The Survival of English Proverbs: A Corpus Based Account

Mieder (1995) raises some very important questions regarding the status and use of proverbs in the modern age; in particular, he asks 'Which texts from former generations are still current today?' 'What are the truly new proverbs of the modern age?' 'How familiar are people with proverbs today?' In this paper I hope to answer some of these questions using a new methodological approach – that of corpus linguistics. In addition, I hope to illustrate some of the types of modification they have undergone, to identify some of the characteristics shared by proverbs that have maintained their currency and to suggest some reasons for their survival into the next millennium.

The traditional practice for paremiologists (e.g. Whiting, 1968, 1977 1989; Simpson 1982 & 1992) has been to establish proverb forms with reference to written sources and to illustrate cases where there are alternative canonical forms. Each entry is supported by quotations, spanning a period of time, commencing with that in which the proverb first occurred in written form. While this approach may be satisfactory from a diachronic point of view - illustrating gradual shifts in proverb forms - it provides little synchronic insight into contemporary proverb use. From this point of view it is problematic for a number of reason: it is dependent on written sources alone; since the initial stage is often a sifting of earlier reference works it is reliant on a large number of texts that are not in any sense contemporary. Lastly, as the primary aim of a paremiological reference work is to identify their standard citation forms and therefore it is unlikely to include a number of the creative modifications that proverbs appear to be undergoing in contemporary English. For example, it is unlikely to include the following uses that occur in the corpus used in this study (with my italics):

Familiarity breeds adoration, not *contempt*
Over-*familiarity breeds* confusion
The hand that rocks the cradle stops the buck
Hell hath no fury like a Tory *scorned*
Hell hath no fury like an arms dealer deprived
Some politicians believe *nothing succeeds like* repetition
But *nothing succeeds like* hot sex and celebrity
The road to hell is paved with case conferences
And it's all work and no play a survey reveals
Has been very much an *all-work-and-no-play* scenario

Yet surely the authors of these texts assumed knowledge of the standard forms since, without this, it would be difficult to understand their authors' stylistic intentions such as humour and irony? Taylor (1999) point out "We know very little about the development and spread of any particular stylistic peculiarities in proverbs" and Mieder 1999 comments "Proverb scholars would do well to pay more attention to the present use of English proverbs". A contemporary account of a paremiological minimum of English proverbs should describe their creative adaptation to the stylistic requirements of modern language use. Indeed, if we include proverb modifications, we may find that the common claim that proverbs are in decline may be much less accurate than has previously been thought and that they continue to constitute an important element in what Hirsch (1987) terms "cultural literacy".

From a methodological point of view Mieder (1995) upholds the traditional practice described above with a claim that a paremiological minimum of Anglo-American proverbs can be established by counting the number of references for particular proverbs in an established reference work such as Whiting (1989). I would like to suggest an alternative methodology based on combining the data available in reference works with that available in a corpus. Corpus linguistics is an approach to language description that is based on a database or corpus of language; a corpus integrates the enormous information storage capacity of computers with a software programme that allows the user to search the database. This allows the user to establish the frequency and other statistical information regarding the occurrence of language patterns. It will allow us to search for any string of words whether it is the full standard citation form of a proverb or its variations. Proverb variations - such as those above - are all those uses of a proverb that in some way differ from the citation form. If the corpus is large and representative enough we may be in a position to realise Hirsch's (1987) goal of establishing the extent to which proverbs constitute part of a minimum of cultural knowledge for an educated speaker of a language. A corpus provides data both on proverb types and their frequency in contemporary language use and I share the view that "a dictionary of cultural literacy ought to be based on frequency analyses" (Mieder 1995)

Corpus based approaches originate in a suspicion regarding the reliability of the intuition of particular individuals as to the typical occurrence of language forms. It stands in antithetical position to traditional linguistics in which it was acceptable for specialists in the field of language to develop an argument regarding linguistic phenomena with reference to sentences that were invented often for the sole purpose of proving the very point they were aiming to illustrate. There was something inherently tautological about claiming that a particular theory accounted for the data when the data themselves had been selected for the purpose of illustrating the theory. Corpus linguistics reverses the relationship between theory and data, in that theory emerges in conjunction with the data rather than determining the data. The developments of the informational storage capacities of computers now allow us to interact with a large body of language data. Once a node or keyword is identified we can search a corpus so as to generate all the lines in which this string occurs. The node can be a word or string of words and the lines provide their collocational contexts. We can see an example for the string *new broom* in the following lines:

Leicester meet Rosslyn Park. The new broom that swept in with the new season
Avenue are back again. But the new broom seems to have swept
he recalls. `It needed a new broom, innovation and marketing. Instead,
in the desert. Under Monty's new broom, and working closely with RAF
John Robins, the new broom at Guardian, the UK composite, has
1m rights issue this week, as new-broom chairman Ron Trenter seeks to beef
Sir: In your editorial `New Broom for Ulster Unionism"
is therefore on hold until a new broom arrives at the Vatican. Given John
reap the rewards of Dieter Bock's new broom. Latest word in the City is to

The citation form *a new broom sweeps clean* occurs only twice in a 330 million word corpus as compared with 110 occurrences of *new broom* used to refer to an attribute of a human or an institution; we may, therefore, infer that *new broom* is a variation

which has become the base form of a proverb. So, in this case, the identification of a proverb variation puts us in a position to establish its base form. The base form of a proverb is a collocationally closed form in which the lexical and syntactical content cannot be further modified. In some cases this will correspond with the citation form – but this will depend on the extent to which proverb variations are both possible and have taken root to become lexicalised word strings or fixed expressions (Moon 1998a and 1998b). An appropriate corpus can enable us to identify the extent to which proverb modifications have become a feature of contemporary language and yields more empirical, quantitative evidence of proverb variations than has previously been the case. In order to be representative of a language the corpus needs to satisfy a number of criteria. It should be large, it should include both speech and writing, it should cover a range of varieties of a language and cover a range of types or genres of texts such as newspapers, books (fiction and non-fiction), radio and television broadcasts and magazines. Only when these criteria have been satisfied can we generalise from what is found in the corpus to the language as a whole.

The corpus used for this study is known as the Bank of English; it is a very large corpus of 330 million words; approximately 80% of these are written and 20% are spoken language. It is divided into 17 separate sub corpora covering newspapers (42%), books (23%), magazines (15%), television, radio and ephemera (posters, advertising, bus tickets etc.). The corpus contains various varieties of English - approximately, in the following proportions: British English 70%, American 20% and Australian 10%. It is held by the University of Birmingham (UK) on behalf of Cobuild. It contains a sophisticated software programme that enables one to search for single words or strings of words. Individual lines can be expanded so as to access the context up to 40 words either side of the node word. As it is one of the largest corpora available (rivalled only by the British National Corpus) it is especially suitable for researching a generally low frequency language phenomena such as proverbs.

The first stage was to identify candidate proverb types and for this purpose I relied on the *Concise Oxford Dictionary of Proverbs*. The procedure was to search for the frequency of all the proverbs in the dictionary. In order to ensure that all proverb variations were identified it was necessary to identify a search term for each of the proverbs. The specification of a search term is that it should identify all variations of the citation form of a proverb but avoid occurrences unrelated to the proverb. The search term was usually the kernel of a proverb form through which it may be possible to establish all related variations of the citation form. For example, a search term for the proverb *let sleeping dogs lie* could be *sleeping dogs* and one for *all that glitters is not gold* could be *all that glitters*. In some cases this was done by selecting the first clause in the proverb e.g. *birds of a feather* or *it takes all sorts*; however, this was not always possible as in some cases the initial phrase may yield a number of lines which are totally unrelated to the proverb. This is the case, for example, if one searches *strike while* for the citation form *strike while the iron is* hot or *small is* for *small is beautiful*. The query language allows the researcher to leave blanks of varying word length; so the first example above was searched using the string *strike while (1,3) hot*. This means that there may be between one and three spaces between *while* and *hot*; this is likely to reduce or eliminate word strings unrelated to the proverb. The corpus is also tagged for parts of speech; so in the second example;

small is beautiful was searched using the string *small is + ADJECTIVE*. In cases where there may be variation in the first clause of the citation form, the search form was comprised of its latter part. For example, in the proverb *do not throw pearls before swine*, an alternative verb such as *cast* could be substituted; therefore the search form is *pearls (0,1) swine*. The results of the searches were checked to ensure that all the lines could be related to a proverb and those that could not be were removed.

The search forms were employed to establish which proverbs occurred most frequently - both in their citation form and as regards variations that could be related in some way to the original proverb. All those that occurred with a frequency of at least ten tokens (either for the citation form or for the variations) were retained for further analysis. There were found to 73 proverb types that had more than ten tokens in their citation form and 110 types that had more than ten tokens in variant forms. A sample of 500 variations were selected to identify the types of modification that occurred. The primary aims of this paper are to identify the paremiological minimum of English proverbs and their most common types of variation and to suggest some characteristics of paremiological vitality – i.e. those that are likely to lead to proverb survival. The findings reported are necessarily dependent on the corpus on which they are based and may need to be modified in the future to accommodate both the expansion of this corpus and the growing availability of other corpora. However, it is hoped that the methodology described above will give us some insight into a paremiological minimum of English proverbs and their role in English cultural literacy.

The following table shows the proverbs that were used most frequently in their citation form based on Simpson (1982). While 73 types showed more than 10 tokens in the corpus we have included those 31 that occurred with a frequency of more than once per ten million words. We have chosen this measure of frequency as Moon (1998b) found that 70% of what she terms phrasal lexemes occur less than once per million words – therefore it seems that for proverbs, which constitute only one sub-type of phrasal lexeme, ten million words is a more delicate unit of measurement.

Table One: Most Frequent Proverbs in English: Citation Forms

Proverb Citation Form **Corpus: Bank of English** **Size: 330 million words**	**Tokens**	**Frequency/ ten million words**
First come, first served	179	5.4
Forgive and forget	103	3.2
Money talks	101	3.1
First things first	100	3.0
Small is beautiful	98	3.0
Better late than never	92	2.8
Horses for courses	83	2.5
Live and let live	70	2.1
Out of sight, out of mind	68	2.1
Boys will be boys	63	1.9
Knowledge is power	51	1.5
Time is money	50	1.5
Charity begins at home	50	1.5
Every man for himself	47	1.4
Time flies	44	1.3
Practice makes perfect	43	1.3
Actions speak louder than words	42	1.3
All good things (must) come to an end	42	1.3
All's well that ends well	42	1.3
Power corrupts	40	1.2
Waste not want not	40	1.2
Fight fire with fire	40	1.2
History repeats itself	38	1.2
Know thyself	38	1.2
There is no place like home	38	1.2
It takes two to tango	35	1.1
Let sleeping dogs lie	34	1.0
Silence is golden	33	1.0
Beauty is in the eye of the beholder	33	1.0
Like father, like son	33	1.0
There is no such thing as a free lunch	33	1.0

We may note some interesting points about this estimate of a paremological minimum; first, the findings are very different from those reported in Mieder (1995). For example from the top 13 proverbs in Albig (1931) none occur - although as we will see later some occurred as proverb variations. Only one of the most familiar proverbs cited in Penn, Jacob and Brown 1988:852) occurred: *all's well that end's well*. Four of the most familiar "sayings" cited in Higbee & Millard (1983: 216-219) did occur; these were *practice makes perfect, better late than never, all's well that ends well,* and *like father, like son*; e*asier said than done* had a high frequency but is not cited in Simpson (1982). Finally, only one of the proverbs that Mieder (1995) cites as high frequency occurred more than once per ten million words in their citation

form (*charity begins at home*) although a further three occur in some form of variation (see table five).

There are at least three possible explanations of the variation in these findings; the first is that since the studies reported by Mieder were based on a different methodology originating in psychology, lexicography and sociolinguistics, they produced different results. Mieder (1995) bases evidence for a paremiological minimum on the number of references cited for each proverb in Whiting (1989). The other studies referred to above are all based on informants and, given the small sample size, it is, perhaps, not surprising that there is considerable variation in findings; for example, Alberg (1931) uses 68 informants; Penn, Jacob & Brown (1988) use 278 informants; Higbee & Millard (1983) use only 50 informants. It may also be the case that familiarity judgements themselves are unreliable – particularly when the informants chosen for each of these studies were college students. One of the motivations of corpus linguistic approaches is that native speaker intuitions are too subjective and reliable and therefore lack the reliability of large corpora, so much more so when these are the intuitions of immature adults. It would seem preferable to make claims for cultural literacy on the basis of those whom we may expect to be culturally literate; and this would not necessarily include college students.

These criticisms seem quite convincing when we consider the significant differences between the findings of the above studies; for example *Make hay while the sun shines* - which occurs in the list of the top thirteen proverbs in Albig (1931) - occurs in Higbee & Millard's (1983) category of most unfamiliar proverbs. Clearly there may be an important diachronic influence here given the length of time between the two studies but would we really expect such a rapid change in status if these lists were reliable? A further possibility worth exploring is that proverbs have undergone important processes of modification that account for the low occurrence of many citation forms. For example if we search our corpus only the first phrase for each of the 13 most frequent proverbs in Albig (1931) we find variations such as:

through the knee of my jeans - *a stitch in time* may have saved ninety! I
A stitch in time from the new Remmington
The `*A Stitch In Time*" series includes Walter
sewing machines to make sure of *a stitch in time*. The syndicate, made up of
dress. The experience proved to be *a stitch in time* for real-life bride-to-be
An apple a day keeps bowel cancer at bay
AN APPLE a day may, judiciously munched in Tokyo, keep the US sanctions *away*.
An apple a day may keep lung diseases at bay
FORGET *an apple a day* - latest research has shown
in a self-obsessed world, where *do unto others* before they do unto you is
inflicts; we are a people who *do unto others*; we like action - Vietnam,
the new leaders have begun to *do unto others* what once was done to them;
I always lived by a book of rules. *Do unto others*, turn the other cheek. And
Normal Judeo-Christian rules - *do unto others*, people in glass houses

It may be illustrative in this respect to compare the number of citation forms in Albig's 13 most frequent proverbs with the number of variations of the type seen above. The results are shown below in table two:

Table Two: A Comparison of Citation Forms and Variations

13 most popular proverbs (Source: Albig 1931)	**Citation tokens**	**Variant form**	**Variant tokens**
A stitch in time saves nine.	13	A stitch in time	41
A rolling stone gathers no moss.	6	none	6
A bird in the hand is worth two in the bush.	0	A bird in the hand	21
Early to bed and early to rise, makes a man healthy, wealthy and wise.	1	Early to bed (0,1) early	7
Never put off till tomorrow what you can do today.	0	Never put off till tomorrow	2
Haste makes waste.	7	none	7
An apple a day keeps the doctor away.	10	An apple a day	25
All that glitters is not gold.	6	All that glitters	30
Do unto others as you would have them do unto you.	5	Do unto others	39
Laugh and the world laughs with you.	9	Laugh and the world	10
Birds of a feather flock together.	6	Birds of a feather	190
There's no fool like an old fool.	7	none	7
Make hay while the sun shines.	6	Make hay	48
TOTAL	76		433

We can see that in this sample of proverbs there is much more evidence of a variant form occurring than the citation form of the proverb. There are variations for ten of the thirteen proverbs and for nine of these there is a much greater likelihood of a proverb occurring in a variant form than there is in its original citation form. The most extreme example of this phenomena in my own sample was with the proverb *the last straw that breaks the camel's back* which occurs only six times in citation form but 421 times in a contracted form *the last straw*. We should then consider the findings for proverb variations and see how far they differ from those for citation forms.

In this account of proverb variation we will use the following typographic conventions: the base form of the proverb will be shown in italics; while a novel element or variation will be underlined. We identified four major types of variation from the citation form of proverbs; these are summarised in table three below:

Table Three: Types of Proverb Variation

Variation Type	Description	Example
Substitution	Lexical substitution of one element while the syntactical pattern is unchanged.	1 *Give them an inch and they will* run *a mile*. 1. *The proof of the* cake *is in the eating* 3. *Out of the frying pan and into* your wardrobe
Contraction	A clause is omitted – usually this is the second clause.	1. He is refusing to give an inch 2. When in Rome 3. Birds of a feather
Antonyms	A form of the proverb which has the opposite meaning to the original (e.g. by omission or insertion of a negative morpheme).	*1. All that glitters* is *gold* *2.* Not *letting sleeping dogs lie* *3. You* can *teach an old dog new tricks*
Expansion	Another linguistic element is inserted into the proverb	*1. Casting* synthetic *pearls before* real *swine* *2. My bark is* definitely *worse than my bite* 3. *The proof if the pudding,* they say, *is in the eating*

Table four shows the results for the types of variation found in a sample of 500 proverb variations:

Table Four: Comparison of Proverb Variations

Variation Type	Number of Occurrences	Percentage of total sample
contraction	316	63%
substitution	127	25%
antonym	28	6%
expansion	29	6%

We can see from table four that the variations summarised in table three vary in the scope of their influence. Contraction is the most common type of variation affecting 63% of all variations; exactly one in four variations are substitutions. Antonyms and expansions were of equal importance and together accounted for 12% of the total variation. It is worth noting that all the proverbs analysed for variations exhibited more than one of the variation types; and some illustrated all four types as we can see from the following lines containing variations of the proverb *all that glitters is not gold*:

A girl's best friend: All that glitters has often been regarded as (Contraction)
Of high fashion where all that glitters has not sold. Luxury in the (Substitution)
There's a lady who's sure that all that glitters is gold and she's buying a (Antonym)

To sandals and sunglasses, all that glitters is gold and black this summer. Who (Expansion)

Or these variations of the proverb *new brooms sweep clean*:

Mrs. Virginia Bottomly, the new broom a the Heritage Department, has (contraction)
Letter headline: New broom sweeps people off the streets and (substitution)
Wishes them success, it shows a new broom doesn't always sweep clean, it just (antonym)
Leicester meet Rosslyn Park. The new broom that swept in with the new season (expansion)

We may also note that in a few cases there is evidence of more than one type of variation at work; for example both contraction and antonymy are evident in the following variation of *don't teach your grandmother to suck eggs*:

Perhaps you should teach your grandmother after all, murmured Dog as they.

However, cases such as these were rare and it was generally possible to classify each of the lines according to the predominant variation type. The fact that the same citation form can undergo a range of variations and even exhibit more than one variation simultaneously should alert us to the fact that when users of English seek some form of creative modification they are not constrained by the nature of the process. In some situations they will select a substitution while in others they will select a contraction – while in a very few cases they may choose both – according to their particular rhetorical intentions. In these cases, probably typical of media language, there is an attempt at creative reinvention of the proverb citation form and, in contemporary language, proverb variation is driven by stylistic considerations that combine innovation with humour. In fact, we may even interpret these variations as a component in a general category of persuasive language of increasing use in advertising, newspapers and other contemporary media.

Given the potential importance of proverb variations in determining a paremiological minimum for English proverbs we should now consider the impact of including variant forms in our analysis of proverb frequency. Table five below shows all those proverb variations that occurred more than once per ten million words in the corpus (i.e. there were more than 33 tokens).

Table Five: Most Frequent English Proverb Variations

Proverbs and Variations (Bank of English) Corpus size: 330 million	**Primary Variation Type**	**Tokens**	**Frequency/ 10 million words**
The last straw	Contraction	421	12.8
Birds of a Feather	Contraction	190	5.8
Small is + ADJECTIVE	Substitution	118	3.6
New broom	Contraction	110	3.3
Early bird	Contraction	104	3.2
have 0,1 cake and eat it	Substitution	98	3.0
Time and tide	Contraction	92	2.8
The hand that rocks the cradle	Contraction	91	2.8
It takes two + VERB	Substitution	81	2.5
All good things	Contraction	74	2.2
Out of sight out of	Substitution	69	2.1
Silence is + ADJECTIVE	Substitution	64	1.9
opportunity knocks	Contraction	63	1.9
swings and roundabouts	Contraction/substitution	62	1.9
The proof of the pudding	Contraction/substitution	60	1.8
Once bitten	Contraction/substitution	58	1.8
sleeping dogs	Contraction/substitution	56	1.7
Give 0,3 an inch	Contraction/substitution	55	1.7
Home is where the 0,2 is	Substitution	54	1.6
Jam tomorrow	Contraction	54	1.6
History 0,2 repeats itself	Substitution	51	1.5
Make hay	Contraction	48	1.5
Every man for himself	Contraction	47	1.4
Hell hath no fury like ….	Substitution	47	1.4
All work and no play	Contraction/substitution	45	1.4
A stitch in time	Contraction	41	1.3
Like father like	Substitution	40	1.2
Life begins at + age	Substitution	39	1.2
Out of the frying pan	Contraction/substitution	39	1.2
When in Rome	Contraction	37	1.1
One law for the rich	Contraction/substitution	35	1.1

We may note that as compared with the 31 proverbs found in table one, table two also contains exactly 31 proverb variations occurring with a frequency of more than once per ten million words. The primary types of variation are contraction and substitution of an element found in the citation form of a proverb.

In five cases the original citation forms were not found at all but variations occurred more than ten times; these are: *Give someone an inch and he/she etc will take a mile* which occurred 55 times*; all is grist that comes to the mill* which occurred 31 times *as grist to the mill; jam tomorrow and jam yesterday, but never jam today* which occurred 54 times as *jam tomorrow;* and *opportunity never knocks twice at any man's door* which occurred 63 times in the contracted form *opportunity knocks.* Surely there

is a case that these phrasal lexemes should be considered as part of the paremiological minimum of English proverbs since their origin is clearly related to the proverb?

My suggestion here is that a paremiological minimum should be based on both the most frequent citation forms of proverbs (shown in table one) and on the proverb variations shown in table five. At this point, I will consider some of the stylistic features of the proverbs that enhance the possibility of their survival into the next millennium. From table one we may notice three stylistic features: brevity, alliteration and repetition. Brevity is clearly a vital characteristic of speech economy central to proverb survival and we can find evidence for this in table five as it shows that contractions are more frequently used than the longer citation forms in the majority of proverb variations. We may note that many of the citation forms in table one are already base forms and could not readily be shortened further while retaining the same meaning; this is also why we found that three of the proverbs taken from Albig (1931) resisted modification. Contraction is nothing other than a process of adaptation by proverbs to the contemporary norms of communication that place an emphasis on brevity and economy of effort. Mieder (1999) finds that the description short occurred in 18 out of 55 informants' definitions of the term proverb. It was decided to test this hypothesis by considering the full sample of 110 proverbs for which there were more than ten tokens. When we divide these into according to the number of words in the citation form of the proverb we can then calculate the average frequency according to proverb word length. Table six shows the findings:

Table Six: Proverb Length and Frequency

Length of proverb	Number of types	Average number of tokens	Average frequency / 10 million words
2 words	5	50.2	1.52
3 words	14	48.7	1.48
4 words	23	34.9	1.06
5 words	7	23.1	0.70
6 words	19	16.0	0.48
7 words	9	12.2	0.37
8 words	16	9.2	0.28
9 words	9	9.9	0.30
10 + words	8	4.8	0.15

We can see clear evidence from table six that the shorter the proverb type, the more tokens occur in the corpus. Minimal two word proverbs such as *money talks*, *time flies, power corrupts* and *know thyself*, occur the most frequently. Conversely, long proverbs such *as never put off until tomorrow what you can do today*, *the grass is always greener on the other side of the fence,* and *sticks and stones may break my bones but words will never hurt me* have to very low frequency. It seems, then, that the key factors in proverb vitality are the brevity of expression quite apart from the apparent value or truth of the semantic content. However, there are other stylistic considerations that we noted above that are found in the most frequent proverb citation forms: in particular, alliteration such as *forgive and forget; fight fire with fire; it takes two to tango; beauty is in the eye of the beholder* and *practice makes perfect*. However, repetition is perhaps, an even more important characteristic of proverb vitality: we may note: *first come first serve; first things first; live and let live; out of sight, out of mind; boys will be boys; all's well that ends well; waste not, want not; fight fire with fire; like father like son*. Of course repetition also involves alliteration and it seems that their combined effect is to serve as an mnemonic that can ensure the ongoing popularity of a proverb form. These phonetic and rhythmic features are found to a much lesser extent in the proverb variations.

It would be particularly interesting, from a comparative point of view, to know the extent to which the stylistic features of proverb vitality in English that have been identified here also apply in other languages. Particularly, since it would be relatively easy to gauge stylistic features such as length, alliteration and repetition as compared, say, with identifying proverbs with similar meanings in different languages (e.g. Charteris-Black 1995b). For English proverbs, the longer the proverb the less likely it is to survive in its citation form and the greater the need for abbreviation to maintain its vitality. Length and the presence of repetition in the form of a proverb are key factors in establishing a paremiological minimum of English proverbs. We may also conclude that it is indeed the case that proverbs are undergoing processes of creative adaptation and modification in contemporary English language use. This is important since it suggests that the use of proverbs may not, in fact, be as much in decline as may have previously been believed. Indeed it is through their ability to adapt that they demonstrate the tenacity necessary for survival.

In this paper I have described some stylistic modifications of English proverbs, in particular, contraction and substitution; the modified form has, in some cases, replaced the original citation form as the base form of the proverb as a result of contraction. I have argued that a paremiological minimum of English proverbs is best identified using a large corpus like the Bank of English and that it should include all types of proverb variation as long as the form can be related to a proverb citation form. I have claimed that length and repetition of form are central characteristics of proverbs that exhibit vitality and repetition of use. I believe that issues of style are of considerable importance to any concerned with how proverbs are used in contemporary English and in the identification of a paremiological minimum. They should also be useful to those engaged in specifying a minimum cultural knowledge of English. Indeed, we may find that, as the influence of corpora on lexicography grows, reference works will need, increasingly, to accommodate the type of variations described here.

References

Albig, William 1931 Proverbs and Social Control. Sociology and Social Research 15:527-535.

Charteris-Black, Jonathan 1995 (b) "Still waters run deep- proverbs about speech and silences: A cross linguistic perspective". DeProverbio: An Electronic Journal of International Proverb Studies.Vol 1 No 2. URL: http://info.utas.edu.au/docs/flonta/

Higbee, Kenneth L., and Richard J. Millard 1983 Visual Imagery and Familiarity Ratings for 203 Sayings. American Journal of Psychology 96:211-222.

Hirsch, E.D. 1987 Cultural Literacy. What Every American Needs to Know. With an Appendix "What Literate Americans Know" by E.D. Hirsch, Joseph Kett, and James Trefil. Boston: Houghton Mifflin Company.

Mieder, W. 1995 Paremiological Minimum and Cultural Literacy. DeProverbio: An Electronic Journal of International Proverb Studies. Vol 1 No 1. URL: http://info.utas.edu.au/docs/flonta/

Mieder Wolfgang 1999 Popular Views of the Proverb. . DeProverbio: An Electronic Journal of International Proverb Studies. Vol 5 No 2. URL: http://info.utas.edu.au/docs/flonta/

Moon. R. 1998 (a) Fixed Expression and Idioms in English. Oxford: Clarendon

Moon 1998 (b) Frequencies and Forms in Phrasal Lexemes in English. In A. P. Cowie (ed.) Phraseology: Theory, Analysis, and Applications. Oxford: Clarendon Press. Pp. 79-100

Penn, Nolan E., Teresa C. Jacob, and Malrie Brown 1988 Familiarity with Proverbs and Performance of a Black Population on Gorham's Proverbs Test. Perceptual and Motor Skills 66:847-854.

Simpson, John A. 1982 The Concise Oxford Dictionary of Proverbs. Oxford: OUP

Taylor, Archer 1999 The Style of Proverbs. DeProverbio: An Electronic Journal of International Proverb Studies. Vol 5 No 1. URL: http://info.utas.edu.au/docs/flonta/

Whiting, Bartlett Jere 1968 Proverbs, Sentences, and Proverbial Phrases from English Writings Mainly Before 1500. Cambridge, Massachusetts: Harvard University Press.

Whiting, Bartlett Jere 1977 Early American Proverbs and Proverbial Phrases. Cambridge, Massachusetts: Harvard University Press.

Whiting, Bartlett Jere 1989 Modern Proverbs and Proverbial Sayings. Cambridge, Massachusetts: Harvard University Press