ISSN:2828-6235
Impact factor:13.77

International Journal of Literature and Language Studies(IJLLS)

USE AND SEMANTIC GROUPS OF IDIOMS, PROVERBS AND PHRASAL VERBS IN ENGLISH PRESS LANGUAGE: A CORPUS-BASED STUDY OF THE GUARDIAN

Zakhiriddin Yusupov Baxtiyorovich

A teacher at Samarkand State Pedagogical Institute, Samarkand, Uzbekistan

Abstract. Press language uses fixed expressions to make texts short, lively and evaluative. This article studies how idioms, proverbs and phrasal verbs are used in English press language and how they can be grouped by meaning. It describes a corpus-based method for articles from The Guardian (culture, sport and world news) and presents a framework of semantic groups for the three types. The article is the first part of a larger English–Uzbek contrastive study of press language.

Keywords: press language; idiom; proverb; phrasal verb; semantic groups; corpus linguistics

1. Introduction

Newspaper writers have little space and must hold the reader's attention. For this reason, they often use fixed expressions, which carry a whole idea in a few words and add an evaluation or an image. Three types are especially important in English: idioms (for example, "break the bank"), proverbs (for example, "no smoke without fire") and phrasal verbs (for example, "step down"). Idioms are word groups whose meaning cannot be found by adding the meanings of the single words (Fernando, 1996; Moon, 1998). Proverbs are short, well-known sentences that give advice or a general truth (Mieder, 2004). Phrasal verbs combine a verb with an adverb particle, and their meaning is often not clear from the parts (Biber et al., 1999).

Earlier studies show that these types behave differently. Moon (1998) found in a large corpus that most idioms are rare, while Gardner and Davies (2007) showed that a small group of verbs and particles forms most phrasal verbs. Proverbs in the media are often changed for humour or surprise (Mieder & Tóthné Litovkina, 1999). However, few studies compare all three types in one press corpus, and almost none connect the results with Uzbek. This article gives the English side of such a comparison and asks three questions:

RQ1. How often, and in what forms, are idioms, proverbs and phrasal verbs used in The Guardian?

RQ2. Into which semantic groups can these expressions be divided?

RQ3. Does their use differ between culture, sport and world news?

2. Methods

2.1. Corpus

The corpus consists of online articles from The Guardian in three sections: culture, sport and world news. It contains [N] articles and [N] words, published between [start date] and [end date]. The same number of words is taken from each section, so that the sections can be compared fairly. The texts are cleaned (menus, captions and reader comments are removed) and stored as plain text files.

2.2. Identifying the expressions

Candidates are found in two steps. First, a concordance program such as AntConc (Sinclair, 1991, describes the method) is used to search for particles, verb forms and known expressions. Second,

the researcher checks every candidate in its context. An idiom or proverb is accepted if it is listed in a dictionary such as the Collins COBUILD Idioms Dictionary or the Oxford Dictionary of Proverbs. A phrasal verb is accepted if the particle is part of the verb and the meaning is not simply the sum of the parts. Literal uses (for example, "he walked out of the room") are counted separately and not included in the main results.

2.3. Classification and analysis

Each accepted item is coded for type, section and form (full, shortened or changed by the writer). For semantic grouping, idioms and proverbs are classified by the source area of their image, following the idea of source domains in conceptual metaphor theory (Lakoff & Johnson, 1980). Phrasal verbs are classified by the general meaning they express in the text. Frequencies are normalised per 10,000 words so that sections of different size can be compared.

Three decisions make the method easier to repeat. First, the same dictionaries are used for the whole corpus. Second, uncertain items are listed in a separate table and discussed, not silently included or removed. Third, all search terms and counting rules are kept in an appendix, so that the same procedure can be applied later to the Uzbek corpus.

3. Results

This section gives the structure of the results. Highlighted fields are to be completed with the corpus data.

3.1. Frequency by type and section (RQ1, RQ3)

Table 1. Frequency of fixed expressions per 10,000 words (to be completed from the corpus)

Type	Culture	Sport	World news	All
Idioms	[__]	[__]	[__]	[__]
Proverbs	[__]	[__]	[__]	[__]
Phrasal verbs	[__]	[__]	[__]	[__]
Total	[__]	[__]	[__]	[__]

3.2. Forms and position in the article

The form of each item is also reported: full form [__]%, shortened form [__]%, changed form [__]%. Headlines and the first paragraph are compared with the main text to show where the expressions appear most often: headlines [__]%, first paragraph [__]%, main text [__]%. Changed forms (for example, a proverb with one word replaced) are listed with their original forms, because they show how writers play with the reader's knowledge.

3.3. Semantic groups (RQ2)

Table 2. Semantic groups with typical examples (frequencies to be added from the corpus)

Group	Typical items of the group	Frequency
Idioms: sport and games	move the goalposts; level playing field; on the ball	[__]
Idioms: war and conflict	bite the bullet; under fire; fight a losing battle	[__]

Group	Typical items of the group	Frequency
Idioms: money and trade	break the bank; pay the price; cut corners	[__]
Idioms: journey and direction	turn the corner; at a crossroads; back to square one	[__]
Idioms: weather and nature	tip of the iceberg; storm in a teacup	[__]
Proverbs: advice and cause–effect	no smoke without fire; look before you leap	[__]
Phrasal verbs: start and end	kick off; wrap up; call off	[__]
Phrasal verbs: change of role or control	step down; take over; hand over	[__]
Phrasal verbs: increase and decrease	go up; cut back; scale down	[__]

4. Discussion

The results should be read against earlier research. If idioms in The Guardian are mostly rare, as Moon (1998) found for English in general, the press would show a stable core of a few frequent expressions and many that appear once. If phrasal verbs are the most frequent type, this would agree with Gardner and Davies (2007), who found that a limited set of verbs and particles produces most of them. For proverbs, the share of changed forms is important: a high share would support Mieder and Tóthné Litovkina (1999), who describe wordplay with proverbs as a feature of modern media.

The semantic groups are expected to show how the press sees events. Sport, war and money images can present politics and culture as competition, conflict or business, which gives the text an evaluation without direct opinion. A comparison between sections will show whether such images belong to the topic (for example, sport expressions in sport news) or are used across all topics. Phrasal verbs, in contrast, are likely to describe common processes such as starting, ending and changing, and to be less dependent on topic.

The results also have practical value. For translators, expressions with a strong image (such as sport or war images) are a known difficulty, because the same image may not exist in Uzbek, and the translator must choose between keeping the image, replacing it or explaining the meaning. For teachers, a list of the most frequent expressions in each section can help students to read English news with less difficulty and to avoid translating idioms word by word. For dictionary makers, frequency data from the press show which expressions deserve a place in a learner's dictionary of newspaper English.

The study has limits. It uses one newspaper, so the results describe The Guardian and not all English press. Identification includes manual decisions by one researcher; a second coder for a part of the data would improve reliability. The corpus is also small compared with general corpora, so rare expressions may be missed.

5. Conclusion

This article presented a method and a framework for studying idioms, proverbs and phrasal verbs in English press language: a section-balanced corpus, a two-step identification procedure and a classification by semantic groups. The completed results will form the English part of a contrastive study with Uzbek press language, where the same procedure can be used for newspapers such as Xalq so‘zi. Such a comparison can help dictionary makers, translators and teachers, because it shows which expressions are common in each language and which images are shared or different.

References

1. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. Longman.
2. Fernando, C. (1996). Idioms and idiomaticity. Oxford University Press.
3. Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. TESOL Quarterly, 41(2), 339–359.
4. Lakoff, G., & Johnson, M. (1980). Metaphors we live by. University of Chicago Press.
5. Mieder, W. (2004). Proverbs: A handbook. Greenwood Press.
6. Mieder, W., & Tóthné Litovkina, A. (1999). Twisted wisdom: Modern anti-proverbs. University of Vermont.
7. Moon, R. (1998). Fixed expressions and idioms in English: A corpus-based approach. Clarendon Press.
8. Sinclair, J. (1991). Corpus, concordance, collocation. Oxford University Press.