

Discoursal Functions of Proverbs. A Corpus-based Study

Gloria CORPAS PASTOR
Translation Studies. University of Malaga

ABSTRACT

Although proverbs play an important role in regulating the verbal behaviour of speakers, they have been almost neglected in discourse analysis. In this paper we attempt to study their discoursal functions by means of two different kinds of corpora: a citation corpus of c. 2,000 examples of multi-word units and a large computerized corpus of several million tokens. Aspects like frequency and lemmatization will be dealt with as well.

1. INTRODUCTION

Proverbs belong to the category of multi-word units, i.e. lexical units composed of a minimum of two orthographical words up to a maximum of a whole sentence (either simple or complex). Some examples are *to cast a spell*, *heavy casualties*, *Better late than never*, *to call it a day* and *How do you do?*

Several kinds of multi-word units can be distinguished according to their function and speech act status. Among the previous examples, *Better late than never* qualifies for inclusion under the category of proverbs because it is a whole sentence, equivalent to a free utterance which is suitable to perform speech acts. Unlike pragmatic idioms such as *How do you do?*,

proverbs are not restricted to set social situations (cf. Gramley & Pätzold, 1992: 53).

Proverbs play an important role in regulating the verbal behaviour of speakers. Nevertheless, they have been almost neglected in discourse analysis. At most they have been mentioned just in passing (cf. Stubbs, 1983). In this paper we are going to study their discursual functions by means of two different kinds of corpora. Our main goals are to find out whether these units are as infrequent as the literature on the topic seems to suggest, and whether our results differ significantly according to the kind of corpus being used.

2. PROVERBS IN CORPUS LINGUISTICS

Most studies on proverbs stem from written collections, rather than from direct observation in the language. There are studies, though, that prove the rule, as the one by Hein (1951) for German. Other authors have also followed the same line of work without much success –they usually gave up due to lack of examples in their materials.

Norrick (1985: 6), for instance, found no clear instances in the Svartvik and Quirk's corpus nor in the transcribed conversations by Crystal and Davy.

His results are no surprising since the actual rate of fixed expressions in discourse appears to be very low indeed (cf. Strässler, 1982: 83; Moon 1992: 494).

In order to find out whether proverbs are as infrequent as previous studies seem to show, we have consulted two kinds of corpora: a citation corpus (CC) and three subcorpora from the Bank of English (University of Birmingham-Cobuild), a large computerized corpus: *British Books* (BB), *The Times* (TT) and *The Independent* (TI) (see "Corpora" in the appendix).

For comparative results, we have selected 20 proverbs from our citation corpus (CC) to see how frequent they are in the three textual subcorpora consulted. As regards the frequency of individual items (see table 1 in the appendix), the results show that proverbs do occur in both citation and textual corpora. Norrick's (1985) findings are certainly biased because of the (small) size of the materials examined.

We have been able to point out some differences, according to the kind of corpus used. Thus, there is a greater number of types in the CC, whereas the three sections of the textual corpus, in comparison, provide more tokens. It should be taken into account that the frequency average of individual proverbs (BB=5.2, TT=2.6, CC=1.25, TI=1.15) can be slightly higher due to the relatively high frequency of some individual units, as in the case of item 6.

As for the percentage of occurrences in the three sections of the Bank of English examined, most proverbs appear less than once per million words or

Table 1: Frequency of individual items

1.	A LITTLE KNOWLEDGE IS A DANGEROUS THING.	CC = 1	BB = 3	TT = 5	TI = 0
2.	ALL WORK AND NO PLAY MAKES JACK A DULL BOY.	CC = 2	BB = 6	TT = 0	TI = 0
3.	BEGGARS CAN'T BE CHOOSERS.	CC = 2	BB = 5	TT = 0	TI = 0
4.	BETTER LATE THAN NEVER.	CC = 3	BB = 6	TT = 2	TI = 1
5.	EARLY TO BED AND EARLY TO RISE, MAKES A MAN HEALTHY, WEALTHY, AND WISE.	CC = 1	BB = 0	TT = 0	TI = 0
6.	EASIER SAID THAN DONE.	CC = 1	BB = 28	TT = 3	TI = 1
7.	EASY COME, EASY GO.	CC = 1	BB = 5	TT = 0	TI = 0
8.	EVERY CLOUD HAS A SILVER LINING.	CC = 1	BB = 7	TT = 17	TI = 10
9.	IT MAKES ALL SORTS TO MAKE A WORLD.	CC = 1	BB = 1	TT = 0	TI = 0
10.	IT'S AN ILL WIND THAT BLOWS NOBODY ANY GOOD.	CC = 1	BB = 7	TT = 3	TI = 2
11.	ONCE BITTEN, TWICE SHY.	CC = 1	BB = 1	TT = 4	TI = 0
12.	OUT OF SIGHT, OUT OF MIND.	CC = 1	BB = 5	TT = 4	TI = 1
13.	TALK OF THE DEVIL, AND HE IS BOUND TO APPEAR.	CC = 1	BB = 1	TT = 0	TI = 1
14.	THE EARLY BIRD CATCHES THE WORM.	CC = 1	BB = 0	TT = 0	TI = 0
15.	THE LEOPARD CANNOT CHANGE ITS SPOTS.	CC = 1	BB = 6	TT = 2	TI = 2
16.	THE PEN IS MIGHTIER THAN THE SWORD.	CC = 1	BB = 0	TT = 0	TI = 0
17.	TIME WILL TELL.	CC = 1	BB = 19	TT = 8	TI = 4
18.	TWO'S COMPANY, THREE'S A CROWD.	CC = 2	BB = 0	TT = 2	TI = 0
19.	WHERE IGNORANCE IS BLISS, IT'S FOLLY TO BE WISE.	CC = 1	BB = 1	TT = 0	TI = 0
20.	YOU CANNOT TEACH AN OLD DOG NEW TRICKS.	CC = 1	BB = 3	TT = 2	TI = 1

Table 2: Tokens per million words

1.	A LITTLE KNOWLEDGE IS A DANGEROUS THING.	BB = 0.09	TT = 0.5	TI = 0
2.	ALL WORK AND NO PLAY MAKES JACK A DULL BOY.	BB = 0.19	TT = 0	TI = 0
3.	BEGGARS CAN'T BE CHOOSERS.	BB = 0.16	TT = 0	TI = 0
4.	BETTER LATE THAN NEVER.	BB = 0.19	TT = 0.2	TI = 0.2
5.	EARLY TO BED AND EARLY TO RISE, MAKES A MAN HEALTHY, WEALTHY, AND WISE.	BB = 0	TT = 0	TI = 0
6.	EASIER SAID THAN DONE.	BB = 0.90	TT = 0.3	TI = 0.2
7.	EASY COME, EASY GO.	BB = 0.16	TT = 0	TI = 0
8.	EVERY CLOUD HAS A SILVER LINING.	BB = 0.22	TT = 1.7	TI = 2
9.	IT MAKES ALL SORTS TO MAKE A WORLD.	BB = 0.03	TT = 0	TI = 0
10.	IT'S AN ILL WIND THAT BLOWS NOBODY ANY GOOD.	BB = 0.22	TT = 0.3	TI = 0.4
11.	ONCE BITTEN, TWICE SHY.	BB = 0.03	TT = 0.4	TI = 0
12.	OUT OF SIGHT, OUT OF MIND.	BB = 0.16	TT = 0.4	TI = 0.2
13.	TALK OF THE DEVIL, AND HE IS BOUND TO APPEAR.	BB = 0.03	TT = 0	TI = 0.2
14.	THE EARLY BIRD CATCHES THE WORM.	BB = 0	TT = 0	TI = 0
15.	THE LEOPARD CANNOT CHANGE ITS SPOTS.	BB = 0.19	TT = 0.2	TI = 0.4
16.	THE PEN IS MIGHTIER THAN THE SWORD.	BB = 0	TT = 0	TI = 0
17.	TIME WILL TELL.	BB = 0.61	TT = 0.8	TI = 0.8
18.	TWO'S COMPANY, THREE'S A CROWD.	BB = 0	TT = 0.2	TI = 0
19.	WHERE IGNORANCE IS BLISS, IT'S FOLLY TO BE WISE.	BB = 0.03	TT = 0	TI = 0
20.	YOU CANNOT TEACH AN OLD DOG NEW TRICKS.	BB = 0.09	TT = 0.2	TI = 0.2

none at all (see table 2 in appendix). The average ratio is BB=0.16, TT=0.23, TI=0.23. Results such as this correspond with the findings in Moon (1992) for fixed expressions in general.

3. LEMMATIZATION

The tables in the appendix show the 20 proverbs under consideration in their canonical form as they appear in Simpson (1988). Nevertheless, these units usually appear in a modified form in actual discourse. More than a half of their total occurrences are either shortened or creatively manipulated in an allusive way. This is a characteristic feature of all multi-word units in most languages (cf. Koller, 1985). In fact, such occasional modifications constitute the typical way of using multi-word units in actual discourse (cf. Sinclair, 1991: 104; Wotjak, 1992: 5). There are plenty of creative manipulations both in the citation and textual subcorpora examined. Some of them are shortenings (examples 1, 2, 3), which may give rise to new multi-word units, and modifications, creative extensions and manipulations of all kinds (examples 4, 5, 6):

- (1) Well, at least that way he would learn if it was a sensible method of getting rid of Eva... *Every cloud had...* (Sharpe, 1978: 67).
- (2) The fall in inflation is *the silver lining* of the prolonged recession (TT).
- (3) If the intelligence briefings represent a dark *cloud* hanging over the planning of allied commanders, *the silver lining* is very thin (TI).
- (4) Still, *buggers can't be choosers* as dear old Sir Winston said at the queer's wedding (Sharpe, 1978: 56).
- (5) *All work and no play makes a woman flat-chested and squinty* (BB).
- (6) *When two's company, but three constitutes unlawful assembly* (TI).

Proverbs are especially prone to modifications of their base forms. Whatever the reasons which determine their actual form in discourse, creative manipulations and shortenings should be taken into account in corpus linguistics: modifications such as these make it very difficult to retrieve proverbs by using simple concordancing methods like OCP or TACT.

4. FUNCTIONS

By discoursal functions of proverbs I understand «their discoursal value in terms of the contribution each expression makes to the information and evaluation structure of its co-text» (Moon, 1992: 493). According to her exemplification of the five kinds of functions performed by fixed expressions, proverbs seem to be mainly evaluative –except for units like *Speak of*

the devil and he's bound to appear, which qualifies for inclusion under the situational heading.

In our materials (both the citation and textual corpora) proverbs are primarily evaluative (with a negative value in most cases), but in an indirect way –these units generate a series of conventional implicatures which appeal to a shared background and culture (cf. Strässler, 1982). For example:

- (7) «I'm all right», he managed to mutter trying to take his eyes off her lips... «I didn't sleep well. That's all.» «*Too much work and not enough play makes Jack a dull boy*», said Mrs. Biggs holding the bag limply.

(Sharpe, 1976: 37)

- (8) You may have difficulty defining your own needs and expressing your needs to others. You will be seen by other people as «*all business and no play*», humourless and «bearing the weight of the world on your shoulders»

(BB)

- (9) «Of course he was before my time but I did hear one or two rather unsavoury stories.» «Youthful excess is one thing», said the Dean, «but mutton dressed as lamb is another». «They say *the leopard doesn't change its spots*», said the Senior Tutor.

(Sharpe, 1976: 14)

Proverbs usually provide a polite form to express an impolite belief in an indirect way –they reflect the politeness principle: «Minimise (other things being equal) the expression of impolite beliefs» (Leech, 1983: 81), with special reference to the tact, approbation, agreement and sympathy maxims. Brown and Levinson (1987: 226) also mention the use of proverbs with respect to strategy no. 13 («over-generalize») in their study on «face threatening acts» as a means to express criticism in a way that is easier on face due to the traditional wisdom –and generalized advice– they contain.

While retaining their evaluative content, proverbs also contribute to the information structure of their co-texts either with or without proverbial affixes (cf. Norrick, 1985). Speakers frequently use them simply to add information (example 10), incorporating them within longer stretches of speech; to support their own assertions (example 11), whose negative evaluation may not be accepted (example 12); or to provide an explanation (example 13).

- (10) Most EC embassies will be burning midnight oil to send dispatches to their capitals, *having been bitten once* already by the surprise of the Danish referendum result.

(TT)

- (11) If he never knows what he is, he'll never know what he's missed. Is that right? Haw Haw. In England we have a saying: *ignorance is bliss*.
(Bond, 1981: 34)
- (12) «It don't hardly matter if dem girls is de limit or not, honey child, «*cause beggars cain't be choosy*». «I am not a beggar!» Fancy replied sharply, stung by the thought.
(BB)
- (13) «She never remarried?» «No. *Once bitten, twice shy*.»
(Lodge, 1992: 48)

On the other hand, proverbs often signal the discoursal structure of texts. The «conventional wisdom» they convey makes them a perfect choice for openings and closings, among other functions. In conversational situations, proverbs can be used for phatic purposes: they are a way to show our interlocutor that we are paying attention, following him and that he can go on holding the floor.

In our materials these units are used either for openings or for (pre-) closings. In the first case they are usually the titles of longer articles which state clearly their correct interpretation. They also function in texts as openings or introductory remarks of longer stretches of text (example 14).

- (14) «*Apples do not fall far from the tree...* And the egalitarianism that afflicted Orwell in his thirties was of a particular fleeting kind. That spasm of sentimental socialism... spluttered out within ten years in the last, tubercular pages of 1984.
(Harris, 1990: 91)

In the same way, they are frequently used as pre-closings or closings (example 15):

- (15) ... The time for action was 18 months ago; it is still *better late than never*.
Yours faithfully, Peter Jackson. Chairman.
(TT)

According to Schegloff and Sacks (1973: 306-307), proverbs and other kinds of fixed expressions are often used to bound a given topic before closing the conversation (the «aphoristic technique»). And, in narrative genres, they tend to occur as realizations of the coda in Labov's (1972) terms (cf. McCarthy, 1991: 138).

Finally, it should be borne in mind that a given proverb is not restricted to a particular function: they can even perform several of them at the same

time. For instance, a proverb such as *A little knowledge is a dangerous thing* conveys information (keeping its evaluative force) in example 16, where it is inserted in the structure of the sentence. In example 17 it backs the speaker's opinion and attitude. In example 18, the proverb is an evaluative comment which acts as pre-closing, whereas in 19 it acts as an introductory sentence to be reversed throughout the following co-text.

- (16) I have been deliberately critical in the appraisal of these methods to show that *a little knowledge is really quite dangerous*. No doubt these methods have worked for their inventors, much as my methods have worked for me.

(BB)

- (17) Besides, harm is inherent because it forces people to make decisions which are not for the good of the game... It represents interference by people who know only a little about the game and, as they say, *a little knowledge is a dangerous thing*.

(TT)

- (18) The hundreds of thousands watching at home many of whom may never have played the game, but enjoy it should not be allowed to think the umpire has made a bad decision that could cost a team the match. *A little knowledge is a dangerous thing*. Yours faithfully, G. D. Woodman.

(TT)

- (19) They say that *a little knowledge is a dangerous thing*, but in my short farming experience I have found the reverse. It is often a single, small piece of knowledge which saves the day.

(TT)

5. CONCLUSION

Several conclusions can be drawn from the preceding study. First, proverbs are not as infrequent as earlier literature on the topic seems to suggest. They appear both in the citation and textual corpora examined. The main difference between both kinds of corpora lies in the fact that the citation corpus contains a greater number of proverbs, whereas the density of proverbs is bigger in the textual ones. This may be explained by the fact that citation corpora provide information on unusual linguistic phenomena while textual corpora are more effective when searching for regularities in the language (cf. Summers, 1993: 183).

As for the functions performed by proverbs in discourse, they are not altered according to the kind of corpus examined. Proverbs are plurifunc-

tional: they are mainly evaluative, but they can also perform situational, organisational and/or informative functions at the same time.

Among the topics for further study, the lemmatization aspect is especially important. The creative manipulations these units can undergo pose additional problems for corpus processing and retrieval. Perhaps a promising step towards solving the problem would be to identify those stable parts (kernels) that enable the receiver to recognise and decode a given proverb in context.

APPENDIX

Corpora

CC = A citación corpus of more than 2,000 multi-word units I have excerpted from 26 novels (two thirds of them published after 1970, mostly fiction) and current newspapers, like *The Times*, *The Sun*, etc.

BB = The 31 million occurrences (tokens) collection of British books, composed of 356 books (300 authors), 257 published after 1985 and divided into fiction (128) and non-fiction (216). It is a subcorpora included in a larger corpus of English: the Bank of English (University of Birmingham, Cobuild).

TT = The 10 million occurrences of the newspaper *Times*, composed of 79 daily issues from June to September of 1992. A subcorpora from the Bank of English.

TI = The 5 million occurrences of *The Independent*, composed of 49 daily issues from 1st October to 18th November of 1990. A subcorpora from the Bank of English.

BIBLIOGRAPHY

- Bond, Edward (1981). *Narrow Road to the Deep North*. London: Eyre Methuen.
- Brown, P.-Levinson, S. C. (1987). *Politeness. Some Universals in Language Usage* (1978). «Studies in Interactional Sociolinguistics». Cambridge & New York: Cambridge University Press.
- Gramley, S.& Pätzold, K-M. (1992). *A Survey of Modern English*. London and New York: Routledge.
- Hain, M. (1951). *Sprichtwort und Volkssprache. Eine volkskundlichsoziologische Dörferuntersuchung*. Giessen: Wilhelm Schmidt.

- Harris, Martyn (1990). *Do It Again*. London: Penguin.
- Koller, W. (1985). «Die einfachen Wahrheiten der Redensarten», *Sprache und Literatur in Wissenschaft und Unterricht* 56. 26-36.
- Leech, G. (1983). *Principles of Pragmatics*. London & New York: Longman.
- Labov, W. (1972). *Language in the Inner City*. Oxford: Blackwell.
- Lodge, David. (1992) *Paradise News*. London: Penguin.
- McCarthy, M. J. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- Moon, R. (1992). «“There is reason in the roasting of eggs”: a consideration of fixed expressions in native-speaker dictionaries», in H. Tommola et al. (eds.) 493-502.
- Norrick, N. L. (1985). *How Proverbs Mean. Semantic Studies in English Proverbs*. Berlin, New York & Amsterdam: Mouton.
- Schegloff, E. A.- Sacks, H. (1973). «Opening up Closings». *Semiotica*. VIII (4). 289-327.
- Simpson, J. (1988). *The Concise Oxford Dictionary of Proverbs*. Oxford & New York: Oxford University Press.
- Sinclair, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford & New York: Oxford University Press.
- Sharpe, T. (1976). *Porterhouse Blue*. London: Pan Books.
- (1978). *Wilt*. London: Pan Books.
- Strässler, J. (1982). *Idioms in English. A Pragmatic Analysis*. Tübingen: Gunter Narr.
- Stubbs, M. (1983). *Discourse Analysis. The Sociolinguistic Analysis of Natural Language*. Oxford: Basil Blackwell.
- Summers, D. (1993). «Longman/Lancaster Language Corpus. Criteria and Design», *International Journal of Lexicography* 6, 3. 181-208.
- Tommola, H.; Varantola, K.; Salmi-Tolonene, T., y Schopp, J. (eds.) (1992). *EURALEX'92 Proceedings I-II. Papers submitted to the 5th EURALEX International Congress on Lexicography in Tampere, Finland*. Studia translologica. Ser. A. Vol. 2. 2 vols. Tampere: University of Tampere.
- Wotjak, B. (1992). *Verbale Phraseolexeme in System und Text*. Tübingen: Max Niemeyer.