*, *magnetic tape*, *come down with*, *lose sight of*, *rid of*, *fountain pen*). He found that (1) learners would prefer *not* to have separate entries for compounds; (2) in nominal compounds the noun is preferred; and (3) in what Béjoint terms *verbal compounds*, French students preferred verbs over adverbs and prepositions, but in the one case of *lose sight of*, which also included a noun, there was a slight preference for the noun.

Tono (1987) investigated the headword choices of 129 Japanese learners of English looking at 62 idioms in specific syntactic patterns. Overall, Tono found a preference for (1) content words over function words; (2) less familiar words; and (3) words with more restricted combinability. As far as I am able to tell, familiarity and combinability were assessed impressionistically and only after the fact.

Bogaards (1990) compared the look-up preferences of a large sample of speakers of French and Dutch in 52 multi-words, and found fairly consistent but L1-dependent look-up strategies. French speakers appeared to have been guided by word frequency, going for the less frequent words, and then by syntactic structure, preferring superordinate (independent) to subordinate (depend-

ent) elements. In contrast, speakers of Dutch seemed to have looked primarily at part of speech, choosing nouns, and then adjectives and verbs, in this order. Bogaards (1991) and Bogaards (1992) explored in more detail the role of frequency differences in determining the choices of native speakers of French.

The EURALEX/AILA Research Project on Dictionary Use (Atkins and Knowles 1990; Atkins and Varantola 1998) does not turn up much useful data on looking up MWEs, primarily because it reports on only three items (a fourth item was found problematic and had to be discarded). We do learn, however, that look-up behaviour does not seem to vary by the L1 of the learner (French, Italian, German, and Spanish), and that the words that learners select for lookup are often not the ones at which the multi-words in question are listed in the leading monolingual learners' dictionaries.

1.2   Possible factors affecting the look-up of multi-words

Bogaards (1991: 204) lists seven factors that might potentially affect the look-up behaviour of dictionary users. These are:

1.   grammatical vs content words

2.   order of the words

3.   word frequency

4.   grammatical category

5.   syntactic structure

6.   semantic value

7.   idiomatic character

The general picture that emerges from past studies on looking up multi-word expressions is that users appear to be guided by word frequency, grammatical category and syntactic structure or word order.

In terms of word frequency, users tend to prefer less frequent words. It is quite likely that frequency is an indirect factor, acting through the mediation of word familiarity, but the latter is harder to measure and is a personal (subjective) attribute of limited use in dictionary design. In contrast, corpus frequency is relatively easy to measure and is collective rather than subjective.

When it comes to word class, users tend not to look up closed-class words such as articles or prepositions, and prefer content words. Amongst the content words, there may be some preference for nouns.

As far as word order is concerned, strangely enough, there is not much in the way of direct reports, and Bogaards (1991: 204) dismisses it as 'fairly uninteresting', but this factor may be hard to distinguish from syntactic role. For instance, Bogaards (1990) found that in French nouns modified by adjectives,

nouns tend to be looked up, and accounted for this in terms of a preference for syntactically superordinate elements. But, in fact, since in French adjectives typically follow nouns, it is hard to judge whether the choices made are not in fact a consequence of simple linear order — that is, users picking the first content word they come across — rather than an awareness of syntactic status. Matters are complicated even further by the same choices being explainable also in terms of a preference for nouns vis-à-vis adjectives. All in all, the role of word position seems an interesting one to examine, if only because it is taken so seriously by dictionary publishers.

Thus, in the present study an attempt will be made to investigate the role of three factors: part of speech, word order, and frequency in attracting users' attention as potential candidates for dictionary lookup.

2.   The study

2.1   Aim

The aim of the study is to assess the effect of part of speech, word position (within the MWE), and lexical frequency on the users' selection of elements in multi-word expressions that they would most readily look up.

2.2   Participants

Participants in the study were 40 Polish secondary school students aged 17 and 18, with males and females roughly equally represented. As learners of English, participants were at the B1 proficiency level as per the Common European Framework of Reference for Languages.[3]

2.3   Instrument

The principal instrument used was the Headword Choice Test designed specifically for this study. The test consisted of 36 English multi-word expressions which were, in equal measure, noun compounds (e.g. *life jacket*) and sentence idioms (e.g. *have a heart of gold*; *still waters run deep*). The items were presented on a single page laid out in two columns, 18 items in each, with instructions in Polish written across the top. There were four versions of the Headword Choice Test (labelled A, B, C, and D) differing only in the order or items, in an effort to counterbalance any order effects.

The selection of items for the Headword Choice Test was guided by the goal to have a balanced representation of words in terms of the combination of the three design factors: lexical frequency, part of speech, and word position within the MWE. And so, it was important to include both frequent and rare nouns, placed initially or otherwise within the MWE. In doing so, we were con-

strained by what is possible in the language. Function words, being closed-class items such as articles or prepositions, tend to be very frequent, and their position relative to lexical words is subject to language-specific syntactic constraints. For this reason, it was not possible to obtain data with all theoretical combinations of frequency, part of speech (POS), and word position.

For word frequency, the Corpus of Contemporary American English (COCA, Davies 2008–) was consulted. Lemmatized frequency counts were used (checked in May 2009). Raw frequency counts were subsequently categorized into three frequency bands: frequent (over 48,000 occurrences in COCA), medium (between 10,000 and 48,000 tokens), and rare (below 10,000). As a result, the 83 content words (tokens) in the MWEs included 31 frequent items, 29 medium-frequency words, and 23 rare items.

In terms of part of speech, items were selected so that at least the three major classes of content words (nouns, verbs, and adjectives) would feature in a variety of word positions and represent a range of frequencies.

When it comes to word order, the literature suggests a special role for the first content word in a multi-word. For this reason, and because the target multi-words varied in length between two and five words, word position was treated as a two-level factor: initial versus non-initial.

The materials and procedure were piloted on a small group of eight subjects similar to our participants in terms of educational level and English proficiency. No problems with the instruction, items, or procedure were noted during the pilot study. All participants in the pilot study completed the task in less than ten minutes.

2.4   Procedure

Participants were provided with printouts of the Headword Choice Test described above. They were instructed by the experimenter in their native language (Polish) to underline, for each item on the list, the one word which they would look up in a dictionary if they wished to find out the meaning of the complete expression. The same instruction was included in writing at the top of the test sheet.

Participants worked individually with no access to additional materials. Based on the results of the pilot study, they were allowed 15 minutes to complete the task. All students started at the same time and when finished, the experimenter collected the sheets. The session proceeded smoothly and all participants managed to complete the task on time.

3.   Results

All word selection data were entered into a database for further processing. Then, for each individual word token, the number of times it had been under-

lined was computed. This number corresponded to the number of subjects, out of the total of 40, who indicated by underlining that they would look up the multi-word item under this specific headword. Such headword selection counts were then analyzed in terms of how they were affected by word position in the MWE, part of speech, and lexical frequency. The measures presented in sections 3.2 to 3.4 below express the mean number of participants who indicated that they would have chosen a given word over other components of the MWE, further averaged for all words with a particular level of a design variable (e.g. initial, verb, medium frequency, etc.). This manner of computing lookup preference measures is unaffected by raw counts of particular categories and so the numbers are directly comparable within each factor.

In what follows, selection counts per item are tabulated (3.1). Further on, the roles of the three factors of interest are presented descriptively in turn (3.2-3.4). Finally, a multivariate GLM analysis is computed to assess the strength of the influence of each of the three factors and portion of variability they explain (3.5).

3.1   Headword selection data

Table 1 below gives complete data on headword selection for the 36 multi-word expressions tested. Each potential headword is followed in parentheses by the number of participants (out of 40) who underlined this particular word. For example, in item 1., *artificial insemination*, 17 participants underlined the adjective *artificial*, while 23 underlined *insemination*. Item 7. is slightly irregular: while most participants went for *red-handed*, two participants underlined just the second portion of this hyphenated word, *handed*. Similarly, in item 29., six participants underlined just the *self* portion of *self-made*. Although this does not cause major problems, hyphenated words are probably best avoided in such designs.

ID	multi-word (underline counts)
1.	artificial (17) insemination (23)
2.	at (0) daggers (37) drawn (3)
3.	back (11) door (29)
4.	be (1) a (0) wet (6) blanket (33)
5.	blind (26) trust (14)
6.	bury (17) the (0) hatchet (23)
7.	catch (10) sb (0) red-handed (28) [handed (2)]
8.	clean (6) slate (34)
9.	couch (28) potato (12)
10.	detached (34) house (6)
11.	estate (30) car (10)
12.	everyday (18) life (22)

13.	first (12) come (3) first (7) served (18)
14.	frame (35) of (0) mind (5)
15.	go (2) to (0) pieces (38)
16.	haste (27) makes (2) waste (11)
17.	have (1) a (0) clue (39)
18.	have (2) a (0) heart (33) of (0) gold (5)
19.	have (1) the (0) last (9) word (30)
20.	hide (27) and (0) seek (13)
21.	in (0) the (0) nick (31) of (1) time (8)
22.	life (10) jacket (30)
23.	like (1) headless (33) chicken (6)
24.	miss (21) the (1) point (18)
25.	new (0) broom (20) sweeps (19) clean (1)
26.	once (2) bitten (21) twice (4) shy (13)
27.	out (19) of (1) the (0) blue (20)
28.	rock (27) the (0) boat (13)
29.	[self (6)] self-made (32) man (2)
30.	sliding (30) doors (10)
31.	still (5) waters (19) run (3) deep (13)
32.	teeter (30) on (0) the (0) brink (10)
33.	time (20) is (0) money (20)
34.	train (17) of (0) thoughts (23)
35.	with (0) open (10) arms (30)
36.	world (13) cup (27)

Table 1:   Target multi-word expressions with selection counts.

An examination of the selection counts suggests that, as in most previous studies, Polish learners of English tend to ignore function words and very frequent words. This becomes even clearer if we focus on the items that all participants ignored (i.e. they were never underlined) in looking up the target multi-word expressions (**Table 2** on the next page).

Those items tend to be frequent function words or relatively delexicalized verbs (*is*, *made*). Other such semantically shallow verbs (*have*, *go*) were underlined only once or twice. The item *sb* is something of a special case, being an abbreviation for *somebody* that is most often used in dictionary metalanguage and other language-teaching materials, but its status as a regular word is questionable. The article *the* is not on the list: while most instances of it were ignored, it was underlined by a single participant in *miss the point*. Such cases emphasize the point that user behaviour is to some degree erratic, and no uniform policy on its own will ensure that *all* users will fully benefit from the entries, however well structured.

word	POS	frequency band	position
a	article	frequent	noninitial
and	conjunction	frequent	noninitial
at	preposition	frequent	initial
in	preposition	frequent	initial
is	verb	frequent	noninitial
made	verb	frequent	noninitial
on	preposition	frequent	noninitial
(sb)	(pronoun/metalng)	(rare)	(noninitial)
to	preposition	frequent	noninitial
with	preposition	frequent	initial

Table 2:   Words never underlined by participants.

Beyond the above observations, it is hard to make reliable generalizations by just scrutinizing tabulated count data. Therefore, we will now attempt to examine how headword selections depend on the three design factors: word position, part of speech, and word frequency.

3.2   Word position

The position of the word within the MWE did not appear to make much difference to our participants. Across all word tokens in the MWEs, the average multi-word-initial word was selected by 15.4 subjects, compared with 14.7 for the noninitial word. This is an unremarkable difference that would probably have little practical significance even if found to be statistically significant (detailed inferential statistics follow in section 3.5 below).

Thus, perhaps somewhat surprisingly, our Polish learners did not exhibit a marked preference for looking up initial components of multi-words. This would indicate that the frequent practice of dictionary makers to list multi-words under the first (content) word is of limited utility, at least for Polish learners.

3.3   Part of speech

Unlike word position, part of speech appears to have had a non-trivial impact on users' decision as to which word to look up (see **Figure 1**). Nouns come out at the top, with a mean of 21.1 selections falling on the noun. Adjectives are the second most preferred word class (16.8), ahead of verbs (10.7). The least often looked up word classes are adverbs (5.0) and prepositions (3.0) (this line-up excludes articles, conjunctions and pronouns, which were not underlined at all, and for which there is little data).
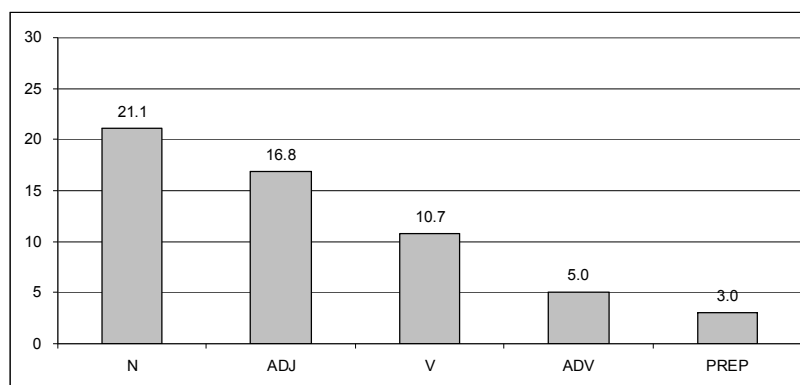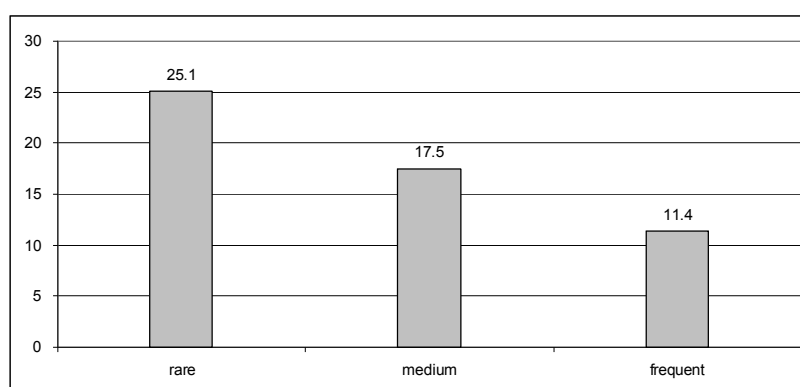
Figure 1:   Lookup preference (in mean selection counts) by part of speech.

The rather poor standing of verbs compared to adjectives is perhaps somewhat surprising. Possibly, this may be related to the relative semantic vagueness of verbs in multi-word expressions.

3.4   Word frequency

Word frequency as expressed in frequency bands again appears to have played a role in guiding the participants' decisions as to which words to look up (see **Figure 2**). Words in the *rare* category registered the highest mean selection count (25.1). Medium frequency words received an average of 17.5 selections, with 11.4 being the figure for frequent words. We will revisit the role of frequency in more detail in section 3.5.1 below.

Figure 2:   Lookup preference (in mean selection counts) for rare, medium, and frequent words.

3.5   A factorial analysis

To assess more systematically the degree of influence that word position, part of speech and word frequency have on the likelihood of the word being selected when looking up MWEs, a factorial General Linear Model (GLM) analysis was conducted on word selection counts as the dependent variable, with the three design factors as predictor variables. This analysis was conducted with the help of the Statistica 8 software suite.

For those unfamiliar with General Linear Modelling, for practical purposes it can be thought of as a generalization of Analysis of Variance (ANOVA) which allows continuous factors, not just categorical ones. Looking at our data, in a conventional ANOVA we would have been forced to use discrete frequency bands as levels of the frequency factor, much as in **Figure 2**. In contrast, the GLM approach has made it possible to utilize the full frequency information and thus obtain a more complete mathematical model of reality. To make frequency figures independent of corpus size, raw frequencies were converted to items per million (ipm, a customary measure in corpus statistics). Further, to reflect the fact that psycholinguistically meaningful differences in word frequency tend to be exponential rather than linear, a common logarithm of ipm was computed.

The data for the less central syntactic categories were not complete in terms of the availability of all combinations of word frequency and word position, so could not be analyzed due to numerous empty cells in the design. For this reason, the GLM analysis was restricted to nouns, verbs and adjectives (these, however, cover 87% of the data; besides, some previous studies also ignored function words).

The results of the GLM analysis are given in **Table 3**. Readers familiar with ANOVA tables should have no problems understanding the results. The table also includes partial η^2 ('eta-squared'), a measure of effect size commonly used in similar designs, as well as observed test power, assuming an alpha level of 0.05.

	SS	df	MS	F	p	partial η^2	observed power (α=0.05)
Intercept	6446.9	1	6446.9	82.2	**0.000**	**0.540**	1.00
Frequency	2680.0	1	2680.0	34.2	**0.000**	**0.328**	1.00
POS	557.4	2	278.7	3.6	**0.034**	**0.092**	0.64
Position	204.7	1	204.7	2.6	0.111	0.036	0.36
POS*Position	233.4	2	116.7	1.5	0.233	0.041	0.31
Error	5488.0	70	78.4				

Table 3:   A three-way GLM analysis of word lookup preference with word frequency, word position and part of speech as factors. Factors in bold are statistically significant.

Table 3 indicates that apart from the intercept (a constant non-zero component, as it were), the two design factors that reach significance are frequency and part of speech. However, the effect size for the latter is much smaller than for the former, which roughly means that frequency predicts a greater portion of the participants' lookup preference. The role of part of speech has already received sufficient coverage in 3.3 above, so let us now turn our attention to frequency.

3.5.1 Frequency

Figure 3 plots word selection counts for individual words against their corpus (COCA) frequency data. Frequency is expressed as a common logarithm of items-per-million, a relative frequency measure often preferred in corpus statistics because of its independence of corpus size.
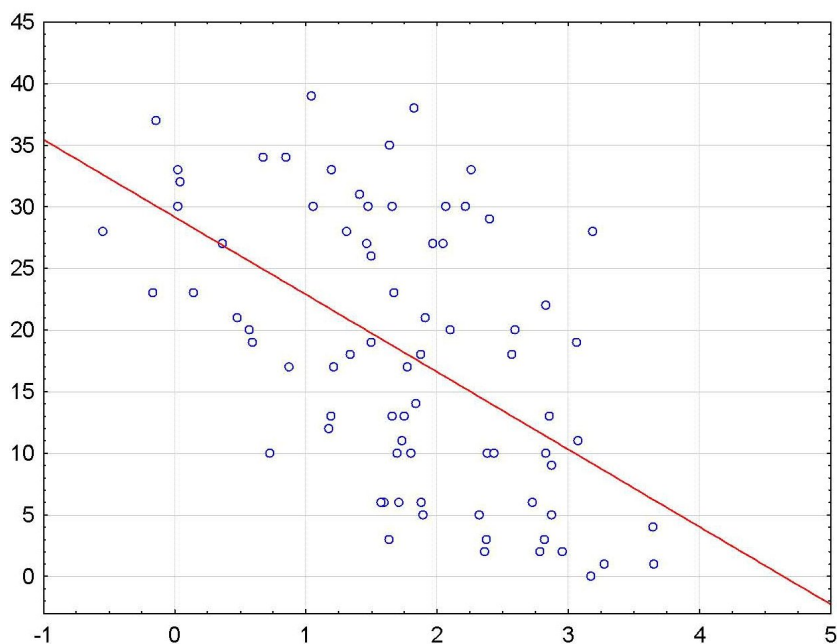

Figure 3:   Scatterplot of word selection counts (a measure of lookup preference) against the common logarithm of relative word frequency, with a regression line fitted (count = 29.2 – 6.3 × log(ipm)).

It can be seen that, in broad outline, the lower the frequency, the greater the tendency for the word to attract attention. To formalize this tendency, a regression line was fitted, and it predicts the word selection count as the intercept of 29.2 minus 6.3 times the logarithm of normalized frequency (formulaically,

count = 29.2 − 66.3 × log(ipm)). While the data points appear to cluster along the regression line, it is also true that they do so rather loosely. This means that lexical frequency only predicts a relatively modest portion of the look-up decisions. There are other factors at play, including of course part of speech. We should also bear in mind that corpus frequency is only a general indicator of word familiarity. Learners are likely to be more directly guided by how familiar a lexical item appears to them, and while the number of times they have encountered a word certainly plays an important role, everyone's experience with words is different. Finally, learners of a language are probably exposed to types of texts in proportions different from those reflected in a general corpus.

3.5.2   Part of speech by word position interaction

The interaction of part of speech by word does not reach significance ($F_{(2, 70)}$=1.5, p=0.2), so one can only speak of tendencies here. The graph (**Figure 4**) patterns into what is often referred to as a crossed interaction. For nouns, it does not matter if the noun is phrase-initial or not. For adjectives and verbs, however, there does seem to be some (albeit not significant) preference in the sample for the initial position.
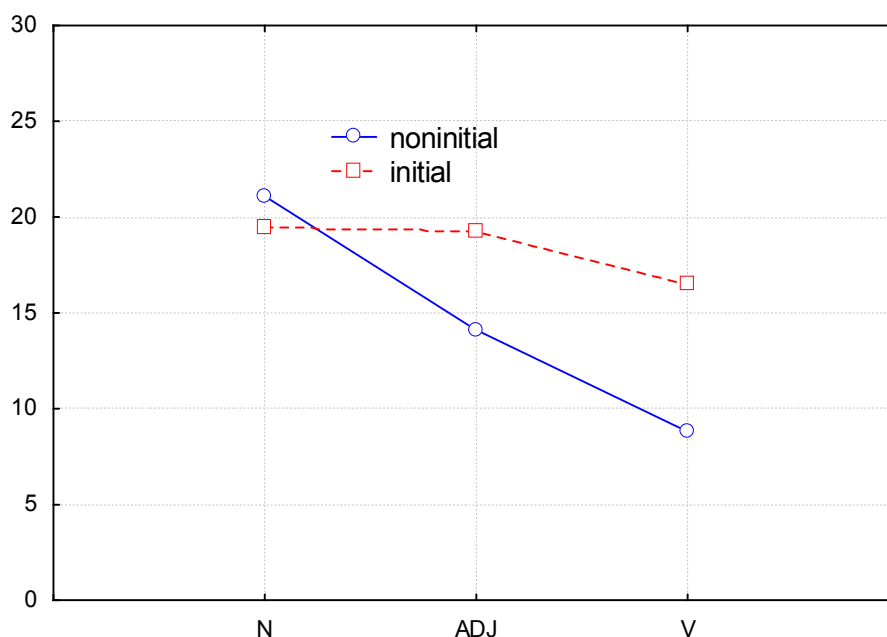
Figure 4:   Interaction plot of part of speech and word position.

Perhaps this pattern means — though at present this is little more than a guess — that participants mostly looked for unfamiliar words and then nouns, but if

these strategies did not yield a clear winner, initial word position may have come into play.

4.   Discussion

When faced with a known multi-word expression, Polish learners of English prefer to look up low-frequency words found in the MWE, probably because those are the words they tend to be less familiar with, and/or because they realize that common words often have very long entries where it is easy to miss something. Apart from the frequency, learners are guided by part of speech, preferring nouns, and then adjectives and verbs, in this order. They tend to ignore function words (articles, prepositions, pronouns) and adverbs, as well as verbs in their delexicalized uses.

Our findings on the whole concur with those obtained in previous studies for native speakers of other languages. The role of frequency features in all investigations, with the possible exception of native speakers of Dutch in Bogaards (1990), and it is telling that in our study frequency stands out as the most robust predictor of headword selection (partial η^2 = 0.328, **Table 3.**). The noun > adjective > verb hierarchy tallies with that noted by Bogaards (1990) for Dutch speakers. The potential POS-dependent role of word position has not been noted before, but this effect was not significant in our study.

Not all the findings overlap, though. On a detailed level, one of the items included in the present study, *artificial insemination*, was also tested by Béjoint (1981). He found a very clear preference (93%) for *insemination*, but in the present study the preference for this word was only marginal (58%). The disparity could be due to the different L1 (French versus Polish), or to divergent dictionary cultures (regular users more or less consciously adapt to what they encounter in dictionaries), or else — perhaps most likely — to a difference in the level of participants (secondary school students versus English majors at university).

This study suffers from a number of limitations. Most obviously, it is limited to Polish learners of English at a specific level.

The task does not exactly mimic an actual look-up situation. As in all previous studies, participants were asked to mark words rather than actually look them up in a dictionary. The advantage of the underlining task is that it is much quicker than actually looking words up, and thus it frees up the time in which to test a greater number of items, but there is no guarantee that learners operate in exactly the same way in the two situations.

Finally, MWEs are presented out of context, which is not how users would encounter them in real texts. In a broader context, learners may not realize they are dealing with MWEs and, instead, believe that they have a problem understanding some sense of a simplex word. It is, however, possible that in such a case they would follow similar strategies in selecting the word to look up.

5.   Implications for lexicographers

The present findings suggest that lexicographers, in deciding where to treat an MWE in full, should be guided primarily by word frequency, going for the least frequent constituent. Doing so should not pose much of a practical problem since in this day and age dictionary compilation is already heavily corpus-based. There may even be potential for a degree of automation here (Kilgarriff et al. 2010). Where there is no clear[4] lowest-frequency word, nouns should be given priority, but in those instances it might be wise to duplicate the full treatment under the second least-frequent item. Cross-references should be given at all nouns, adjectives, and verbs except extremely frequent ones such as *be* or *have*.

All these decisions on the treatment of multi-word-expressions should be described in the front matter of the dictionary. Even if the average user will not make good use of that information, there is a chance that their teacher might.

6.   MWEs in paper and electronic dictionaries

The issue of where to place multi-word expressions is a particularly relevant one for paper dictionaries, where restrictions of space make it rather impractical to present such items under many headwords at the same time. If one has to pick one lemma under which to embed the MWE, it is important that it is a lemma that most users would expect the expression to be placed under. Other lemmas can, and often do, include cross-references to the headword with the full treatment, giving users access to the expression, even if through an indirect route.

An unorthodox solution was adopted in *Cambridge International Dictionary of English* (CIDE, Procter 1995): this dictionary included a complete index of multi-word expressions in a separate section. Later editions did not retain this feature, and such an index is probably not an effective solution.

In electronic dictionaries it is perfectly possible to store an MWE in a single place, but present the full treatment under multiple lemmas. While this is not a huge technical problem, it is not at all obvious that this is indeed the best option, as doing so would significantly inflate entries, making them harder to navigate. This is especially important on devices with small displays, such as mobile phones, where presentation space is radically limited (Lew 2010: 299, in press). Thus, the issue of which component word of an MWE is the one users would most readily look up remains at least partially relevant for electronic dictionaries. It will become less of a problem once the dictionary can reliably recognize multi-word items typed directly into the search box. In fact, such a capability is slowly becoming a reality (Lew 2011, 2012), though progress is hampered by the fact that multi-word expressions often exhibit significant variation in form.

Still, success in the above case would be contingent on the dictionary users realizing that they are dealing with a multi-word item in the first place. There is no doubt that MWEs sometimes go unrecognized, and yet users may still choose to look up one of their components when faced with a comprehension problem they see as being due to a particular problem word within the scope of the MWE. In such a case, they may still chance upon the MWE within the entry, provided it is salient enough. Thus, felicitous placement of MWEs remains important even in those electronic dictionaries which are capable of finding them independently of headwords.

7.   Educating dictionary users

Dictionary users in formal educational settings should be given training in dictionary (reference) skills (Lew and Galas 2008; Bae 2011; Ronald and Ozawa 2011). As part of that training, they should be made aware of the importance of multi-word expressions and taught to identify them in texts. They should receive hands-on practice on how to effectively find MWEs in dictionaries. Further, users should become aware that a good candidate to start the search is the word that looks the least familiar, but if this fails, they should try the noun. Regular users of a specific dictionary should make an effort to find out what its MWE placement strategy is, if there is one (of course, explicit advice in the front matter will help, see 5 above). For electronic dictionaries, they should check if multi-word expressions may be typed directly into the search box, and if so, follow this strategy. If this does not work, they might consider switching to a dictionary that does offer this functionality.

Acknowledgements

Dedication

This article is dedicated to the memory of Paul Bogaards.

Notes

1.   In this model individual words neatly fill the terminal nodes of a syntactic structure, with word combinability mostly restricted to syntax.

2. For example, *Macmillan English Dictionary* (Rundell 2007: ix) instructs users to '[l]ook for fixed expressions at the entry for the first main word in the expression'.

3. The CEFR is a system proposed by the Council of Europe which aims at harmonizing standards of attainment in foreign language learning.

4. Bogaards (1992) proposes a minimum difference in rank of about 2500 for French, but I have doubts whether a difference so expressed is a useful measure across a broad range of frequency. For example, a difference between frequency ranks of 10 and 2510 is dramatic, but one between 50,000 and 52,500 will be rather hard to notice, if not somewhat arbitrary.

References

Atkins, B.T.S. and F.E. Knowles. 1990. Interim Report on the EURALEX/AILA Research Project into Dictionary Use. Magay, T. and J. Zigány (Eds.). 1990. *BudaLEX '88 Proceedings. Papers from the 3rd International EURALEX Congress, Budapest 4–9 September 1988:* 381-392. Budapest: Akadémiai Kiadó.

Atkins, B.T.S. and K. Varantola. 1998. Language Learners Using Dictionaries: The Final Report on the EURALEX/AILA Research Project on Dictionary Use. Atkins, B.T.S. (Ed.). 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators:* 21-81. Lexicographica Series Maior 88. Tübingen: Max Niemeyer.

Bae, S. 2011. Teacher-Training in Dictionary Use: Voices from Korean Teachers of English. Akasu, K. and S. Uchida (Eds.). 2011. *ASIALEX2011 Proceedings, LEXICOGRAPHY: Theoretical and Practical Perspectives:* 45-55. Kyoto: Asian Association for Lexicography.

Béjoint, H. 1981. The Foreign Student's Use of Monolingual English Dictionaries: A Study of Language Needs and Reference Skills. *Applied Linguistics* 2(3): 207-222.

Bogaards, P. 1990. Où cherche-t-on dans le dictionnaire? *International Journal of Lexicography* 3(2): 79-102.

Bogaards, P. 1991. Word Frequency in the Search Strategies of French Dictionary Users. *Lexicographica* 7: 202-212.

Bogaards, P. 1992. French Dictionary Users and Word Frequency. Tommola, K.H., K. Varantola, T. Salami-Tononen and J. Schopp (Eds.). 1992. *EURALEX '92 Proceedings. Papers submitted to the 5th EURALEX International Congress, Tampere, Finland:* 51-59. Tampere: Department of Translation Studies, University of Tampere.

Davies, M. 2008–. The Corpus of Contemporary American English: 425 Million Words, 1990–Present. *Available online at http://corpus.byu.edu/coca/.*

Kilgarriff, A., V. Kovár and P. Rychlý. 2010. Tickbox Lexicography. Granger, S. and M. Paquot (Eds.). 2010. *eLexicography in the 21st Century: New Challenges, New Applications. Proceedings of ELEX2009:* 411-418. Louvain-la-Neuve: Cahiers du CENTAL.

Lew, R. 2010. Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries. *Lexikos* 20: 290-306.

Lew, R. 2011. Online Dictionaries of English. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography: The Internet, Digital Initiatives and Lexicography:* 230-250. London/New York: Continuum.

Lew, R. 2012. How Can We Make Electronic Dictionaries More Effective? Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography:* 343-361. Oxford: Oxford University Press.

Lew, R. In press. Space Restrictions in Paper and Electronic Dictionaries and Their Implications for the Design of Production Dictionaries. Banski, P. and B. Wójtowicz (Eds.). In press. *Issues in Modern Lexicography.* München: Lincom Europa.

Lew, R. and K. Galas. 2008. Can Dictionary Skills Be Taught? The Effectiveness of Lexicographic Training for Primary-School-Level Polish Learners of English. Bernal, E. and J. DeCesaris (Eds.). 2008. *Proceedings of the XIII EURALEX International Congress, Barcelona, 15–19 July 2008*: 1273-1285. Barcelona: Universitat Pompeu Fabra.

Nowakowski, M. 1990. Metaphysics of the Dictionary Versus the Lexicon. Tomaszczyk, J. and B. Lewandowska-Tomaszczyk (Eds.). 1990. *Meaning and Lexicography:* 3-19. Amsterdam/Philadelphia: John Benjamins.

Procter, P. et al. (Ed.). 1995. *Cambridge International Dictionary of English.* (CIDE.) Cambridge: Cambridge University Press.

Ronald, J. and S. Ozawa. 2011. Electronic Dictionary Use: Identifying and Addressing User Difficulties. Akasu, K. and S. Uchida (Eds.). 2011 *Asialex2011 Proceedings, Lexicography: Theoretical and Practical Perspectives:* 436-446. Kyoto: Asian Association for Lexicography.

Rundell, M. (Ed.). 2007. *Macmillan English Dictionary.* Second Edition. London: Macmillan. (MED2)

Tono, Y. 1987. *Which Word Do You Look up First? A Study of Dictionary Reference Skills.* Unpublished M.Ed. dissertation. Tokyo Gakugei University, Tokyo.

Zgusta, L., V. Cerný, Z. Hermanová-Novotná and D. Heroldová. 1971. *Manual of Lexicography.* Janua Linguarum. Series Maior 39. The Hague: Mouton.