Colouring Meaning

Colouring Meaning

Collocation and connotation in figurative language

Gill Philip
University of Bologna

John Benjamins Publishing Company

Amsterdam / Philadelphia

2nd proofs

For mum and dad

2nd proofs

Table of contents

2nd proofs

List of tables and figures

2nd proofs

Acknowledgements

Several people have played an instrumental role in the creation of this book.

In the first place, that it came to be written at all is thanks to Encarnacion Hidalgo Tenorio, who prodded me into writing up a monograph on my "hobby-horse" topic of colour words and connotation. Hopefully in the pages and chapters to come, it will become apparent it is in fact a very serious topic as well as being an excellent vehicle for delving into the mysterious relationship between meaning in the mind and meaning in use.

The ideas developed here, as well as having been heavily influenced by J. R. Firth and John Sinclair's writings, owe much to Bill Louw's studies into collocation, contexts of situation, and their relationship to special effects in text. I count myself lucky to have enjoyed his friendship and support over the past decade.

I would like to take the opportunity to acknowledge the contribution that the Researching and Applying Metaphor International Association (RaAM) and its members have made to my professional growth, through exposure to other fields of linguistics (especially psycholinguistics and cognitive linguistics) as well as through their enthusiasm for my work. The wide-ranging references to be found in this book are very much a result of having participated in RaAM conferences and having been exposed to a very wide range of approaches to the study of figurative language.

As far as the manuscript which you are holding in your hands is concerned, I owe a particular debt of gratitude to Geoffrey Williams. Without his detailed and insightful (yet encouraging) comments on an earlier draft, this book would have been dull as ditchwater and quite unreadable. That said, I take full responsibility for its remaining shortcomings.

Finally, thanks are due to Elena Tognini Bonelli, editor of this book series as well as the person who introduced me to corpus linguistics in the first place, and to Kees Vaes at John Benjamin's, who has been exceptionally patient as my deadlines slipped. You are all *one in a million* and, hence, statistically (as well as personally) significant.

2nd proofs

A search for meanings

1.1 Introduction

This is a book about phrases and how their meanings are affected by the addition, subtraction and repositioning of the words that form them. It does this by investigating the interrelationship between phrases and words in conventional language and in more unconventional uses. The study of language norms and their exploitations can be addressed from various angles, for various purposes. Here, the intention is to discover how creative variations of familiar phrases communicate meanings above and beyond those that are associated with the normal wording. These include associative and connotative meanings, as the title to this book indicates, but may also involve less striking effects such as making a phrase more contextually relevant, or expressing emphasis.

To this end, the reader will be guided through some of lesser-known waters of phraseology. Some knowledge of corpus linguistics is taken for granted, *viz.* what a corpus is, how it can be interrogated, and what kinds of information can be extracted automatically before human intervention and interpretation come in. However, the theoretical background of corpus phraseology will be covered in some depth. For the rest, the reader is assumed to have only a hazy knowledge of metaphor, connotation and other "secondary" meaning, and none at all of psycholinguistics. These areas of study converge and overlap in how they account for phrasal meaning, making a working knowledge of them essential to the argument presented in this book, which is that meanings choose words, and deliberate changes to those words may add to a meaning but not change its fundamental message.

The germ from which this study springs is an investigation of connotation focusing on colour words (Philip 2003). The present study is not *about* colour words, however, but about meaning, exemplified through colour words. Colour words are an ideal lexical set for testing hypotheses about meaning because their basic meanings are unambiguous: something either is or is not *pink*, or *blue*, or *green*. At the same time, colours – both as entities and as words – are associated with dozens of extended, metaphorical and/or connotative meanings, so have immense meaning potential which can be exploited creatively. Not all of

the examples used in this book feature colours, but most of the longer worked examples do. Colour word idioms and collocations make it possible to give a clear picture of the kinds of mechanisms that come into play when meanings are manipulated. But where do those meanings come from in the first place?

1.2   Meanings and words

Pick up a book near to hand, open it at random, and find the first general noun that appears on the left-hand page. Make a note of what you think it means. Now, read the full sentence in which it occurs, and consider this question: does the noun have the same meaning in that sentence context as it did when you were considering it in isolation? Whether it did or did not depends on several factors – how common the meaning is, how familiar it is to you, how often you have come across it recently, whether or not it is literal (or *basic*), the genre and topic of the text… Can you explain why you interpreted the contextualised noun in the way that you did, especially if that meaning was not the same as the context-free meaning?

One way of accounting for the difference is to consider words as containers for meanings. This is a very pervasive notion in the English language (Reddy 1979), and probably in many others too. Some words "contain" a great many meanings, while others have a more limited range of reference, but the meaning is most definitely considered to be inherent in, and part of, the word. When we try to understand a word, we delve into the container and pull out the meaning that we want. If we are not sure which meaning we want (because there is no context to guide us), we will look for the most prominent, or *salient* one, rather in the way a child will try to grab the biggest packet in a lucky dip. If we are looking for a particular meaning, we will be more selective, feeling for a package of a particular shape, size or weight to suit our needs.

This is one account of how we access meanings through words. Another one, which is diametrically opposed to it, is Sinclair's "empty lexicon" (1996a): words are devoid of meaning. This view is rather drastic, possibly provocative, certainly counter-intuitive, and perhaps better expressed as the *underspecified* lexicon (Frazier & Rayner 1990). So words are not entirely devoid of meaning, they are, instead, potentially meaningful, but the actual meaning they are to express cannot be selected without contextual cues. The principle can be illustrated with reference to the world of sport: a midfielder cannot play football alone, in fact a midfielder is not even a "midfielder" but only a person kicking a ball unless all the other members of the team are in place. And while everyone knows that it is possible to play a version of football with limited resources (e.g. two in goal, two

kicking the ball) and minimal rules, the fully-fledged game requires a set number of players in particular positions obeying (hopefully) the rules of play. Words too team up in order to realise their expressive potential and to convey specific and complex meanings.

The idea that meaning is determined by co-selection is familiar to linguistics working within the neo-Firthian tradition, but in spite of this familiarity, the idea that words are empty or underspecified often meets with resistance. This can occur especially when the empty lexicon is misconstrued as stating that words mean nothing, when in fact it states that words have no meaning. The difference is subtle but important. As every speaker of a language knows, words mean: things have names, and the function of those names is to act as verbal substitutes for their associated things. This is such a central and basic fact of language that it blinds most speakers to some other equally central and basic facts of language, i.e. that words need not refer to things at all, and that once past infancy, normally-functioning human beings do not usually use single words as substitutes for pointing at things. Once we move past using words as "lexical substitution-counters" (Firth 1935: 20), we find that meaning is only realised by words in combinations, each single word relying on its neighbours to stabilise the overall message being conveyed. This is what the empty lexicon refers to. Meaning is not held inside each discrete word form but extends over a number of words, and their co-occurrence cancels out the inappropriate or irrelevant meanings that might have been activated in other lexical company.

One of the main concerns of this book is to uncover how word meanings and the empty lexicon interact by studying conventional phrases and creative variants of them. Corpus analysis establishes the regular lexical patternings which fix meaning, and this aspect of phraseology is quite well documented. Less well explained is what happens when the "normal" patterning is modified. When modifications are made, the basic meaning seems to remain, but is often enhanced or added to, and sometimes double meanings, allusions and connotations are inferred. This study attempts to explain how this happens, considering meaning as an essentially phraseological phenomenon (cf. Sinclair 1991; Wray 2002) which sometimes allows single word choices to be significant. It is therefore an investigation of how deliberate and idiosyncratic word choices affect conventionally-established wordings-as-meanings. In other words it looks at how Sinclair's (1990) open choice principle operates *within* the idiom principle and how the mental lexicon intrudes on the empty lexicon.[1]

---

1.   The term "mental lexicon" is widely used in cognitive and psycholinguistics to refer to the inventory of lexical items – the personal dictionary – each of us has stored in our mind. Neither the term nor what it refers to are generally accepted within the neo-Firthian tradition, in which

It will be seen that creativity seems to be constrained by conventional language patternings. Creativity is often thought of as a very free act of expression, but while this may be true, it is also true that if it is not obviously meaningful, its expressive effect is diminished. Even if meaning can be extracted from such famous nonsense examples as *colorless green ideas sleep furiously* (Chomsky 1964: 149) or *'Twas brillig, and the slithy toves / Did gyre and gimble in the wabe* (Carroll 1872), neither is meaningful in the sense of conveying an unequivocal message. Creativity that is also meaningful involves exploitation of conventional, familiar language, and in doing so generates multiple layers of meaning such as that which occurs in Examples (1)–(6). The aim of this study is to determine what parameters govern exploitations like these; in short, to explain how they work.

(1) **Light relief** from the winter blues.

(2) Greens need to **grasp the nettle**.

(3) Old Lady **takes a hammer to** her furniture.

(4) "… well, it's no good **crying over spilt potion**, I suppose … but **the cat's among the pixies** now."

(5) Embittered, disillusioned, increasingly shrewish and expertly violent female workaholic seeks **horizon-dwelling male** to make nothing else matter.

(6) "Don't you know that in the Service one must always **choose the lesser of two weevils?**"

Examples (1)–(6) are all creative exploitations of fixed phrases which are elucidated in the footnote provided.[2] Substitution is perhaps the easiest to identify (Examples (4) and (6)), but creativity can just as easily involve paraphrase (Example (5)), intentional ambiguity or double meanings (Example (3)) or the use

'the meaning of words lies in their use' (Wittgenstein 1953: 80). The mental lexicon is inextricably related to the Chomskyan notion of *competence*, which is unquantifiable and therefore at odds with an empirical approach to the study of language. For more on competence/performance from a corpus linguistic perspective, see Tognini Bonelli (2001: 68ff.)

2.   Sources: Ex. 1: *The Guardian* Tuesday January 10, 2006, article about light boxes in the treatment of Seasonal Affective Disorder (SAD); Ex. 2: *The Guardian* Monday September 10, 2007, article about population management and the environment; Ex. 3: *The Financial Times* Monday July 12, 2004, article about the Bank of England ("The Old Lady of Threadneedle Street") auctioning off (hence "hammer") furniture due to office modernisation; Ex. 4: J. K. Rowling, 2003. *Harry Potter and the Order of the Phoenix*. London: Bloomsbury (p. 27), variants of "crying over spilt milk" and "the cat's among the pigeons" respectively; Ex. 5: C. Brookmyre, 2003. *The Sacred Art of Stealing*. London: Abacus (p. 44), variant of "[somebody/anybody] on the horizon"; Ex. 6: P. O'Brien, 1979. *The Fortune of War*. London: Collins (p. 49), pun based on "the lesser of two evils".

of a figurative expression in a context which simultaneously activates its literal meaning (Examples (1) and (2)).

Although semantically complex, these kinds of creative meaning-making are in no way unusual. They are the bread and butter of modern journalism, as a glance at the headlines and sports section of any British daily newspaper will confirm (Examples (1)–(3)); and literature is elevated above normal language because its creative manipulation of conventional forms creates aesthetically pleasing and meaningful effects (Examples (4)–(6)). Because of this, the apparently specialised topic of this investigation becomes relevant to a range of practitioners. Corpus stylisticians and literature scholars will glean a better understanding of how creative exploitations of conventional language generate enriched meanings, and will do so in ways which can be traced collocationally, replicated, and even reproduced. The same knowledge can be put into operation by creative writing specialists, but also by translators who face the task of reproducing the effects of an author's idiosyncrasies and creative manipulations in the target language. The chapters to follow shed some light on how language users fuse novel elements with conventional language patterns to generate those enriched meanings that make language a joy and a pleasure.

Returning one last time to the football analogy, conventional phrases provide an example of how words play a team game. Each word is an essential part of the phrase, and has its own form, position and function. Under normal circumstances, the conformation as a whole acts in accordance with the moves it has rehearsed, each element contributing the necessary input to ensure a successful result. Occasionally some minor modifications have to be made so that the conformation fits seamlessly with the preceding play; this is akin to minor grammatical modifications so that a phrase fits the syntax and context. But idioms and fixed phrases, like football teams, have scope for the introduction of substitutes to replace an under-performing player or to boost performance. Introducing untested substitutes to a game can have interesting consequences. These new members have to fit into the pre-established set-up, but not only do they replace the outgoing player, they further enhance their team's performance by contributing something of their personal style and strengths, and in so doing bring new energy to a flagging game. The words that substitute the regular lexical members of a phrase too make an additional contribution to its overall meaning, not merely preserving the original meaning but also adding something extra to it. That extra meaning is something of an unknown quantity, and it is one of the major focus points of this study to find out what meanings appear to come into play, how their presence can be ascertained, and what they appear to contribute to the overall meaning expressed.

1.3   Words, collocations, phrases

What makes a sequence of words qualify as a collocation, and when does a collocation meet the requirements for classification as a phrase? What is a phrase, anyway? These questions do not have entirely straightforward answers, and depend heavily on the kinds of preconceptions about the relationship between words and meaning that were outlined in §1.2. Phrases come in so many guises that it is unlikely that any single definition could ever account for them all. To make matters worse, there is no single definition of "collocation" either (see §3.3).

In extremely reductive terms, although collocations are made up of more than one word, they only express one meaning. In most definitions, this meaning must be semantically complete, i.e. express a single thought or concept. This is known technically as its "semantic integrity" (Cruse 1986). When two (or more) collocates co-occur – not necessarily side by side or even in a particular sequence – this meaning is the only meaning that is understood. In other words, the various meanings that each of those collocates might express in other contexts, with different collocates, simply fail to materialise. But collocations, although semantically complete, are not as semantically complete as phrases, which can be said to express a *complete* thought or concept rather than just a *single* one. How it does so, and what a complete thought is as opposed to a single one, will be treated at some length in the course of this book. For now, suffice it to say that since phrases can easily be used as stand-alone items – in newspaper headlines, as punch-lines to jokes, as sound-bites, and as final comments to wrap up a discussion or even an academic paper – and collocations on the whole cannot, it can be inferred that phrases express more complete meanings than collocations do.

Tied up with the greater meaningfulness of phrases is their form. Not all linguists agree with the axiom that form and meaning are inseparable, but when it comes to phrases, it is difficult to find a dissenting voice on this matter. The vast majority of collocations can withstand variation to word form (inflections of the lemma) and positioning, while by definition the sequencing of a phrase must be fixed, or at least semi-fixed; and the meaning of the phrase relies not only on the right words occurring in the right sequence, but requires that those be particular word *forms*. Even a change from singular noun to plural, or a change of tense or aspect can be enough to prevent a phrasal meaning from being activated. Simply put, collocations have more formal flexibility than phrases do, and this translates semantically into a less specific or less complete meaning.

The comparative freedom of a collocation compared to a phrase incorporating that same collocation can be appreciated by a glance at how apparently similar forms collocate. Tables 1.1 and 1.2 show the collocation profiles for *black and*

Table 1.1 Collocation profile for *black and white*

L3	L2	L1	Centre	R1	R2	R3	R4	R5
the	a	in	black	and	white	and	of	the
of	the	the				photographs	and	a
and	of	a				the	the	and
a	and	of				in	in	of
in	in	and				photograph	it	in
is	with	between				film	on	to
to	colour	small				"	are	is
with	on	both				images	colour	was
it	is	with				"	is	with
be	to	old				prints	which	are
by	it	"				television	with	which
was	was	to				pied	that	on
are	be	or				one	a	at
were	down	as				I	I	I
one	his	his				striped	but	were
there	her	on				stripes	to	his
than	there	its				to	you	that
"	–	that				but	was	no
from	by	little				cat	they	by
I	an	distinctive				or	there	not
colour	all	simple				with	"	she
it's	that	striking				people	or	colour
only	its	for				it	by	as
into	black	red				minstrel	from	you
for	"	only				photos	were	from
up	as	large				on	at	for
red	white	is				illustrations	for	also
she	these	by				pictures	"	an
all	like	all				–	we	have
have	for	were				marble	so	had
	at	not				tiled	show	do
	it's	from				was	have	he
	out	their				of	had	all
	some					photo	yeah	it
	are					picture	will	can
	or					photography	who	"
						as	as	used
						that	if	like
						films	dairy	
						there	one	
						print	black	

Table 1.1 (*continued*)

L3	L2	L1	Centre R1 R2	R3	R4	R5
				so	not	
				is		
				minstrels		
				pattern		
				work		
				for		
				plumage		
				plates		

Table 1.2 Collocation profile for *in black and white*

L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
the	a	the	to	down	in	black	and	white	and	the
of	the		it	it					the	
	it			there					"	
	and			be					in	

white and *in black and white* respectively.[3] The profiles not only list which collocations occur (in descending frequency) but also indicate where these occur relative to the search term (L = left, or before the node; R = right, or after the node). The range of collocates for the collocation *black and white* (see especially R3) is strikingly more varied than that for *in black and white*. In generating Table 1.2, all that happened was that the three-word sequence *black and white* was expanded into a four-word sequence by adding the most frequent collocate to the left (L1). Such a dramatic difference, emerging from the addition of just one word (*in*, which is not even a content word) illustrates nicely how meanings become increasingly specific as the length of a collocation/phrase increases. The extreme point would be when the phrase appears not to collocate consistently with anything, at which stage it can be said to be meaningfully complete (cf. units of meaning, §3.2). When it is meaningfully complete, it may also be available for use as a stand-alone item; and when functioning as a stand-alone item, it does not attract regular or reliable collocates, except perhaps punctuation.

In theoretical linguistics, words, collocations and phrases are often treated as if they each obeyed their own set of rules. Yet it is obvious that words have a relation to collocations and that collocations have a relation to phrases. Perhaps less obvious, but no less true, is that there are also many points of similarity

---

3. Data: British National Corpus processed using WordSmith Tools 4.0 (Scott 1998).

between discrete words and phrases. The point to note is that lexis follows not *rules* as such, but *tendencies*. So words tend to combine into collocations in order to express more specific meanings, and collocations tend to grow into phrases when even greater specificity is required; but this does not entail that words can never be used singly, for their "intrinsic" value (see §1.4): they just tend not to.

One place where words are used for their intrinsic value is in creative exploitations of conventional phrases. Creativity of this sort is an exceptionally interesting aspect of language use, because it seems to fuse the two dominant views of meaning that were outlined in §1.2: the mental lexicon or "words have meanings" view, and the empty lexicon or "meanings have words" view. These outlooks on how lexis works are largely incompatible and mark a fundamental split within the linguistics community. One either does or does not accept that there is such a thing as the mental lexicon; one either does or does not accept that words only have meaning when in use. Yet in order to do justice to creative language, the fence has to be straddled, at least for a while, to make sense of how the extremes interact.

If any group of linguists believes in the mental lexicon, it is psycholinguists. Unlike other "words have meanings" proponents, including language philosophers and cognitive linguists, psycholinguists seek empirical experimental evidence for the claims they put forward. Their findings are thus observable, replicable, and derived from the study of others' interpretations of language, not just those of the researcher. Research conducted by psycholinguists has interesting parallels with that conducted by corpus linguists, who seek empirical evidence for the "meanings have words" claims they put forward. The findings from psycholinguistic investigations are also observable and replicable, and are derived from the study of texts written by others for the consumption of others.

Because a great deal of psycholinguistic research has been done on the interpretation of phrasal meaning, the work presented in this volume makes reference to the main psycholinguistic schools of thought regarding idioms as well as to other corpus linguistics studies in the area. Combining experimental research findings and corpus-derived ones adds an important dimension to both fields. In psycholinguistics experiments, phrases are presented (usually) in pairs – one version deemed to trigger a phrasal interpretation, the other a fully compositional one. In corpus linguistics, the single instance is subordinate to the typical patternings in which it and other instances occur. Discoveries relating to interpretation of idioms can be juxtaposed with observations regarding their typical patterns of use to enhance the existing body of knowledge, to mutual benefit. Interpretations which seem to find no corroboration in corpus data, on the other hand, indicate where shared linguistic knowledge ends, and the representation of language in the mind begins.

1.4   Meaning in the mind

This study acknowledges that some meaning is unobservable and exists in the minds of language users – a departure from most corpus linguistics studies, which would abhor such an admission (Hoey 2005 is a noble exception). This is necessary because two aspects of so-called secondary meaning – connotation and semantic association – come into play in creative uses of language. Connotations are occasional, associative and individual meanings held in the memory (Bussmann 1996:96) and typically pegged to single words, e.g. *red* can connote blood, communism, fire, Santa Claus, and so on. Semantic association (Hoey 2005) is an abstraction of collocation in which abstract semantic features shared by a word's collocates, such as negativity/positivity, are transferred onto or associated with the word.[4]

    In language play, the conventional patterns of usage do not tell the whole story: secondary meanings are voluntarily evoked. Since secondary meanings are only occasionally present, something special must be happening when they do appear. That particular something is rarely addressed in studies which investigate linguistic creativity, where the spotlight is invariably directed straight at the variation with scant regard for the underlying form. Admittedly, conventional phraseology is something of a let-down in comparison with creative exploitation (consider Examples (4)–(6), above) because it is just a platform for something more interesting and appealing to take flight from. But if only the novel variants are investigated it is impossible to understand how the novelty actually works with respect to more normal uses. The view taken here is that unusual language is actually very similar to normal language, but the differences lie in use. This means that they must be studied in context, and they must be studied comparatively, using data from similar sources aimed at similar audiences. And this is where

---

4.   These concepts will be explained in the detail they deserve in §3.5–3.6, though a few words need to be said immediately about the decision to use the term *semantic association* for features that *semantic prosody* is often used to cover. Semantic prosody as originally presented comprises two aspects – a semantic component and a functional component (Sinclair 1996b). For many years the pragmatic component has been largely overlooked, with the more easily-identifiable semantic traits being highlighted (e.g. Partington 2004; Morley & Partington 2009). These semantic traits are typically inferred from the collocates and are essentially an abstraction of the semantic preference. Referring to this with the term "semantic association" (after Hoey 2005) makes it possible to separate out the abstract semantic features of a word's collocates (which may have evaluative force, yet do not express the pragmatic, functional meaning of the unit of meaning), leaving "semantic prosody" to refer to the functional meaning which 'ranges over combinations of words in an utterance rather than being attached to just one' (Sinclair 2003:117).

corpora come in. Later chapters will present the findings of corpus analysis and explain, amongst other things: (i) how it is possible to recognise a phrase even if it has been altered; (ii) whether or not the unaltered phrase would be meaningful and appropriate in the same context as its variant; (iii) what happens in the phrase and/or its context to trigger a meaning that is not normally there; and (iv) how the conventional and unconventional aspects combine in "optimal innovation" (Giora 2003) i.e. the creation of meaningful, appropriate and aesthetically pleasing language.

1.5   Meanings in use

Up to this point, for simplicity, it has been suggested that phrases either appear in their conventional form or as creative modifications of it. This is certainly the assumption made when an exploitation is deciphered as "phrase + novel element". The reality, uncovered through corpus-based work into phraseology, is that most phrases can undergo a degree of variation before any novelty actually stands out. These degrees of unmarked variation are usually overlooked, but in the discussion to follow, the reader will encounter the full range of variation encompassing forms which can barely be classed as variants at all, to those which are undeniably novel. All are present in the same data. Finding variants in a corpus needs no particular tools other than patience and a certain amount of lateral thinking, and will be outlined in due course (§6.2). But this is not the only way in which corpora are used in unusual ways in this study. A focus on variation entails a departure from standard corpus analytic techniques which prioritises regularity and downplays variation. Some might wonder why a corpus is necessary at all.

    The reason why corpus data is used, and why variation is included alongside standard wordings, is to find out how phrases fit into the wider textual environment. Examining phrases in context is expected to shed light on how context sets up favourable conditions for a potential meaning to be actualised, which has two areas of direct application in this study. One of these is to test the various hypotheses put forward by psycholinguists on how phrasal meanings are understood, using naturally-occurring samples of language rather than examples constructed for experimental purposes. Experimental texts are intended to favour particular meanings very clearly and explicitly, while real-world use can be much fuzzier. It is therefore interesting to compare the contextual features of both kinds of texts to see what similarities arise, and why particular contextual features might trigger particular meanings. The second application is to include context in a comparison of normal and variant forms of phrases. This makes it possible not only to discuss overtly creative forms such as those seen in Examples (1)–(6), but also to

explain why some apparently normal realisations are perceived as unusual, as in Examples (7)–(9).

(7)   Actor John Inman is **in the pink** as he and co-star Fleur Bennett begin work on a new series of Grace and Favour.

(8)   Tim has a collection of space age memorabilia which would make most extra-terrestrials **green with envy** – if they weren't that colour already.

(9)   A customer **saw red** when her bank said she was overdrawn – by £40 million.[5]

Unusual uses of language are a feature of all creative texts, but they are normally judged in comparison with some imaginary idealised norm. This kind of comparison is not unsuccessful, but it lacks depth because it can only analyse the variant in terms of superficial differences. Analysing phrases in context with corpus linguistics techniques provides access to the deeper layers of meaning which are easily missed, most importantly the semantic preferences, semantic associations, and semantic prosody. These are all abstract features of phrasal meaning, and as such can only be identified by looking at large numbers of examples, as in a corpus. It will be reiterated throughout the chapters of this book that these abstract elements of meaning are the ones which are always present, in spite of variation, anchoring the phrase to its complete, functional-pragmatic meaning.

Creativity is individualistic, but not too much so. If there is not enough common ground shared between writer and reader, speaker and hearer, communication simply cannot succeed. The line of argument that runs through this book is that a phrase can withstand all sorts of lexical and syntactic variation on the condition that it is not used to express a different pragmatic function. Varying the most visible of cotextual elements – the collocates and colligates – can indeed change the force and focus of the meaning, but does not change the underlying message, and novelty is merely an optional extra. It is the cherry on the cake, not the cake itself.

1.6   Navigating variation

The relationship between meanings and words is a complex one. The chapters of this book seek to introduce this relationship and exemplify it with colour-word

---

5.   Source: Bank of English 450m version. Ex. 7: British Magazines, *pink* also refers to Inman's "camp" persona; Ex. 8: Oznews, Martians are stereotypically *green*-skinned; Ex. 9: Today, being overdrawn is being in the *red*.

idioms and their variants, illustrating progressively more elaborate examples of language in use. Chapter 2 explains idiomaticity in general, while Chapter 3 outlines in detail the theoretical framework adopted in the study: the unit of meaning (Sinclair 1996b), extended and complemented by a discussion of connotation, metaphor, and other secondary meanings. The levels of analysis inherent in the unit of meaning – collocation and phraseology, syntactic and semantic relations, and secondary meanings – are dealt with in turn in Chapters 4–7, with a series of examples and case studies which illustrate how the various levels interact. The threads of the argument are brought together in Chapter 8, in which the creative exploitations provided in Examples (1)–(6) above (§1.2) are subjected to examination using the spectrum of tools introduced and utilised in the other chapters. These final worked examples demonstrate just how intricate meaning-making is, especially when read between the lines of a corpus.

Idioms and idiomaticity

This chapter introduces idioms and idiom-like features in phrases and collocations. There is no shortage of descriptions and classifications of idiom in the literature (see especially Makkai 1972; Fernando & Flavell 1982; Fernando 1996; Moon 1998), but most existing work is theoretical in nature and requires testing against the reality of idioms in use. The essential features that must be present in order to define a phrase as an idiom are outlined, and the notion of *collocational harmony* is introduced to explain how literal and non-literal meanings are perceived in text. On the basis of these features, the difference between idiomatic and non-idiomatic collocations is also explained. The chapter ends with a description and discussion of the data selected for inclusion in this study.

2.1 What is an idiom?

One of the frustrating things about defining language technically is that the most simple and accessible explanations often turn out to be misleading, if not virtually meaningless, on closer inspection. Idiom is a prime example. It is defined in contrast to "normal" vocabulary, normal here implicitly meaning language which is built up *compositionally*, i.e. alternating grammatical and lexical elements to express meanings and their relationships to one another. Considered in this light, idioms are easily defined: an idiom is a combination of two or more orthographic words whose meaning, taken together, 'cannot be predicted from the meanings of the constituent parts' (Kövecses & Szabó 1996: 326).[1] In other words, there is a discrepancy between what the phrase as a whole means and what its constituents would mean if read compositionally. This, however, is only one of the features of idiomaticity. For an idiom to expresses the meaning that it does, it must also be *institutionalised*. Institutionalisation of lexis means that there is a received meaning which is understood by all. Idioms are 'highly overlearned word sequences that comprehenders have experience with as holistic units' (Titone & Connine

---

1. "Word" here means "discrete word form" or "orthographic word"; some scholars admit smaller units, such as phonemes (Hockett 1958) or morphemes (Makkai 1972).

1999: 1655), and as such they feature in dictionaries and in speakers' personal vocabulary stores as if they were a kind of long word.

The meaning of an idiom cannot be predicted or built up by adding together the meanings of its components, but there is no need to do so because it is already an established item of vocabulary with an established meaning attached. However, if a phrase is not institutionalised, or is institutionalised but unfamiliar, the case is different: hearers and readers have to unpackage the meaning word by word to try and make sense of it, and each person can potentially infer a different meaning from the same unfamiliar idiomatic word string (Keysar & Bly 1999).

The twin features of institutionalisation and non-compositionality are not the exclusive preserve of idioms. One additional feature that is generally attributed to idioms, but not to other phrasal types, is that they have a homonymous literal counterpart, so *kick the bucket* can mean both *die* and *hit a bucket with your foot* while a non-idiomatic phrase such as *in the middle of* has only one reading.[2] A further claim made in all theoretical accounts of idioms (see especially Makkai 1972; Gläser 1988) is that idioms are *transformationally deficient*. Underlying this assumption is the view that normal lexis is very flexible and can tolerate any change that is compatible with the grammar. Idioms, in contrast, are considered to be far less flexible in this sense. Their meaning is tightly bound up with their wording, and so any changes to morphology or syntactic sequencing, as well as substitution of any of the constituents, will almost certainly cause the idiomatic meaning to revert to a compositional one.[3]

---

2. Those idioms which have both an idiomatic and a literal meaning, e.g. *kick the bucket* and *skating on thin ice*, are referred to in the psycholinguistics literature as "ditropically ambiguous". Not all idioms are ditropically ambiguous, for example *by and large* and *hue and cry* have no literal counterpart.

3. Gläser (1988: 268–269) provides one of the most comprehensive yet succinct discussions of transformation tests which are applied to idioms. True idioms are believed to fail all the tests, i.e. the idiomatic meaning cannot withstand the transformations, while the meanings of other phrases can. The tests can be grouped into two areas: lexical transformation tests and grammatical transformation tests. The main lexical tests are

i.   augmentation (addition of lexical constituents)
ii.   elimination (deletion of constituents)
iii. substitution (replacing a constituent by a semantically-related word), and
iv.   permutation (rearranging constituents whose order is fixed)

The main grammatical tests are

i.   blocking of predication
ii.   blocking of the formation of comparative and superlative forms of adjectives,
iii. blocking of nominalization, and
iv.   blocking of passivisation

It is not only idioms whose meaning cannot be predicted from their constituents: the same is also true of collocations, as will be seen in §2.2. This is a problem for formal description which considers it normal for discrete meanings to be associated with discrete words, and not normal, or at least much less normal, for them to be associated with longer sequences. Viewed from this "words have meanings" perspective, phrases are palpably different from normal language. Yet from the "meanings have words" standpoint, where meanings are associated with words in context, it becomes much more difficult to draw a line between the meanings expressed by idioms, non-idiomatic phrases, collocations, and discrete orthographic words.

Fortunately, a number of features set out the parameters for idiomaticity, making it possible to differentiate idioms (e.g. *not the end of the world*) from lexical bundles (e.g. *at the end of*) or collocations (e.g. *world view*). These are compositionality/analysability, salience, semantic transparency, and adherence to truth conditions. The first three are phrase-internal, pertaining to the relationship between the discrete words within the idiom, while the fourth is identified by contrasting the idiom with its context. The four features function together, not in isolation, and represent clines or gradations of idiomaticity rather than discrete categories. "True" or "pure" idioms display all four features prominently, while others are more variable: some idioms are more idiomatic than others. For example, *flying a kite* or *having kittens* are both extremely idiomatic, strongly representing all four features, while *cherry picking* and *skating on thin ice* are moderately idiomatic, as the reader should be able to appreciate after reading the discussion in §2.1.1–2.1.4, where the features of idiomaticity are outlined.

2.1.1   Compositionality and analysability

It has already been mentioned briefly that idioms are non-compositional sequences. They are pre-constructed vocabulary items which are learned whole and used whole, and this means that they are not built up from the lexical and grammatical resources of the language each time they are used.

The same phenomenon can be viewed from the other direction. Being learned whole implies not only that idioms are not freshly constructed each time they are used, but also that they are not broken down (or analysed) into their component

---

There are valid objections to the blanket use of these criteria. Chafe, for example, argues that "deficiencies" are likely to be related to the underlying meaning of the idiom. Citing the case of *kick the bucket*, which cannot be passivised, he makes the valid point that the literal counterpart, *die*, cannot be passivised either: it is therefore unreasonable to expect an idiom to undergo transformations which its literal counterpart does not or indeed cannot (Chafe 1968: 122).

parts when used or learned. The analysability of an idiom refers to the ease with which it can be taken apart and its meaning understood from the meanings and syntactic roles of its constituents. Analysability has semantic and syntactic aspects: in this discussion the semantic analysability of idioms is treated in §2.1.3, while for now we will limit the discussion to syntactic aspects.

Syntactic analysability governs the flexibility of an idiom. All idioms have a *canonical form*,[4] which is how it is intuitively remembered and how it is recorded in dictionaries and lexicons. This non-compositional form is 'a lexical complex which is semantically simplex' (Cruse 1986: 37), and although convention dictates that these discrete words appear on the written page as such, *raining cats and dogs* might be better represented with a single, uninterrupted string: *rainingcatsanddogs*. However, the discrete words permit syntactic analysis to take place, which in turn makes some kinds of variation possible: *rain cats and dogs*, *rained cats and dogs*.

A well-formed idiom can appear in a range of variant realisations as well as in its canonical form. For example, verbal idioms such as *blaze a trail*, *blow the whistle on sb*, or *turn the tables on sb* can also be converted into a range of derived compound verbs, compound nouns and nominalisations, adjectives, adverbials or prepositional phrases (Gustawsson 2006): *trail-blazer*, *trail-blazing*, *to trail-blaze*; *whistle blowing*, *whistle blowers*, *to whistle-blow*; *the turning of the tables*, *table-turning* (ibid.: 173–174). These are essentially non-compositional because they exploit the wording and idea of the existing idiom, but they are more compositional than the canonical form is (it is a single lexical item, while the modification is not).

If an idiom is asyntactic (grammatically ill-formed) it cannot be modified in this way because the normal grammatical rules of transformation cannot be applied. For example, *by and large* is considered ill-formed (the preposition-conjunction-adjective sequence is agrammatical), and therefore resists change.[5] When the

---

4. This canonical form may vary in minor details from person to person, dictionary to dictionary. In particular, a dictionary may only include the core collocation at the heart of an idiom rather than the entire phrase, especially if there is variation permitted outside that collocation. For instance, *caught red-handed* is typically listed as *red-handed* (OED, COBUILD) because the verb can inflect regularly, even if *caught* is the most common form. *Raining cats and dogs*, on the other hand, is a canonical form, because even though the verb can inflect in principle, it tends not to.

5. As with many idioms, the form only seems agrammatical because the relevant meanings of the words have fallen out of use. *By and large* was a metaphorical idiom used in the navy. It means "broadly speaking" because 'sailing "by" means to steer a ship very close to the line of the wind, and sailing "large" means the wind is on the quarter. This technique made it easier for helmsmen to keep a ship on course during changing winds and in difficult conditions but not

syntactic roles of the constituents cannot be analysed, regular transformation is impossible, and the idiom in question is described as frozen. Frozen idioms are generally perceived as being more idiomatic than less fixed idioms are.

2.1.2   Salience

Most speakers of a language, if asked, can say what "the" meaning of a given word is. This meaning may well be one of dozens of possible meanings that are realised with that word, but there is something about "the" meaning that makes it the first one to spring to mind. That something is salience.

Understanding what salience is and what its implications are for language comprehension makes it possible to appreciate where the widely-held belief that "words have meanings" stems from. Belief is the key word here: salience is effectively a belief about what words mean, which can be contrasted with the evidence of what words mean when used communicatively in meaningful events (§4.2). Being a belief about meaning, salience is defined qualitatively and cannot be quantified empirically. So although the concept itself is easy enough to grasp, it must be remembered that salience is subjective and unfixed, because the salient meaning of a word, collocation, or idiom is the most dominant (prominent) one *for an individual* (Giora 2003: 40).

> To be salient, meanings of words, phrases, or sentences (e.g. the conventional interpretations of idioms or proverbs) have to be coded in the mental lexicon and, in addition, enjoy prominence due to their conventionality, frequency, familiarity, or prototypicality. Meanings not coded in the mental lexicon (e.g. conversational implicatures constructed on the fly) are nonsalient. Coded meanings that are less familiar or less frequent are less salient.   (Giora 2002: 490–491)

In §1.2, choosing meanings for words was likened to a children's game of lucky dip, where the first choice is typically the biggest package – children tend to unanimously equate bigger with better. However, imagine for a moment that some of the children know that there is a smaller package in the sack containing something small but valuable – money, the latest gadget, chocolate... In this case, those children are likely to be more selective and try to find the parcel that best fits their idea of what a parcel containing something valuable looks and feels like. Both of these strategies (bigger is better, and small but significant) are illustrations of how salience works.

---

in a particularly accurate way' (Jack 2004: 3). Thus *by* is an adjective, not a preposition; but as it is no longer recognised as such, the idiom is considered ill-formed.

How any particular meaning comes to be the dominant one for an individual at any particular point in time is impossible to state with any certainty, but frequency is assumed to play an important part, a view advanced in the "corpus to cognition principle" (Schmid 2000). Constant reinforcement of the connection between a word and a meaning (or a real-world object, event, or entity) is a major contributing factor to the strength of association that builds up between words and meanings. In the total absence of external cues, we use our memories of previous encounters with the word to provide it with an overall best-fit meaning which we consider to be both typical and conventional. The most conventional meaning is generally considered to be the word's basic meaning.[6] Crucially, that basic meaning will require no support from other words in the context for its realisation: it is essentially a 'context-free literal meaning' (Lakoff 1986). There is a perennial problem with context-free meaning, however, in that nothing in human experience is truly context-free. Words are learned in context, even if that is a point-and-name context; and significant experience in any given context makes one meaning more prominent than it might otherwise be. The most frequent meaning of *note* in British English is 'short letter'. However for a musician, musical sounds and/or their representation on a stave are far more frequent, both as events and as meanings, than short letters are, making the musical meaning of *note* more prominent and almost certainly more salient for these individuals.[7]

Salient meaning is sometimes confused with literal meaning, but it is not the same thing at all. In fact idioms, irrespectively of how easily their meaning can be interpreted, are salient if familiar, non-salient if unfamiliar (Schweigert et al. 2003). What this means in practice is this: an idiom is composed of distinct word forms, each of which has its own salient meanings; but the idiom too has a salient meaning as a phrase. The meanings of the distinct words are *no more salient* than the meaning of the phrase, effectively making an idiom (once learned) no different semantically to a discrete word (once learned). Importantly, and because they are salient in their own right, idioms are not subservient to the presumed greater

---

6. This judgement reflects a folk-linguistic view of "literal" meaning (see especially Gibbs et al. 1993) in which the literal merges with the prototypical and the frequent. The end result may not be literal at all in linguistic terms: many words which are considered to be literal by the populace at large are in fact metaphorical in the strict sense (see Steen 2002; Pragglejaz Group 2007).

7. Of course, there is a great deal of individual variation in perceptions of salience, making it dangerous to over-generalise or take for granted that what is salient for one will be equally salient for another. It is therefore unfortunate that arguments relying on salience judgements (see, for example, Giora 2003: 43, 54) rarely specify which meaning the authors believe is (or should be) the salient one; and how salience is supposed to be determined or predetermined at all is circumvented in the psycholinguistics and cognitive linguistics literature.

salience of their component words. If anything, the reverse is true: when the idiom belongs to a ditropically ambiguous pair, i.e. when a single word string can be interpreted either idiomatically or by semantic-syntactic analysis, e.g. *break the ice*, *kick the bucket*, it has been found that the idiomatic meaning is processed (understood) in tandem with, and sometimes faster than, its compositional counterpart (Ortony et al. 1978). Additionally, some psycholinguists affirm that the idiomatic meaning is *more* salient, and *more readily associated* with the institutionalised phrase than a non-idiomatic (compositional) meaning is (Gibbs & O'Brien 1990). This is supported by corpus linguistics research into lexical priming (Hoey 2005), which hypothesises that one interpretation is always preferred over others, not merely because it is contextually more appropriate, but because it is more familiar and salient in absolute terms (ibid.: 82ff.).[8] In an unfamiliar idiom, however, the constituent words have individual salience, while the idiomatic meaning is non-salient. As it is not stored in the memory, the only way its meaning can be understood is by inference, using context to help (Peleg et al. 2001).

Even the presence of highly polysemous words within the unfamiliar idiom offers no guarantee of its correct interpretation (Keysar & Bly 1995). It is not enough for an idiom to be institutionalised in its non-compositional form: it must also be familiar for its phrasal meaning to be conveyed. Salience is not institutionalised, nor is it institutionalisable because it is determined by familiarity and, within this, by the degree of importance that each single person assigns, consciously or unconsciously, to a particular meaning.

2.1.3   Semantic transparency

Having discussed salience at length, we now come back to the phrasal meaning of idioms. This meaning is judged on a transparency/ opacity scale, with opaque idioms being the most obviously idiomatic. If an idiom's meaning is relatively clear from a compositional reading of its component words, it is said to be transparent: you can see though the words and arrive at the meaning, e.g. *wave a red flag before a bull*, "provoke", *take a hammer to sth*, "break up, destroy, attack". In opaque idioms, the words obscure the meaning, either because they are not salient (e.g. *by and large*, "broadly speaking"), because it is difficult to appreciate how they can be connected with the meaning (e.g. *kick the bucket*, "die"), or because they do not seem to make sense (e.g. *go cold turkey*, "stop taking an addictive substance").

---

8. Hoey's argument is centred on polysemous words as discrete units, but the principle applies equally to ditropically ambiguous idioms. If an idiomatic meaning is salient, it is reasonable to suppose that the string will not normally be used compositionally, and if it is, both meanings will be accessed at once (see Giora 2003: 11). Humour is likely to ensue, whether intended or not.

The perception of semantic transparency or opacity is not an absolute one nor is it entirely objective. Like saliency judgements, it is especially connected with familiarity. A transparent idiom is one whose meaning is familiar and/or whose component words have familiar meanings. This familiarity facilitates and speeds up the search for meaning. An opaque idiom, as we have already seen, lacks this aspect, but some psycholinguistics have found that even the most opaque of idioms 'may be perfectly transparent once one has learned their meaning' (Keysar & Bly 1999: 1572). When the meaning of an opaque idiom is unknown, the reader or listener takes some time to arrive at a plausible interpretation because it is necessary to make metaphorical connections between the words, the context, and the message that seems to be being conveyed. If, however, the component words themselves are unfamiliar, no metaphorical links can be constructed and the interpretation must be based entirely on contextual cues.

2.1.4   Truth conditions

The idiom *wave a red flag before a bull* was used to illustrate a transparent metaphor in §2.1.4. This is the American English alternant form of *(be) like a red rag to a bull*, which is examined below (§5.6). *To wave a red flag before a bull* nicely illustrates the difference between literal and idiomatic readings when these are determined by context. If a person is acting in a provocative manner, s/he can be said to be *waving a red flag before a bull*, or *like a red rag to a bull*, even when there is no bull in the vicinity, and no red fabric of any description is being waved at all. The same expressions can, in theory at least, be used to talk about a person – archetypically a Spanish matador – who is indeed waving his red cloak at a bull, inciting it to attack him. The link between the phrase and its idiomatic meaning is perfectly transparent, but what makes the former meaning idiomatic (or metaphorical) rather than compositional (or literal) is that it is untruthful.

It is important to stress that truth conditions are not determined by the internal composition of an idiom but by context. The truth condition requirement is the only one of the established criteria for judging idiomaticity to acknowledge that meaning is the product of the item (phrase) in an environment (situational or textual context). In doing so, truth conditions mark a boundary line between literal and figurative interpretations of idioms, just as they do for literal and figurative interpretations of single words and collocations (see §3.6.4). Figurative meanings are activated when there is a mismatch between the contextual situation and the truthfulness or relevance of the lexical item in question, be it an idiom, a metaphor, irony, or any other form of non-literal language. In other words, the literal reading of *wave a red flag before a bull* is untruthful (there is no

bull present, and no waving of anything), but since this "lie" is not functioning as lies normally do (the function of a lie is to deliberately mislead, which is not the case here) it is understood not as an untruthful statement but rather as a figure of speech. The greater the contrast between the context and the surface meaning of the idiom, the more metaphorical (and therefore idiomatic) it is.

It is their untruthfulness that ultimately differentiates idioms from other phrasal sequences (such as *from bad to worse*), just as it is untruthfulness which signals that a word is being used metaphorically, not literally. Yet "literal" is a term fraught with difficulties. It is subjective and even within linguistics it has at least five definitions: 'conventional meaning, subject-matter meaning, nonmetaphorical meaning, truth-conditional meaning, and context-free meaning' (Gibbs et al. 1993: 387). It can, however, be helpful to consider it as a phenomenon of collocation. A word (or collocation or phrase) is interpreted literally when it is in *collocational harmony* with its salient context. Previous encounters with a word fix its salient meaning(s) (§2.1.2), so when the word appears in such a context, the expectation of its meaning and the meaning it conveys are consonant. When on the other hand the word appears in a different context, i.e. one whose collocates do not support the salient meaning, a clash is perceived between the salient meaning and the meaning that is set up in context. This will be explored further in §3.6, but for now the important point to note is that adherence to truth conditions as a feature of idiomaticity is a collocational phenomenon. It is not inherent in the phrase's meaning but determined and signalled by context.

2.1.5   Dimensions of idiomaticity

The features of idioms described above are not the only features that have been mentioned in the literature, but they are the ones that are most consistently mentioned in both theoretical and descriptive accounts (see also Svensson 2008). Those which are semantic in nature (salience and truthfulness) are the most difficult to pin down, because of the subjectivity that lies at the heart of meaning judgements. Although easily understood as folk-linguistic categories, "literal" and "figurative" meanings are very difficult to separate out from one another. This is mainly due to the fact that they are treated as if they were inherent properties of words (discrete word forms), yet we have just seen (§2.1.4) that they are contextually determined. Corpus linguistics copes with this fact rather well because its methodology is based on Firth's (1957) contextual theory of meaning; theoretical, philosophical, cognitive and psycholinguistics fare rather worse, because they consider context to be of secondary importance. In downplaying the role of context, and magnifying the properties of decontextualised words,

they view idioms and idiomaticity only in terms of their formal make-up and "anomalous" behaviour instead of investigating how they mean what they do and, indeed, the full range of meaning that they express. These aspects of meaning can be accounted for within Sinclair's idiom principle (1991), as articulated within extended units of meaning (Sinclair 1996b), as will be revealed in Chapter 3 and throughout the subsequent discussion. For the remainder of this chapter, however, we will compare and contrast idioms and collocations on the basis of the features of idiomaticity outlined in this section.

2.2   Idioms and collocations

Many of the features of idioms that were described in the previous section seem to be equally applicable to collocations. Collocations are also recurrent combinations of words, making them institutionalised lexical items or, at the very least, institutional word combinations, just like idioms are. Some collocations are completely non-compositional (though perfectly analysable). This is the case with binominals and compounds (*salt and pepper*, *blood red*). Others are semi-compositional, including most verb-noun collocates, which can tolerate a high degree of morphological and syntactic position change, and "long distance collocations" (Siepmann 2005) which are spaced out over longer stretches of text (e.g. *not only … but [also]*). The salient meaning of a collocation also differs from the sum of the salient meanings of its constituents, even though this difference is often masked by familiarity. For example, *contact lenses* do not resemble other types of lenses used in optics due to their size, material and flexibility; and contact is made between the lens and the eyeball rather than between visual perception and the object viewed. Collocations may be transparent or opaque, and here again familiarity interferes: *fish and chips* does not refer to any fish accompanied by chips, but to a certain kind of fish (typically cod, haddock or plaice) which has been dipped in batter and deep-fried; and it is served with thick-cut potato chips, not with French fries, and typically bought as street-food from a fish and chip shop (cf. Cruse 1986: 39–40). Finally, collocations may express untruths if compared to a literal reading: where is the height in *highly dangerous*, the having in *have a look*, the ink or vehicle in a *pen drive*?

So where do the differences lie? Leaving aside for a moment those collocations which are idiomatic, it should already be apparent that collocations occupy the space somewhere between discrete word forms and idioms. The most obvious difference between an idiom and a collocation can be found in the relationship each has with the salient meanings of its constituents. An idiom expresses an idea which cannot normally be inferred from the meanings of its constituents, while

a collocation expresses an idea which *can* be inferred to some extent without the contribution of contextual cues. Word co-occurrence in collocation causes the component words' salient meanings to be *restricted* or *specified* in some way, but not changed as such. *Sunglasses* are spectacles that shield the eyes from the sun, not beakers or spectacles that look like or depict the sun, or spectacles which you can use to look at the sun with, for instance. Yet while the salient meanings of *sun* and *glasses* indisputably contribute to the meaning of *sunglasses*, the same cannot be said for the salient meanings of *white* and *elephant* contributing to *white elephant*, "useless and costly object". When the meaning of a collocation cannot be inferred from the salient meanings, this is a sure sign that it is idiomatic. Similarly, if the collocation looks transparent but its analysed, salience-derived meaning turns out to be incongruous in context (Examples (1)–(3)), it is metaphorical, hence idiomatic.[9]

(1)   It is likely that travelling by plane is a **red herring**.

(2)   Too fat to adopt a baby 95 kg Liz fights **red tape**.

(3)   A certain sort of handsome young man; charming **black sheep** who are irresistible to elderly women.

Metaphorical collocations are like idioms in that once learned, the salience of the collocation is at least as high as the salience of the collocates when viewed independently. This seems to make them attract kinds of exploitation which "normal" collocations do not do, or do not do to nearly the same extent. This is partly because metaphorical collocations can be part of idiomatic expressions: they often occur as the invariable or minimally variable cores around which particular phrasal configurations build up (e.g. *black sheep*, *blue moon*, *green light*). Being central to the idiom, these collocations act as focal points for the salient phrasal meaning and are for this reason able to convey that meaning in ellipsed form – without the aid of the complementary collocates which complete the phraseology associated with the canonical form. Chapter 4 investigates how such collocations function as multi-word nodes within units of meaning (Sinclair 1996b). In particular, it looks at the role of phraseological context in fixing meaning, and how the impact of salient word-meanings is diminished within phrases (delexicalisation).

Another kind of exploitation found to occur with metaphorical collocations is that, like idioms, their contextual environments can be tweaked so that the salient meanings of the individual collocates are reactivated. The original meaning is not lost but is instead supplemented and enriched. How this happens, and how it is possible to state what additional meanings are activated, is treated in

---

9.   Sources: Bank of English 450m version. Ex. 1: Guardian; Ex. 2: Oznews; Ex. 3: Guardian.

some detail in Chapter 5. Finally, in the data investigated, it was found that the underlying conceits associated with metaphorical collocations and idioms can themselves be exploited metaphorically. Different aspects of this are treated in Chapters 5, 6 and 7, where a comprehensive explanation of how metaphor can be used to extend and generate meanings is provided.

2.3   A search for meaning(s) in idioms and metaphorical collocations

The title of this book, *Colouring Meaning*, is itself an example of a double meaning, as this is an investigation into how meanings can be modified to introduce additional nuances and associations, and does so by focusing on colour words in particular. The choice to study colour words rather than any other lexical set, or idioms in general, stemmed from curiosity regarding their connotative meanings and how they are transmitted through language.

Colours are imbued with a bewildering range of connotative and associative meanings, and these are reflected in language in two principal ways: in fixed and semi-fixed phrases such as idioms and metaphorical collocations, and in language play. However, conventional expressions which on the surface seem to embody and perpetuate particular associations do not always convey those meanings. Given the preceding discussion, it is clear that if the surface wording does not reflect the meaning of the whole, idiomaticity is at work. But this does not explain how the same wordings sometimes *do* express the connotation. This study goes some way in explaining why that happens.

2.3.1   Idioms: Word meaning and beyond

Idioms and metaphorical collocations which include a colour word in their surface realisation, e.g. *green with envy*, *red in tooth and claw*, *whiter than white*, provide evidence for the existence of particular associative meanings. For example, *green* is associated with jealousy, as demonstrated by *green with envy* and *the green-eyed monster*; *red* is associated with blood, as in *caught red-handed* and *red in tooth and claw*; *white* is associated with purity, hence *white lie*, *white as the driven snow*, and *whiter than white*. Despite this explicit knowledge about the meanings of colours – and in spite of appearances – expressions such as these do not necessarily express the connotations that they appear to embody, or do not express them in the ways we might expect them to. *Whiter than white*, for instance, typically describes people who are corrupt in some way, making it primarily ironic (Philip

2006), and *caught red-handed* does not seem to be used where blood would normally be present (ibid.; see also §4.3).

Colour-word idioms feature in the popular imagination as proof of the existence of connotative meanings. These associations typically relate to the colour's conventionalised metaphorical extensions (*red in tooth and claw* = blood, violence; *be green* = in harmony with nature) or symbolic value in the culture (*whiter than white* = purity; *red flag* = danger). In addition to these culture-bound associations, it is also true of colours that they refer metonymically to any entity which manifests itself in that colour (Niemeier 1998), thus *black* is related to black or dark things including the night and dirt, *white* to light, cleanliness, and purity, *red* to blood, fire and strong emotions (which cause the face to flush "red"), *blue* to the sky and the water, and so on (Philip 2006).

Although not all the symbolic values associated with colours find their way into conventional lexicalised forms (ibid.), many do, and in so doing the connotative meanings are reiterated through language. For this reason to *be green* still has "immature" as one of its conventional meanings, even though the connection between the green wood of a sapling and the malleability of the young mind has effectively been lost. The meaning is not a particularly salient one now, and the expression is relatively uncommon. If the phrase were to fall out of use, the association itself might be expected to die as a consequence.[10]

Although the surface wording of colour-word idioms suggests various symbolic and connotative meanings, these may not normally be present at all in the established meanings of the phrases they are said to have generated. This can be inferred by the fascination that the population at large has with etymology and the increasing number of books for popular consumption which reveal the lost meanings of idioms (notably Ayto & Crofton 2006; Cresswell 2007; Games 2006; Jack 2004; Wilton 2004). If the meanings were self-evident in the everyday meaning attributed to the phrase there would be little need for such etymological dénouement.

It is not only the amateur linguist who turns to etymology, however. The theoretical distinction between literal and non-literal meanings is essentially that the etymologically-prior one is literal, and all others are derived from it. This makes it

---

10. The "immaturity" sense of *green* occurred only 20 times (six of these in one text) in the 450m Bank of English, in contrast with several hundred occurrences of the nature-loving/environmentalist sense. Similarly, the association of *yellow* with cowardice is virtually obsolete: the non-zoological use of *yellow belly/bellied* occurred 14 times in the same data set, 10 occurrences of which required the additional presence of "coward" (or synonym) in the immediate cotext, e.g. "I'm such a coward! I'm such a yellow-bellied chicken! I'm so ashamed!" (US Books). *Yellow streak* occurs 19 times in the same data set, 17 of which as the metaphorical collocation.

tempting to believe that all idioms once had a literal meaning whose lexicalisation became frozen and, with the passing of time, started to be used metaphorically (see Wray 2008: 30–33). Unfortunately, there is very little historical data available to support this, with the result that etymological unpacking of idiomatic meaning is often a matter of conjecture. That said, in later chapters we will see that the literal meaning makes a contribution to the enriched meanings expressed through idiomatic language by feeding an "image schema" (Lakoff 1987b; see §5.4–5.6). To anticipate that discussion, consider briefly the case of *a bolt from the blue*, which means something like "unexpectedly".

   If the sky is black with thunder clouds and a lightning bolt is seen to shoot from it, this would be a dramatic but not particularly exceptional event. If, however, the weather is fine and the sky a perfect blue, a lightning bolt would come as something of a surprise. This is the "literal" meaning which is said to underlie, or "motivate" the expression. The problem is appreciating how this image translates into uses and contexts where neither lightning bolts nor the sky (whatever the weather) are relevant.

   The things that come as *a bolt from the blue* are mainly unexpected and unwished-for events including (sudden) death, divorce, political coups, arrest, or news of the same (Examples (4)–(7)).[11]

(4)   Not long after that, she had a massive stroke – literally **a bolt from the blue** – while she was travelling on the tube.

(5)   One of the most miserable cases that I have ever had to deal with in my constituency involved the death of a 14-year-old from glue sniffing. For his family, the fact that he had been experimenting with glue came as **a bolt from the blue**.

(6)   Family friend David Quinn said: "David rang and told her. She had no inkling." <p> Sandra's sister Catherine said: "It's a **bolt out of the blue.** Sandra and David had a rock-solid marriage. They were the perfect couple."

(7)   That same weekly, in its diary column, says that the office of British Prime Minister Mr John Major was forced to deny the persistent rumour that a Foreign Office spokesman had described the coup as "**a bolt from the blue**", since it begs the question, "What was British intelligence up to?"

While none of these things are like lightning bolts in appearance, the *effect* they have on the individuals concerned is analogous to the *effect* that a literal *bolt from the blue* would have, namely devastation (literally and figuratively). Thus there is

---

11.   Source for examples and synthesis of data presented: Bank of English 429 million version. Ex. 4: Guardian; Ex. 5: New Scientist; Ex. 6: Today; Ex. 7. BBC.

a suggestion that the connection between literal and metaphorical readings lies in their effect and function – what they *do* communicatively – rather than what they *say*. This connection between "origin" and function, and the implications that it has for idiomatic meaning is treated throughout the chapters which discuss idiom variation.

2.3.2   Colour words in collocations

Colours have symbolic meanings, but colour words, being mere words, can only express those (and other) meanings in collaboration with other words. For this reason, conventional collocations and idioms were selected for this study so that it would be possible to assess how, where and when their associative meanings are present in text. Before explaining how the data set was established, a brief note on colour words is necessary.

   Not all languages recognise the same number of colours. Berlin & Kay's (1969) groundbreaking study in this area found that although every language has at least two colour terms (*black* and *white*, or, more correctly, "light" and "dark"), and potentially hundreds, it is useful to distinguish between "basic" and "non-basic" colour terms. This study examines the former, but reference will be found throughout the data to the latter.

   Basic colour terms are superordinates, while non-basic colour terms are their hyponyms. So non-basic colour terms, which are specific both in terms of the precise hue they refer to and also the domain or register in which they are used, can ultimately be considered as optional subdivisions within the superordinate category. For example, *green* is basic, but *emerald*, *lime* and *chartreuse* are not; *red* is basic, but *crimson*, *vermilion*, and *burgundy* are not; *blue* is basic, but *light blue*, *turquoise* and *navy blue* are not. Two points need to be made here. The first is that it is not so much the hue as the word that is of importance here, because a speaker of any language can distinguish between differently-coloured plastic chips even if s/he refers to them with the same word. An English speaker will use the word *blue* to refer to the spectrum that Italian and Russian speakers would split into two distinct basic colours (*azzurro/blu* for Italian, *goluboj/sinij* for Russian), perhaps specifying *light* and *dark blue*, but still considering them the "same" colour, unlike *red* and *pink* which are separate. The second point is that different languages have different numbers of basic colour terms. English has eleven basic colour terms, which, according to Berlin & Kay (ibid.: 3), is the largest number of basic colour terms that any language has. These are *black*, *white*, *red*, *green*, *yellow*, *blue*, *brown*,

purple, pink, orange and *grey*.[12] More recent studies have ascertained that some of the languages which did not feature in the 1969 study have more than eleven: Italian, Russian and Turkish certainly have twelve, possibly more (Uusküla 2006, 2007; Özgen & Davies 1998; Davies & Corbett 1994).

The basic colour terms in English were taken as the basis for identifying a comprehensive set of metaphorical collocations and idioms to study. These were identified from a range of print sources rather than directly from the corpus, because it was essential that their meanings were established and familiar to the language community at large. An investigation of how meanings are exploited for stylistic purposes can only be carried out if the basic meaning is already established, and inclusion in a dictionary is a good indicator of its status in the language as a whole.[13] The print sources consulted were the *Oxford English Dictionary* (henceforth *OED*), *Collins COBUILD English Dictionary for Advanced Learners* (henceforth *COBUILD*), *COBUILD Dictionary of Idioms*, *Dictionary of Idioms and their Origins*, *Thesaurus of Traditional English Metaphors*, and *Wordsworth Dictionary of Idioms*. The same procedure was followed for the initial Italian data set, only a small part of which is included in later chapters, using the dictionaries *Lo Zingarelli 2002* and *Paravia Dizionario della lingua italiana* supplemented by the bilingual (English-Italian) *Ragazzini*.

Initially, all the collocations and phrases listed under the basic colour term headwords were compiled. This list was then pruned to exclude denotative nomenclature in which the colour word had a specifying or disambiguating function (e.g. *greenfinch* and *red kite* are birds belonging to the finch and kite families respectively; *white blood cell/red blood cell* are the two types of blood cell, terminologically differentiated by their colour; *red wine* denoted the dark shades of wine obtained from *red grapes, and so on*). However, any of these expressions which were polysemous were retained, e.g. *grey cells* and *grey matter* which can refer to intelligence in general as well as denote the composition of the brain; *red carpet* which can refer metonymically to important people.

---

12. This order of presentation reflects to the order in which colour terms are acquired in world languages. If a language only has two terms, they will be *black* (or "dark") and *white* (or "light"). The third term is always *red*; the fourth *blue*. *Yellow* and *green* tend to come next, in no particular order (and some languages only have one term covering both hues). *Brown* appears in the lexicon after both *yellow* and *green* are in place. *Purple*, *pink*, *orange* and *grey* do not come in any fixed sequence, but if one is present, all four are, thus completing the set of eleven.

13. There is still no consensus view on what precisely does and does not qualify as a "phrase" (see especially Gries 2008), so relying on published sources was felt to be the most reliable solution.

After this initial winnowing down, the list comprised a range of idioms (*scream blue murder*, *red rag to a bull*, *the pot calling the kettle black*), proverbs (*the grass is always greener*, *red sky at night*), conventional similes (*as red as a lobster*, *whiter than white*), and polysemous or metaphorical collocations (*black and white*, *grey matter*, *green fingers, white wedding*). Each of these was verified for currency by an initial search in the Bank of English corpus for English and the *Corpus di Italiano Scritto* (CorIS, Rossini Favretti 2000) for Italian. Those which did not appear were removed from the list, as were those which occurred less than five times.[14] In the end, the number of colour word expressions on the list was 126 (72 English expressions, 54 Italian expressions). These are listed and pegged to their symbolic or connotative meaning (after Philip 2006) in the Appendix.

Because variant forms of the expressions were also to be taken into consideration from the outset, it was decided to search for these within the corpora studied rather than look serendipitously for exploitations in other reading matter. Finding variants is now easier than it was, thanks especially to phraseological retrieval software such as *ConcGram* (Greaves 2009; see §6.1.1), but this cannot run on remotely-accessed corpora such as the Bank of English or CorIS. Instead, series of searches can be carried out on a trial and error basis, making the most of wildcards and variable gaps between the collocates in the query. First of all the search for variations was limited to see whether the colour words could be replaced by different colour words (e.g. *talk until you are blue/red/black* etc. *in the face*), but was later extended to include as many lexical and syntactic variations of the expressions as possible. The approach taken is outlined in Philip (2008), and details are provided in §6.2 to illustrate how variants of *the grass is always greener* were located both in a corpus and on the Internet. As well as variants to canonical forms of the expressions on the list, some phraseologically similar forms with no colour word were also studied; *caught red-handed / in flagrante delicto / in the act (of)*; *once in a blue moon / lifetime*; *passare una notte in bianco / insonne*; *into the red / debt*; *yellow / tabloid / gutter press*; *out of the blue / nowhere*. This made it possible to contrast highly salient vocabulary (colour words) with less salient vocabulary in similar structures.

---

14. Removal of absent/very low-frequency items was rarely necessary for English; more items were discarded from the Italian list because this smaller data set (80 million words) inevitably yielded poorer results. The eliminated expressions are listed in square brackets in the Appendix.

2.3.3   Colour words in language corpora

Corpus data is far from universally used in studies of figurative language and phraseology, mainly because these focus primarily on the formal and structural composition of phrases rather than tying phrases to their contexts of use. Cowie (1998a) identifies three major strands in phraseological research – the theoretical-formal stemming from Russian lexicology (see Cowie 1998b for an overview), the cultural-anthropological (discussed thoroughly by Piirainen 2008), and the corpus-based – the last of these being intimately linked to corpus-based lexicography for foreign language learning where usage information is an integral part of lexicographical description (Cowie 1998a: 3).

Using corpora to study the figurative uses of colour words in fixed expressions shifts the research emphasis away from etymology and the semantic analysis of the components, and onto meaning in use. This is not to say that the internal make-up of phrases is not taken into account in corpus-based studies, nor that the cultural values expressed through phrases are disregarded, but rather that the meaning of a phrase cannot be wholly accounted for without taking its contextual environment into consideration. Aspects of the better-established strands of phraseological research will enter into the discussion at various stages of this book, but will always be integrated into the "bigger picture" provided by corpus data.

What corpus data contributes to the study of figurative language is primarily the relationship of the item (phrase or collocate) in a textual environment. If 'meaning is function in context' (Halliday 1992: 16), then context cannot be considered an optional extra in determining the meaning of a phrase and/or its component words. Corpus data provides important information about the expressions' normal contexts of use, and this in turn makes it possible to build up a fuller account of what they are now used to express rather than what they have meant in the past.

Corpus data is also a rich and reliable source of variant forms: rich, because it reveals the extent to which idioms and other fixed expressions appear in modified form as well as the range of modification types which occur; and reliable in that these variants occur in the same types of texts as the canonical forms do, rather than being collected from different sources. Instances of word play (and this is one of the more prominent effects of variation) are often studied in isolation, and compared to their underlying canonical form, this too in isolation. The choice to use the variations present in corpus data means that the non-canonical forms are not only attested but can also be observed in relation to the wider context in which they have been used (see Gustawsson 2006; see also Moon 1996, 1998). The problem with collecting instances of variation from other sources is that these are all likely to be marked, and thus clear-cut cases of modification, otherwise they

would most probably have passed unnoticed. Using the spectrum of variation present in corpus data resets the balance between marked variation and less eye-catching non-canonical forms. In this way, variants can be compared with other variants as well as with the canonical base form, and the mechanisms underlying variability unravelled more successfully.

2.4   Into the wide blue yonder

This chapter has explored the main distinguishing features of idioms and idiomaticity, and sought to explain why some phrases and collocations are idiomatic while others are not on the basis of these features. It has also provided an outline of the data to be used in the bulk of the analyses to follow, explaining the choice to study colour words in particular, and why they are to be studied within idioms and other institutionalised collocations. Finally, it has put forward a case for corpus-based, contextualised studies of phraseology which will be reinforced in the chapters to come.

While the discussion in this chapter has made a good deal of mention of collocation, and some passing remarks about the idiom principle in language (Sinclair 1991) and extended units of meaning (Sinclair 1996b), it has not treated them in any depth. This gap is filled in Chapter 3, which focuses on collocation and its abstractions into phraseology. The extended unit of meaning will be seen to be a crucial notion in understanding the meaning of idiomatic expressions and, even more so, in explaining the parameters which govern variation.

Co(n)text and meaning

Chapter 2 outlined what idioms are and how they relate to collocations. This chapter deals with idiomaticity as a central feature of "normal" language, and does so within a Sinclairian framework. By way of background, the contrasting open choice and idiom principles (Sinclair 1991), mentioned in passing in the previous two chapters, are explained in more detail. Following from this is a much more thorough discussion of collocation, colligation, semantic preference and semantic prosody, which together make up the 'extended unit of meaning' (Sinclair 1996b). These four features – none being entirely unproblematic in terms of their descriptions – are supplemented with a fifth, borrowed from Hoey's (2005) work on lexical priming: semantic association. The addition of this fifth element, inserted between semantic preference and semantic prosody, is necessary in order to make a clear distinction between the pragmatic function of the unit of meaning (one aspect of semantic prosody) and the connotative, transferred, evaluative and other secondary meanings which can be inferred from the collocates (also commonly referred to as semantic prosody). Later chapters will demonstrate how important it is to differentiate between these related but fundamentally different aspects of meaning.

3.1   The idiom principle

There is something about natural-sounding language that makes it sound natural. That something is idiomaticity, intended here in the broad sense of "proper to and typical of a given language". True idioms are simply more proper to, and more typical of the language than other word combinations are. Linguistics is still very much dominated by a word-centric view of language, in which words have meaning and syntactic roles, and generally behave in an orderly fashion, words slotting neatly into the positions left open by the grammar. Idiomaticity (in both broad and narrow senses) does not quite conform to this model, and for a very long time indeed was simply ignored and often disparaged.

Corpus lexicography, most notably the COBUILD project, marked a sea change in this respect. The original Birmingham Corpus was about 20 million words of running text (Renouf 2007: 30), now considered a trifling size, but in the 1980s this was an unprecedented and unimaginably rich source of language data.

There was sufficient data available for it to be possible to identify patterns recurring in the neighbouring cotext for thousands of word forms, and this 'showed up the great influence which phraseology exerts over word meaning' (Clear 1993: 271). Patterns give rise to meaning: it does not come from words in isolation. This discovery turned traditional word semantics on its head. The distinct meanings expressed by any single word were seen to be entirely dependent on the word's co-occurrence with other words in regular lexicogrammatical configurations.

The idiom principle (Sinclair 1991) emerged out of this observation. In essence, it affirms that language tends to be phraseological, not compositional, and that as a result some word combinations are more privileged – more frequent, more meaningful, more significant – than others. The corollary of this is that word combination choices are more restricted than speakers intuitively suppose. The idiom principle alternates with the open choice principle, a "slot and filler" view of language whereby lexical and grammatical choices alternate in syntax to create meaningful units, these units typically corresponding to sentences.

> The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments.   (Sinclair 1991: 110)

Semi-preconstructed phrases may be very short sequences known as "collocational frameworks" (Renouf & Sinclair 1991) which

> consist of a discontinuous sequence of two words, positioned at one word remove from each other; they are therefore not grammatically self-standing; their well-formedness is dependent on what intervenes.   (ibid.: 128)

What is being described here is essentially a grammatical structure with a variable lexical slot, such as *a…of*; these frameworks are completed by the presence of the intervening word(s) and the collocates that this word attracts to the right of the structure. 'A group of single occurrences may seen to be constituting a class' (ibid.), as in Figure 3.1 (after Renouf & Sinclair 1991: 142) in which the slot-filler, *accident*, serves as an "idiomatic platform" (ibid.) for completion to the right by any of a series of semantically-related words.

> an accident of birth
> an accident of history
> an accident of history or birth
> an accident of fate
> an accident of post-war politics
> an accident of war

Figure 3.1 Completed collocational framework

> the faintest idea
> the least idea
> the slightest idea
> the foggiest idea
> the remotest idea
> the slightest conception
> the slightest notion
> the foggiest notion
> the remotest notion
> the foggiest

Figure 3.2 Semi-prepackaged phrase *the faintest idea*

As well as such grammar-centric phrase structures, there are others which are more obviously semantically-based. "Lexicogrammatical frames" (Moon 1998) consist of an invariable lexical item and a range of collocates which are, however, semantically related, e.g. *beyond belief/ doubt/ question/ recognition* (ibid.: 39).

> There is a common structure which contains a variable slot; the variable element is lexical, rather than grammatical, and the variations found tend to belong to the same semantic set. Because of the similarities in grammatical structure and lexical content, the meanings of the phrases can be said to be roughly synonymous.
> (Moon 1998: 145–146)

Even more variability can occur in what Francis (1993) calls "semi-prepackaged phrases", which are understood to be variants of one another, but in which 'there is no single lexical item which is essential' (ibid.: 144). Here, no invariable lexical item is required, unlike in a lexicogrammatical frame; rather the structure is determined by the colligation of grammatical categories whose lexical value is semantically related. Thus *the faintest idea* can be seen to be one possible realisation of a semi-prepackaged phrase whose 'only essential elements are the "superlative" marking of the adjective and the definite article *the* which accompanies it' (Francis 1993: 144). Other attested forms are listed in Figure 3.2.

These structures go some way to illustrate the principle of idiomaticity in language. In all, Sinclair notes seven features of language which point to the tendency towards idiomaticity being 'far more pervasive and elusive than we have allowed so far' (Sinclair 1991: 111). These are listed below.

1. Many phrases have an indeterminate extent;
2. Many phrases allow internal lexical variation;
3. Many phrases allow internal lexical syntactic modification;
4. Many phrases allow some variation in word order;
5. Many uses of words and phrases attract other words in strong collocation;

2nd proofs

6. Many uses of words and phrases show a tendency to co-occur with certain grammatical choices; and

7. Many uses of words and phrases show a tendency to occur in a certain semantic environment.   (ibid.: 111–112)

What is interesting to note in this list of features is the presence of the word "phrase". The idiom principle, both in itself and as expressed in the more detailed extended unit of meaning (Sinclair 1996b; §3.2) has tended to be applied to discrete words as focal points of phrases rather than to phrases as focal points within the wider cotext. Yet, as this book intends to demonstrate, phrases too operate within the idiom principle, generating what can be thought of as concentric rings of phraseology starting with the node phrase then moving out into its habitual cotextual environment, then farther out again into the text as a whole. We will see that idioms very often have indeterminate extent, that they allow various kinds of internal variation (to words, sequencing and syntax), that they attract collocates and grammatical choices, and that they tend to occur in particular semantic contexts. In order to do so, it is necessary to understand how these features operate in detail, and this is done with reference to the extended unit of meaning.

3.2   Extended units of meaning

The features of the *phraseological tendency* in language (Sinclair 1996b: 82) listed in the previous section can be seen as the precursor to a more fully developed statement of the idiom principle: the unit of meaning (Sinclair 1996b). In this view, meaning extends well beyond the boundaries of the single word and is better expressed as being the sum of the word, its cotextual environment, its contextual (situational) use, and the expressive and pragmatic meaning being conveyed in the communicative act (features 5, 6 and 7 on the list above; see also §3.1). In more technical terms, then, the focal point or *node* (which is the word or phrase under study) enters into lexical relations (collocation) and grammatical preferences (colligation) in the cotext. Viewing the range of collocates provides evidence of the preferred semantic field(s) of reference (semantic preference) of the node in its semi-fixed configuration. Beyond these information-centric aspects of meaning is the 'functional choice which links meaning to purpose' (ibid.: 88): the semantic prosody.

Collocation, colligation, semantic preference and semantic prosody have all been mentioned in earlier chapters, but no strict definition has as yet been given of any of them. As with many technical terms, they are all used to refer to slightly different phenomena by different authors, and while the distinctions are often subtle,

they do need to be noted. What also needs to be addressed is how they relate to idioms which, as we saw in Chapter 2, are in some respects like long words, and in other respects like phrases. This is especially important when addressing the more abstract notions of the unit of meaning, especially semantic prosody, as it has long been noted that what unites idiomatic expressions as a class 'is that the words are used with a particular pragmatic intention by the utterer, instead of or in addition to the intention to communicate information' (Hanks 1987: 134). To give a fleeting example, 'in the most common meaning of [*bitch*], what is at stake is the utterer's intention to insult, not the semantic convention associated with the sense' (ibid.).

The following sections discuss each of the features of the extended unit of meaning in turn, using idioms to illustrate the points made, rather than the more ubiquitous single words or phrasal verbs. True to the description offered earlier of concentric rings of phraseology (§3.1), each of these elements will then have a chapter-length survey dedicated to them in which colour-word idioms – both in their canonical forms and in their attested variants – are subjected to more rigorous analysis and comment from their collocational patternings to their connotative values, and, indeed, beyond.

3.3   Collocation

Collocation as a technical term in linguistics is most commonly associated with the name of J. R. Firth, although he was not the first linguist to use it nor is his use of the term the only one. However, within corpus linguistics, Firthian collocation is the one that is most frequently referred to, so we will use it as a starting point for our discussion.

Firth introduces collocation in his 1951 paper *Modes of Meaning*, where he writes:

> I propose to bring forward as a technical term, meaning by "collocation", and to apply the test of "collocability".   (Firth 1951: 194)

He goes on to demonstrate how 'part of the meaning of the word *ass* in modern colloquial English can be by collocation' (ibid.). Firth does not define collocation clearly in that paper, nor elsewhere; what precisely it refers to has to be reconstructed from comments and examples scattered throughout his works.

In the most general terms,

> The habitual collocations in which words under study appear are quite simply the word accompaniment, the other word-material in which they are most commonly or most characteristically embedded.   (Firth 1957: 180)

The meaning that words convey in text is a product of collocation, and this is to be differentiated from meaning in the mind.

> Meaning by collocation is an abstraction at the syntagmatic level and is not directly concerned with the conceptual or idea approach to the meaning of words. One of the meanings of *night* its collocability with *dark*, and of *dark*, of course, collocation with *night*. (Firth 1951: 196)

These citations tell us what collocations *do*. The remainder of this section looks at what collocations *are*.

3.3.1   Defining collocation

Firth's conception of collocation is usually articulated with reference to a few favourite citations, mainly from *A synopsis of linguistic theory 1930–1957* (Firth 1957). One of these is:

> The collocation of a word or a "piece" is not to be regarded as mere juxtaposition, it is an order of *mutual expectancy*. The words are mutually expected and mutually prehended'. (ibid.: 181)

Another is that 'a word in a usual collocation stares you in the face just as it is' (ibid.: 182).

However, looking more closely at Firth's writings, we discover that collocation is deemed to be particularly important for 'key-words, pivotal words, leading words' (Firth 1956: 106) occurring in the general language, in restricted languages, or indeed in the language of an individual (ibid.), and may be 'personal and idiosyncratic, or normal' (Firth 1952: 18). Collocation concerns word forms, not lemmas (Firth 1957: 181), and although it is primarily linked to common and high-frequency co-occurrence, it can also refer to 'a-normal' word pairings. Interestingly – because this seems to have been overlooked by later writers – collocations 'are statements of the habitual or customary places of that word in collocational order but not in any other contextual order and emphatically not in any grammatical order' (1957: 181).

Firth's examples of collocations mainly feature word pairs, but longer collocations are also discussed (notably in Firth 1951 and 1952) as are "extended collocations" (1956: 105); and while most of his examples involve content words, there is evidence of his admitting particles and other function words too.

Most linguistic research into collocations stems from lexicographic traditions. On the one hand, there is corpus linguistics whose methodology, initially developed in the corpus-based lexicography of the COBUILD project (Sinclair

1987), is characterised by the analysis of collocations in digital data. Collocation also has a long "analogue" history (Louw 2007) in classical Russian phraseology. Here it has a distinctly semantic rather than frequency-derived slant: collocations are still recognised as being frequent co-occurrences of words, but those combinations must not only be frequent, they must also be semantically salient. As a result, collocation is sometimes restricted to content words only (e.g. Hausmann 1984; Mel'čuk 1995, 1998), and the semantic relationship between the constituents changes: meaning by collocation is no longer a blend in which the constituent words make an equal contribution, but is instead *directional.* In other words, one of the words (the semantically dependent "collocate") modifies the meaning of the other (the semantically autonomous "base").

The base–collocate classification is usually described formally by part of speech, e.g. adjective+noun, verb+noun, etc., except by Mel'čuk (1995, 1998) who classifies them by *lexical function*, e.g. intensification, support verb, causative, etc. In either case, collocations are pairings of two content words whose combined meaning is constant and institutionalised. The semantic relationship between the constituents can be expressed as follows.

> A collocation **AB** of language **L** is a semantic phraseme of **L** such that its signified 'X' is constructed out of the signified of one of its two constituent lexemes – say of **A** – and a signified 'C' ['X' = 'A⊕C'] such that the lexeme **B** expresses 'C' only contingent on **A**. (Mel'čuk 1998: 30)

Regarding collocations as a combination of a base and a collocate can be seen to be of practical importance for the inclusion of collocations in dictionaries. This is not unproblematic. Importantly, it is not always obvious which of the words should be considered the base and which the collocate. For instance, in *a pack of dogs*, should the base be *pack* or *dogs*? Either one could be said to qualify the other, the former as a quantifier (*how many* dogs?), the latter as a qualifier (a pack of *what*?). A dictionary user would presumably look under the headword *dog* to discover what to call a group of dogs (cf. Hausmann 1985), but this does not entirely solve the problem. Bartsch (2004: 37) points out that the meaning of *pack* is determined within the collocation *a pack of*, and that in a different structure, even in collocation with *dogs*, this meaning might not be activated. In short, function words also contribute to phrasal meaning, and should not be ignored.

The base–collocate view of collocation is not simply a matter concerning positioning in dictionaries, but has semantic implications too. Viewing the constituents of a collocation as semantically unequal has repercussions on how the meaning itself is perceived. The implication is that a base is coloured by its collocate, but not changed, and that the collocate contributes a nuance or indeed changes its meaning. Again, this is problematic, because not all collocations behave this

way. Meľčuk provides a number of collocations whose base is *assurance* (French: 'insurance'). Depending on its collocate (e.g. *vie* 'life', *auto* 'car' *maladie* 'illness', or *incendie* 'fire') *assurance* means 'protection of' or 'protection against' (Meľčuk 1998: 41). The simple solution is to exclude from the category of "collocation" any word combinations that do not fit the model: thus collocations whose semantic relationship is opaque, whose semantic directionality cannot be ascertained, and/or which in combination acquire additional meaning not accounted for in the meanings of the constituents become "quasi-phrasemes" or "quasi-idioms" (Meľčuk 1998: 29–30). These include the collocations that I have been referring to as metaphorical collocations (§2.2).

Although neither Hausmann or Meľčuk consider function words as collocates, they are not entirely excluded from consideration. Benson et al. (1986a) provide both grammatical and lexical collocations in their *BBI Combinatory dictionary of English*. Grammatical collocations consist of 'a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or a clause' (ibid.: ix), while lexical collocations 'normally do not contain prepositions, infinitives and clauses' but typically 'consist of nouns, adjectives, verbs, and adverbs' (ibid.: xxiv). It is of note that only grammatical collocations include a "dominant word", because this suggests that the constituents of lexical collocations have equal semantic status. What is necessary is that the collocation has psychological salience as a unit (Benson et al. 1986b: 253), which, it will be recalled, is also a criterion for the classification of word strings as phrases or idioms (cf. §2.1.2).

Collocations can therefore be viewed as co-creating meaning (in Firth) or as modifying meaning (in the Russian tradition). There is of course a third way. Collocation may also sharpen the focus of the constituents' meaning. As Sinclair explains,

> *night* does not distinguish one of the meanings of *dark*, nor does *dark* distinguish one of the meanings of night. The phrase *dark night* has its own meaning; roughly speaking, the notion "dark" is already present in the notion *night* (though not all nights are dark, it is characteristic of a night to be dark). So the adjective *dark* is not selecting among all possible nights, the dark ones, but is reinforcing the dark element already in *night*.   (Sinclair & Teubert 2004: xxi)

This view of collocation has important parallels with how metaphor works in the creation of meaning, as will be seen in §3.6.4. In this book, unless stated otherwise, meaning by collocation is implicitly taken to mean restriction or focusing of the meaning of all constituents of the collocation. The constituents are referred to as "collocates", but with no implication of a base–collocate relationship.

3.3.2   Collocation in the digital age

The Firthian concept of collocation is the one that most corpus linguists refer to; and since computers have greatly facilitated access to information about language, it is inevitable that corpus linguists should have dedicated so much of their energy to collocation. Indeed, collocation lies at the heart of corpus query software applications and the analysis of concordances.

The most important contributions of computers to the study of collocations are (i) retrievability and (ii) quantification. As far as retrievability is concerned, corpus query software extracts all instances of a node, each in its original context, and therefore with all its attested cotexts. This means that high-frequency collocations whose salience is middling or low can be extracted in addition to those whose salience is high and which are therefore intuitively recoverable. In particular, it is possible to identify grammatical collocations including low-salience support verb constructions (*take a look*, *have a go*) and combinations such as *Friday night* or *Monday morning* which look compositional on the surface but which carry cultural implications as collocations which are not suggested by their constituents.

Within corpus linguistics, a word combination can be considered as a collocation if it occurs recurrently in the data, which means at least twice (Sinclair 1991) though Clear considers three occurrences as the minimum requirement (1993: 277). Co-occurrence need not be contiguous nor need its sequencing be fixed, but the collocates are expected to occur within five words left or right of each other, and the closer they are, the more significant their relationship.[1] This can be appreciated by looking at the collocation plot for *the pot calling the kettle black* from a data file comprising all the canonical forms and variants attested in the data (Table 3.1).[2] The words that comprise the canonical form are highlighted in bold type so that their distribution is more immediately visible. From the distributional profile of the collocates, it is evident that there is a preferred surface form comprising both lexical and grammatical collocates (i.e. the canonical form), but we also see that the collocates' positions are reasonably flexible (*pot* is present at L4, L3, L2 and R3). The plot also demonstrates the importance of proximity; the number of collocates (distinct word forms, or types) increases with greater distance from the node, while the number of tokens of each collocate (occurrences of the same word form) increases with greater proximity.

---

1.   'We must realize that the wider the span, the lower is the significance in general. The span ±2 can still be very powerful, but once we come to ±4 or ±5, significance will decrease' (Sinclair & Teubert 2004: xxvii).

2.   The full set of concordance lines can be consulted in Figure 3.3 (§3.3.3).

2nd proofs

Table 3.1 Collocation plot for *the pot calling the kettle black*

L5	L4	L3	L2	L1	NODE	R1	R2	R3	R4	R5
of	**the**	pot	**calling**	the	kettle	**black**	the	pot	to	to
like	of	words	**pot**	and		and	I	is	is	the
is	**pot**	pots	and			**calling**	and	I	**black**	more
that							but	a	in	it
the								and		a
about										and
a										in
be										
heard										
case										

Table 3.2 Collocation plot for *con le mani nel sacco*

L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
di	di	di	per	preso	NODE	in	si	con	un	a
e	che	il	lo	colto		ma	per	fa	a	la
allo	scopo	ti	è	colti		e	il	a	qui	se
	un	un	perché	cogliere		il			s	i
	ma	è	abbiamo	beccato		poi			o	l'
	decine	si		sorpreso		è			che	
	a	che		pizzicato		mentre			i	
		li		presi		non			la	
		quando								

Because this plot charts an institutionalised phrase, the number of collocate types that occur is limited by the words that comprise the phrase itself. However, one can consider a canonical phrase as a multi-word node and observe the collocational patterns of that node more closely, as in Table 3.2, which shows the collocation plot for the Italian equivalent of the idiom *caught red-handed, con le mani nel sacco*.[3] Here, the entire phrase was substituted by "node" in a data file. That file was then run through a concordance program, so that information about the collocates of the phrase as a whole, not just one of its constituents, could be charted.

The greater number of collocates listed here is partly due to the more complex morphology of Italian (for example, *il*, *lo*, *l'* and *i* would all appear as *the* in English); however, it is apparent that the phrase has a lexical collocate immediately to the left, usually a past participle. It is interesting to note that the presentation

---

3. For discussion of this data and full concordance listing, see §3.4.2.

of the data in Table 3.2 masks another collocational phenomenon: restriction of polysemy (but see Figure 3.8 below). The string *mani nel sacco*, which could be used literally to say *hands in the/his/her bag*, collocates with the function words *con* and *le* immediately to the left in every one of its occurrences, and each of those occurrences in turn collocates, again to the left, with a verb belonging to the semantic class "take/catch". A similar point is made by Sinclair et al. (2004) with regard to the collocates of *herring*. Its literal senses collocate with fishing and cooking terms, but its idiomatic sense (*red herring*) avoids these (2004: 94–98). Hoey (2005: 82ff.) has also demonstrated that polysemous senses tend to avoid each others' collocational patternings, with less frequent senses in particular avoiding the collocations patternings of more frequent senses.

The collocation profiles in Tables 3.1 and 3.2 use raw frequency data, which is adequate when working with small quantities of data (< 50 occurrences). With larger quantities, relative frequency scores are required, either calculated with reference to standard deviations, e.g. z-score, t-score or by reference to observed/expected co-occurrence relative to random distribution, e.g. Mutual Information (MI, Church & Hanks 1990). Without going into too much detail, the former provide a good indication of collocational significance for function words and other high-frequency items, while MI provides richer information regarding lexical collocations (Clear 1993: 281; see Barnbrook 1996: 87–101 for an overview and description of statistical scores used in corpus linguistics). Consider the most frequent collocates of *wedding* by t-score and MI (Table 3.3).[4]

As a final note on frequency, it should be borne in mind that

> by far the majority of lexical items have a relative frequency in current English of less than 20 per million. The chance probability of such items occurring adjacent to each other diminishes to less than 1 in 2,300,000,000!   (Clear 1993: 274)

Viewed in this light, it seems obvious that measurements based on chance occurrence are not entirely satisfactory (see Sinclair & Teubert 2004: xxi–xxii; see also Kilgarriff 2005), but in the absence of anything more suitable they provide a starting point for collocation analysis that is vastly superior to human introspection.

3.3.3   Idioms as complex collocations

The collocation plots in Tables 3.1 and 3.2 illustrated that idioms are a kind of long collocation, and that they also attract collocates outside the notional boundaries of the phrase. There are good reasons for considering idioms as complex

---

4. Source: Bank of English 450m version.

Table 3.3 Top 20 collocates of *wedding* by t-score and MI

Collocate	Freq.	t-score	Collocate	Freq.	MI score
white	126	11.196524	gown	6	10.008875
a	88	7.002894	dress	25	8.986883
dress	24	4.889586	white	135	8.624997
big	14	3.629657	marry	7	8.344866
traditional	11	3.288344	dresses	3	8.039408
church	8	2.778094	cake	3	7.279469
her	11	2.663763	display	7	7.013867
had	12	2.527452	traditional	11	6.774913
gown	6	2.447269	church	8	5.713379
full	5	2.090126	married	5	5.665011
in	29	1.986662	wearing	3	5.640284
married	4	1.953995	big	15	5.062603
wanted	4	1.879131	social	7	4.776010
she	7	1.759833	wants	3	4.315294
dresses	3	1.725903	couple	3	4.136683
cake	3	1.721640	wanted	4	3.949345
want	4	1.712849	full	5	3.838345
I	14	1.709323	having	3	3.099578
wearing	3	1.699626	thing	3	3.085337
long	4	1.675887	want	4	2.700854

collocations, rather than considering collocations as a subset of set phrases (cf. Melčuk 1998:23). One is related to the psychological processing of idioms, which will be discussed in this section. Another is that idioms involve lexical collocates which are typically connected by grammatical collocates to create a hybrid, lexicogrammatical collocation. On the whole, the lexical collocates in idioms seem to be more psychologically salient than the grammatical collocates. This can be appreciated when canonical forms and their non-canonical instantiations are arranged sequentially as in Figure 3.3.

Figure 3.3 shows all the occurrences of *kettle* co-occurring with *pot* and/or *black*, the concordances having been arranged to illustrate the progressive "distillation" of *the pot calling the kettle black* into its core collocation *pot+kettle*, as attested in corpus data.[5] From the canonical form which, as can be seen in the first eight concordance lines, collocates with *a case of* to the left, and punctuation to the right, we find the phrase slowly being boiled down to its bare essentials, typified by such examples as *the phrase, 'Pot, kettle and black,' springs to mind* (line 34) and *Pot and kettle, or what?* (line 38).

───────────────

5.   Source: Bank of English 450m version.

```
 1  endship". A case of the pot calling the kettle black, I fear. <p> In 1983, the
 2  tc, it is a case of the pot calling the kettle black # mdash; P. Hudson, Jind
 3  t is this a case of the pot calling the kettle black? Could holidaymakers get
 4  is rather a case of the pot calling the kettle black? The RSPCA used to run an
 5  certainly a case of the pot calling the kettle black. Keatingspeak is now as t
 6  it is not a case of the pot calling the kettle black. <p> McEnroe genuinely b
 7   In a prime case of the pot calling the kettle black, 48-year-old Iglesias, wh
 8  s a classic case of the pot calling the kettle black, MEPs were obliged to cle
 9  which is a bit like the pot calling the kettle black. <p> As if to answer it,
10  one. <p> It's like the pot calling the kettle black," said parish council cle
11  her, because it was the pot calling the kettle black. <p> That Scotland cheat
12  MPs, this really is the pot calling the kettle black." He believes that the I
13  mpetitive would be `the pot calling the kettle black." It's hard to see how an
14  mind, that would be the pot calling the kettle black with a vengeance. And I
15  99 </dt> TALK about the pot calling the kettle black # Linfield actually won t
16  ry. WVW: Isn't that the pot calling the kettle black? After her time in office
17  980703 </date> Sir: The pot calling the kettle black! Press tells doctors that
18  M0X> <ZGY> Rather <ZG0> pot calling the kettle black <ZGY> <M0X> Absolutely. <
19  elvin Grove. January 12 Pot calling the kettle black? IT is not often that I h
20  eat in the kitchen and pots calling the kettle black. I wondered if it was a p
21  e happy. I've heard of pots calling the kettle black, but this is more in the
22  is this a case of the pot-i calling the kettle black # <p> Mahoney laughed and
23  n out of steam. <p> The pot calling the kettle grey? Labour will be a Viking o
24  p> Surely a case of the pot calling the kettle schwarz. <p> Germans have remai
25  `It is time the pot stopped calling the kettle `noir # The scandal is revealed
26  look like a case of the pot calling the kettle... Hannahs make-over has all th
27  aste, because that would be calling the kettle black, but I don't like the way
28   low she says to me the pan calling the kettle blackbottom and I had to tell h
29  sn't that a case of the Doc calling the kettle black? <h> John, Neil or Paddy?
30  been for years as well. So we've got a kettle calling the pot black round her
31  inion would be a classic example of the kettle calling the pot black. And er I
32  enever anything's gone wrong. The words kettle, black and pot suddenly spring
33  enner spent on petrol the phrase, `Pot, kettle and black," springs to mind.
34  so clever and witty, the words pot and kettle do spring to mind about his clo
35  awful afternoon, which called to mind `kettle" and `pot", and culminated in t
36  new boss is a useless jerk – a pot and kettle case if ever there was to try t
37  wonder, ever heard the words `pot" and `kettle # Take Bruce Anderson, the auth
38  ng out demons and evil spirits. Pot and kettle, or what? KEITH PORTEOUS WOOD G
39   of humility. Talk about Mr Pot and Mr Kettle? Finally I must chide you over
40  superiority. There is a bit of pot-and-kettle about its outrage. Growth aroun
41  urope. North and south are like pot and kettle and neither out-shines the othe
42  u know, really. This is the pot and the kettle getting together and painting e
43  e the speaker `sound stupid" (pot, meet kettle). She was sitting next to me in
44  no-smoking area?" Hello pot, my name's kettle. I have a phobia of dirty ashtr
45  it," he declared. Pat Pot meets Mariah Kettle. liamfay@clubi.ie <xr> 9108 </x
46  er there was a case of dear pot, yours kettle. Mandy didn't bother too much
47  mind you. The case of Pot versus Black Kettle (1927). A BAND in Texas have c
```

Figure 3.3 Progressive distillation of *the pot calling the kettle black*

What is important to point out is the ubiquity of the truncated form of the idiom. The canonical phrase with its "canonical cotext", *BE […]a case of the pot calling the kettle black*, occurs eight times; the canonical form in a freer contextual environment, a further eight times, and very minor modification in another four lines, bringing the total number of (near)-canonical forms to twenty. In

comparison, there are sixteen occurrences of just the collocates *pot* and *kettle* occurring in highly variable (though often formulaic) contexts, always alluding to the idiomatic meaning associated with the canonical form. Furthermore, the meaning of the canonical form is preserved despite the break-down of the syntactic unit, further supporting the idea that the idiomatic meaning is fixed by lexical collocation involving *pot* and *kettle* (usually but not necessarily in that order).

If the lexical collocates in idioms are always salient – and the data analysed in this study support that notion very strongly – this helps to explain how idiomatic readings are activated. Psycholinguists who favour the Configuration Hypothesis of idiom comprehension (Cacciari & Tabossi 1988; Tabossi & Zardon 1993) have noted that idiomatic readings are activated when a "key" is encountered in the string.

> The recognition of an idiom does not seem to proceed incrementally, that is, gradually over time. There is no evidence of any idiomatic meaning activation at all until the key word in the configuration is encountered, at which point the idiomatic meaning is fully activated.   (Glucksberg 2001: 71)

Precisely at what point the idiomatic meaning is activated varies from idiom to idiom, but there is a key 'portion of an idiom that allows access to the idiomatic meaning' (Hillert & Swinney 2001: 113). The idiom key remains something of a mystery, in the sense that it has not yet been defined in linguistic terms. All that is known is that once encountered, reading times speed up (provided the idiom appears in its canonical form, McGlone et al. 1994), which is taken as evidence of comprehension having taken place.

If considered as a collocational phenomenon, idiom keys may be regarded as nothing more mysterious than the final salient collocate whose presence confirms the idiomatic meaning and causes all other possible meanings to be rejected. Idioms, as was mentioned earlier, are 'highly overlearned word sequences that comprehenders have experience with as holistic units' (Titone & Connine 1999: 1655). For this reason, once an idiomatic string is started, the reader/hearer can anticipate its completion because 'the words and phrases that have occurred have induced a sense of semantic-structural incompleteness' (Sinclair & Mauranen 2006: 136). The last essential collocate which fixes the idiomatic meaning can be anticipated because 'prospection arises from experience, and so common patterns will be more securely expected than rare ones' (ibid.: 137).

Cacciari & Tabossi (1998) note that readers do not process the salient meaning of *cielo* 'heaven' within the idiom *al settimo cielo* 'in seventh heaven'.[6] The

---

6.   Processing is determined by choosing one of three semantically-related key words: in this particular case, they were *santo* 'saint', which is linked to the salient/literal meaning;

Table 3.4 Collocates of *al settimo* at node +1

al settimo	cielo (64)	*in seventh heaven*
	piano (40)	*on the seventh floor*
	posto (38)	*in seventh place*
	mese (26)	*in the seventh month*
	giorno (10)	*on the seventh day*
	comma (8)	*in the seventh [legal] paragraph*
	secolo (5)	*in the seventh century*

idiomatic meaning, fixed in collocation, is clearly dominant in terms of psychological salience. This is not entirely surprising if we take two things into consideration. One is that the idiom in normal use and in the experimental setting (Example (1), Cacciari & Tabossi 1998: 671) is located in a context that prospects an expression of happiness.

(1)   After the excellent performance, the tennis player was **in seventh heaven**.

The other is that *cielo* is the most frequent right-hand collocate of *al settimo* (Table 3.4), and in this form always means 'delighted'.[7]

Thus context sets up the expectation for happiness to be lexicalised, and this expectation is met with *al settimo cielo* in its most common, idiomatic meaning. With the collocates being contiguous, there is no need to wait for the idiom key to appear, so the idiomatic (collocational) meaning is immediately activated: this involves height as the relevant attribute of *cielo*, enhanced by *settimo* (progressive numbering), and despite the origins of the expression its meaning has no religious or heavenly significance.

The meaning of an idiom may well hinge on its lexical collocates, but its grammatical relations are also important in consolidating the relationship between phrasal form and meaning. However, sometimes the collocational relationship operates at a more abstract level, featuring members of particular lexical and semantic sets within a grammatical structure. These abstractions are described in the next section.

---

felice 'happy', related to the idiomatic meaning; and the unrelated control *ombrello* 'umbrella' (Cacciari & Tabossi 1998: 671).

7.   Source: CorIS.

3.4   Grammatical and lexical abstractions

While collocation is a feature of the actual co-occurrence of words, all the remaining elements of the unit of meaning operate at a more abstract level. Colligation and semantic preference are both at one remove from collocation, while semantic prosody signals a further step into abstraction (see §3.5).

3.4.1   Colligation

Colligation is a Firthian concept relating to co-occurrence of grammatical classes.

> The statement of meaning at the grammatical level is in terms of word and sentence classes or of similar categories and of the interrelation of those categories in *colligations*.   (Firth 1957: 181; emphasis in original)

In this original conception, 'colligations cannot be of words as such' (ibid.: 182). This means that unlike in the current use of the term which describes grammatical classes co-occurring with a given node, there is no concrete node to speak of. Colligation can thus be contrasted with collocation, which is 'actual words in habitual company' (ibid.: 182).

> Colligations of grammatical categories related in a given structure do not necessarily follow word divisions or even sub-divisions of words […] A colligation is not to be interpreted as abstraction in parallel with a collocation of exemplifying words in a text.   (ibid.: 182–183)

Something approaching colligation was illustrated above in Figure 3.2 with the instantiations of the lexicogrammatical frame associated with *the faintest idea*. In that frame, there are no obligatory words, so no node can be isolated; what there is instead is the colligation of the definite article, a superlative adjective, and a noun. The meanings of the related instantiations of the phrase rely on this colligation in combination with the semantic preference for the adjective to refer to small quantities, and the noun to refer to knowledge or belief.

Although based on Firth's concept, the more widespread Sinclairian use of colligation describes the co-occurrence of a class of grammatical items with a specified node.[8] For instance, regarding the node *true feelings*, Sinclair notes that 'there is a strong colligation with a possessive adjective' (ibid.: 89). Other kinds of

---

8.   Sinclair acknowledges that his use of the term is different from Firth's, which he calls "full colligation" (1996b: 85), but justifies its looser application on the grounds that 'it is an extremely useful concept at this stage of our investigations' (ibid.).

colligation might be a preference for a particular verb tense, negative particles, modal verbs, participles, *that-* clauses, and so on. The notion that words may prefer (or indeed, avoid) particular positions in text is picked up by Hoey (2005) in his more detailed definition of colligation.

> The basic idea of colligation is that just as a lexical item may be primed to co-occur with another lexical item, so also it may be primed to occur in or with a particular grammatical function. Alternatively, it may be primed to avoid appearance in or co-occurrence with a particular grammatical function.
>   (Hoey 2005: 43)

Hoey attributes his use of colligation also to refer to sentential position as deriving from Halliday (ibid.: 42–43); it can, of course, also be seen as a natural extension of considering punctuation as a grammatical class, because punctuation is one of the most obvious indicators of positioning in text.

The concordance lines for *red in tooth and claw* shown in Figure 3.4 demonstrate some aspects of colligation working together.[9] By far and away the most frequent collocate is *nature*, which occurs nine times immediately to the left of the phrase and again at two positions to the left. *Red in tooth and claw* is a noun modifier, but, unusually for English, it has a pronounced preference for postpositive position (occurring after the noun it modifies, rather than before), and does so almost always in clause-final position.[10] The only exception to this is *red-in-tooth-and-claw masculinism*, where the phrase is hyphenated to ensure its integrity and thus preserve its non-compositional meaning. As with most modifiers, *red in tooth and claw* can also be used predicatively, though this is uncommon, this particular phrase being typically used for noun modification.

The colligation of *red in tooth and claw* as postpositive modifier is particularly striking because this is an unusual form in English, usually connected to foreign borrowings *(steak Diane, duck à l'orange)* rather than home-grown forms. Sometimes colligation can be less easy to spot because it is taken for granted as part of the grammar and, unlike collocation, does not 'stare you in the face just as it is' (Firth 1957: 182). One way of overcoming this limitation is to look at foreign language data, because it can be easier to notice regularities in an unfamiliar language. To this end, let us look at *pecora nera della famiglia*, the Italian form

---

9.   Source: Bank of English 400m version.

10.   Both the collocate *nature* and the odd syntactic positioning are a result of this being, in origin, a citation: "Who trusted God was love indeed / And love Creation's final law / Tho' Nature, red in tooth and claw / With ravine, shriek'd against his creed" (Tennyson 1850. *In Memoriam A.H.H.* Canto 56).

```
oo is just competition. <h> Red in tooth and claw: Comparative advertising (1402)
stirrup-cup socialists. <p> Red in tooth and claw? And politics. <p> D'they ken J
as a competitive business # Red in tooth and claw' or 'a struggle for existence'
the royal and ancient game. Red in tooth and claw and unashamed of it, he became
 real spirit of capitalism, red in tooth and claw, than corporate men, becoming m
icense granted to commerce, red in tooth and claw, permitted him to assert more t
 commented one contributor, red in tooth and claw. <p> Never red in politics, tho
Saxon wheeler-dealer model, red in tooth and claw, is pitted against systems like
 In Memoriam about 'Nature, red in tooth and claw' had struck many a responsive c
ky. <p> Wertheimer: Nature, red in tooth and claw. Is Montana ready for its retur
e nearest we got to nature, red in tooth and claw, was once again Rachel the Rece
ins the primordial animals 'red in tooth and claw"; it contains the potential see
bring guilt. Carnivores are red in tooth and claw. The only really jolly word I c
work for companies that are red in tooth and claw in terms of private enterprise
denounced the opposition as red in tooth and claw because it suggested that Lord
 and not the one we know as red in tooth and claw. But, still, the challenge is v
 to characterize nature as 'red in tooth and claw," and what led Herbert Spencer
lec Baldwin as a capitalist red in tooth and claw – Fuck you is my name. I drove
d one on predators entitled Red in Tooth and claw. Which is more about fleas and
 sleazy tale of a Hollywood red in tooth and claw but with a big yellow streak ru
lities of a society that is red in tooth and claw compared with what people here
ive, green feminism meeting red-in-tooth-and-claw masculinism: all men are rapist
the thing we do that's most red in tooth and claw, we with our easy lives in the
 the bush to take on nature red in tooth and claw. It grew darker and darker as w
porary thinking. <p> Nature red in tooth and claw/ Disasters great and small/ Ato
ecosystems. Watching nature red in tooth and claw is certainly one way to see cou
iled wanly at Anna. 'Nature red in tooth and claw. That's what it feels like, tea
s, it is not so much nature red in tooth and claw" that keeps populations within
appreciate this. <p> Nature red in tooth and claw," I told her, 'is only one way
 alternative to the nature 'red in tooth and claw" that characterizes the rest of
te that Labour is no longer red in tooth and claw, but a serious and sober politi
ians?" <h> A great shade of red in tooth and claw </h> <b> By ROS COWARD </b> <p>
had the hearts of predators red in tooth and claw. <p> But unlike Zuckerman, who
d to it. <h> Books: Testing red in tooth and claw </h> <b> By MIKE HOLDERNESS </b
```

Figure 3.4 *Red in tooth and claw*

of *black sheep of the family*, illustrated in Figure 3.5.[11] Speakers of any of the Romance languages should be able to make some sense of the Italian text, though its unfamiliarity allows its formal aspects to be discussed without meaning getting too much in the way.

The concordances listed (all occurrences of *pecora nera*, minus lines referring to actual sheep which have black fleece) show that the idiom is of indeterminate length, the collocation *pecora nera* being the only constant. Collocates to the immediate left of the node include the definite article *la* (11 times) and its fusion with the preposition *su* 'on' (as *sulla*) a further twice, plus the demonstrative pronoun *quella* (1). There is also the indeterminate quantifier *qualche* (2) and one possessive, *la sua* 'its'. To the immediate right, there are two striking collocational patterns: the first is clause-final punctuation (seven full stops, one colon, comma

---

11. Source: CorIS.

```
stirpe americana, di cui egli era la pecora nera. Non sto cercando di dargli dell
, questo non significa che ci sia la pecora nera. Il mio consiglio è quello di pr
i occhi brucianti di vita, Tonia, la pecora nera. Chinò la testa in ossequioso s
a vide. Tonia la contadina. Tonia la pecora nera. Tonia la serva. Come si era in
Poi nel gregge si è scoperta qualche pecora nera. Può essere fisiologico, no?". N
grande come la nostra ci sia qualche pecora nera. Purtroppo, scoprire il ladro no
era tempo che si separasse da quella pecora nera. "Sento la mancanza di un compag
nsorzio, la decima che rifiuta è una pecora nera: non riesce a dare molto fastid
 della loro famiglia, fosse pure una pecora nera come Josef. «Venga» disse il gen
re in loro favore. La considerano la pecora nera dei Wagner. Mi sento, piuttosto
 parti, l'Italia, fino a due mesi fa pecora nera del gregge europeo, viene ora pr
quio, che vive a Madrid, è il cugino pecora nera del re Juan Carlos, figlio di Al
steban Trueba veniva informato sulla pecora nera della famiglia attraverso la spo
ambi diplomati. Ma Maria? Era lei la pecora nera della famiglia, la più spregevo
figlio minore del re, considerato la pecora nera della famiglia, l'unico realmen
 in un temporalesco Cumulonimbus, la pecora nera della famiglia. Il meteorologo c
uppi con loro. La mia reputazione di pecora nera della famiglia costituiva, fortu
qualche giorno fa la Olivetti era la pecora nera dell'industria italiana e molti
nel), prima delle donne con 3'47"00; pecora nera della manifestazione invece Fab
va le pecore, fece affidamento sulla pecora nera della sua famiglia: Lazzaro. Co
a cooperativa, la Santa Fara, sia la pecora nera della zona: quella che non vuole
i a Margaret, additandola come vera "pecora nera" di casa Windsor. E per Snowdon
chio, perché ormai mi considerano la pecora nera del quartiere. Così quando hanno
riazioni genetiche. "Qui io sono una pecora nera" dice David C. Rowe dell'Univers
he in tutte le professioni esista la pecora nera ed è per questo che la nostra at
scandalo'. – Ogni famiglia ha la sua pecora nera – sentenziava la Maria del Cason
```

Figure 3.5 *Pecora nera*

and dash), the second is the preposition *di*, in all but one of the fourteen instances fused with the definite article. Wherever there is the preposition *di* (*dei*, *del*, *della*, *dell'*), a noun follows, the only recurring form being *famiglia* 'family'.

The immediate right and left collocates of *pecora nera* therefore tell us that it colligates with the possessive, almost always in the structure *la…di…* (the equivalent of the English *the…of…* possessive), but also with *la sua…* (the equivalent of *its/their…*). When it does not colligate with the possessive, the phrase colligates with clause-final position. Other colligates which are less immediate because both their forms and their positioning are more variable, are verbs in the subjunctive mood (indicated in bold type) and others, notably *considerare* 'to consider' (see underlined words), which express opinion/belief and which are therefore related to the function of the subjunctive in Italian. Therefore there are three colligations present: the possessive immediately surrounding the node, clause- or sentence-final position, and the subjunctive mood realized in the grammar and reinforced lexically in the preceding cotext.

3.4.2   Semantic preference

While colligation is an abstraction of the grammatical patterning associated with a node, semantic preference, the third element of the unit of meaning, is an abstraction of the semantic patterning which pulls together collocates and non-recurring near-synonyms into semantic sets. One way of looking at semantic preference is with reference to the (completed) collocational frameworks and lexicogrammatical frames that were introduced in §3.1. In both of these structures, an invariable lexical element (*an accident of…*, *beyond…*) is completed by any one of a range of words, the range defined by semantic similarity.

Idioms can withstand variation, but they can also be viewed as particularly stable realisations of lexical and grammatical preferences. For instance, *once in a blue moon* can be seen to slot effortlessly into the *once in a lifetime* frame, named after its most frequent realisation (Figure 3.6). In order to belong to this lexicogrammatical frame, the right-hand collocate of *once in a…* must refer to a period of time, typically a period long enough not to recur in a person's lifetime. There are some very frequent realisations of the pattern (*lifetime*, *while*, *blue moon*), and some nonce forms (*career*, *house on fire*) whose interpretation is partially coloured by previous encounters with the more frequent forms (Sinclair & Mauranen 2006; Kerbrat-Orecchioni 1977). This aspect of meaning will be dealt with in §3.5, so no further comment will be made at this stage: suffice it to say that recurrent collocates in this variable slot all refer to extended periods of time, and although each of the conventionalised phrases in the paradigm has its

```
Once in a…   lifetime (500)
             while (851)
                  great while (8)
                  long while (3)
                  wee while (2)
             blue moon (68)
             generation (27)
             decade (13)
             century (12)
             millennium (8)
             thousand year(s) (8)
             hundred years (6)
             million years (2)
             Preston Guild (2)
             career (1)
             house on fire (1)
```

Figure 3.6 Lexicogrammatical frame for *once in a lifetime*

own shades and nuances, they all convey similar information can therefore be expected to occur in similar kinds of contexts.

As well as lexicogrammatical frames, Moon identifies a further type of idiom variation which she calls "idiom schemas" (Moon 1998: 161–170). These describe a peculiar kind of idiom which has no canonical form as such, and, unlike lexicogrammatical frames where there is an invariable lexical component, idiom schemas are defined through colligation, the lexicalisation of the colligational patterning depending on semantic preference. Figure 3.7a shows the general variation patterns found inside the idiom schema for *[shake] in one's [shoes]* (after Moon 1998: 161), with the verbs reduced to their infinitive forms and the possessive pronoun normalised to the lexicographically conventional "one's". Figure 3.7b shows the rather messier instantiations of the same schema found in the British National Corpus (BNC).

On the face of it, idiom schemas seem no different to semi-prepackaged phrases. There is one important difference, however: in a semi-prepackaged phrase, the semantic preference for each of the variable slots conveys meaning as information. *The faintest idea*, for example, follows the classic collocational relationship of base (a noun meaning "idea") and collocate (an adjective expressing minimal

```
shake in one's shoes
quake in one's shoes
shake in one's boots
quake in one's boots
shiver in one's boots
quake in one's Doc Marten's
quake in one's size 11s
```

Figure 3.7a Idiom schema for *[shake] in one's [shoes]*

```
murmured that West Indies must be "quaking in their boots" at the thought
Aid made the likes of The Redskins quake   in their Doc Marten's and feel totally
Snow White's the one who should be quaking in her   boots. On reflecti
ted's giant striker, has Stuttgart quaking in their boots at the prospect of to
he dealing room bully, and dealers quaked  in their boots at the thought of
n, should have the wretched Estall quaking in his   boots. He is not
for Monklands, East (Mr. Smith) is quaking in his   shoes as the Labour Front
    Every time I do a stall I I quake   in my   boots up until I do it an
ce of the fact that Raskolnikov is shaking in his   shoes. The author kno
John walked across to them. He was shaking in his   shoes, but said as calmly
   spoke to me I was ju literally shake   in me   shoes, he gave you that
d despite a clear conscience began shaking in my   boots. I lit another
 to win the title, I'm not exactly shaking in my   shoes about facing Benichou.
rs, politicians and industrialists shaking in their boots. The
o make a decent Human-eating giant shiver  in his   boots and think shame!
```

Figure 3.7b Instantiations of *[shake] in one's [shoes]*

quantity). In an idiom schema, however, the meaning of the whole is idiomatic, so the base–collocate relationship is much fuzzier, if relevant at all. Instead, the sets of semantic preferences must be consistent with what is often called the "image schema" (Lakoff 1987a) of the idiom. That image schema corresponds to a meaning which is established and institutionalised, even though the meaning is not connected to an equally established and institutionalised wording. So unlike semi-prepackaged phrases, idiom schemas

> are characterized by an underlying conceit [...] and an over-lying preferred lexical realization, usually with connoted evaluation. The exact form of words may vary or be exploited, but is still tied to the underlying conceit which provides the driving or motivating force.   (Moon 1998: 163)

Semantic preference can be instrumental in completing the meaning expressed by a phrase, as above, but more often it is thought of as being a contextual complement to a meaningful core. Semantic preference accounts for some of the variability in phrasal idioms, in which a core collocation acts as node, and attracts collocates, colligates, and so on. While several instances of this will be investigated in Chapter 4, here we will go back to Italian data and look at the functional equivalent of *caught red-handed*, expressed in Italian as *(preso) con le mani nel sacco* 'taken with one's hands in the bag' (Figure 3.8).[12] Here – as in the English equivalent – there is a core collocation, clearly visible as *con le mani nel sacco* (the query carried out was just for *mani nel sacco*). The verb, however, is variable, and not only because of the complex morphology of Italian. We can find a range of alternative verbs to the immediate left of the phrase, and on the basis of frequency alone, it is unclear which of *prendere* (16) and *cogliere* (15), should be considered canonical. These are synonyms, and are joined in meaning to *beccare* (5), *pescare* (2), *pizzicare* (2), and *cuccare* (1), the main difference being that these lower frequency items are all also more colloquial. Other repeated verbs related to the same general sense are *sorprendere* 'surprise/catch' (5), *scoprire* 'discover/catch' (2), and *trovare* 'find' (1), which highlight the surprise element of being caught. The remaining two verbs, *arrestare* 'arrest' and *finire* 'end up' are one-off occurrences, the former fitting with the general sense of the other verbs, the latter referring only obliquely to being caught, although the surrounding cotext leaves us in no doubt.

This kind of variation in verbal idioms is very common, and one of the reasons why idioms are more likely to be listed in dictionaries under nouns than

---

12. Source: CorIS.

```
 compagnia di Aversa avevano arrestato con le mani nel sacco tre estorsori del
 t," che sfidò i giornalisti a beccarlo con le mani nel sacco e si meritò quell
a. Una carogna vestita a festa beccata con le mani nel sacco. - Cosa ci fa qui
 scura e brillante . Li abbiamo beccati con le mani nel sacco! gridò Don Diego.
naski! Questa volta ti abbiamo beccato con le mani nel sacco!" Ero stanco. Non
u of Investigation (Fbi) hanno beccato con le mani nel sacco un uomo dei servi
 un ladro attento a non farsi cogliere con le mani nel sacco. Perché se gli es
ttere solo allo scopo di far cogliere "con le mani nel sacco" il pubblico uffi
tto marchingegni alla 007 per cogliere con le mani nel sacco il marpione. Ma i
 metterlo fuori in un istante. Lo colsi con le mani nel sacco, intento a rubare
 che si era intascata 350 milioni Colta con le mani nel sacco. È finita in guai
i suoi nastri . Patrick li aveva colti con le mani nel sacco e, per quanto inf
a - insieme a decine di colleghi colti con le mani nel sacco - in un ufficio p
 scambiarono una rapida occhiata. Colti con le mani nel sacco, trovavano all'im
  la fiducia che, se proprio si è colti con le mani nel sacco, basterà restitui
amo che li abbiamo letteralmente colti con le mani nel sacco. Gli risponde con
ualcosa di prezioso senza essere colto con le mani nel sacco per pura fortuna.
 pubblici i termini. Il governo, colto con le mani nel sacco e perciò, secondo
entato al capitano, eh? Ma ti ho colto con le mani nel sacco. Matteo si era gi
a 9.000 sterline di multa perché colto con le mani nel sacco: custodiva gelosa
? Che in Italia un reo confesso, colto con le mani nel sacco, magari inchiodat
non sopporta, quello che si fa cuccare con le mani nel sacco, quello che ha be
 ronti ad arrestarlo. L'operaio finisce con le mani nel sacco, viene arrestato
 Pure se colgono uno di questi malnati con le mani nel sacco, devono metterlo
olizia forse vuole prendere il maniaco con le mani nel sacco, ma questo signif
ta certezza di avere preso una persona con le mani nel sacco". Attualmente il
vora per la Cia e che è stata pescata "con le mani nel sacco mentre era impegn
uo indirizzo esatto. Una volta pescato con le mani nel sacco, si potrà inviare
 un complice, lo sconosciuto pizzicato con le mani nel sacco aveva già caricat
ssato inosservato ed è stato pizzicato con le mani nel sacco mentre rovistava
droga da questi a quelli. Per prendere con le mani nel sacco i polacchi e bruc
veramente audace. Ma noi lo prenderemo con le mani nel sacco. Vieni, Sharp. Do
 Qualche tempo di attesa per prenderli con le mani nel sacco, poi i carabinier
sa, su quella jeep. Dobbiamo prenderlo con le mani nel sacco.» «In fondo, fa s
e anche domani, quando vorrà prenderti con le mani nel sacco. Ma quando il flu
i, ci sono ma decine di deputati presi con le mani nel sacco siedono tranquill
 rischio per i giovani * * * * * Presi con le mani nel sacco. O meglio con la
 - Psi, che nei primi anni 90 fu preso con le mani nel sacco di una brutta qua
re è quello di smettere quando è preso con le mani nel sacco. Poi gli tengono
unto da farlo sentire un bambino preso con le mani nel sacco. In realtà Arnold
ium abbia successo sia che venga preso con le mani nel sacco. Lo spiritismo è
ita". Nino Polacco, praticamente preso con le mani nel sacco, non rinunciò a m
a nell'atto di compiere il male. Preso con le mani nel sacco, non vivrà più di
broker compresi) che verranno scoperte con le mani nel sacco. In vista della d
osse stata uccisa da un ladro scoperto con le mani nel sacco, le occhiate di c
ttuata allo scopo di fare sorprendere "con le mani nel sacco" il funzionario i
' aria , smuove qualcosa, la sorprendi con le mani nel sacco. È il beccheggio,
ino portata in caserma perché sorpresa con le mani nel sacco. Soprattutto per
ome la sua patta. La Ronda lo sorprese con le mani nel sacco, la mano sul pacc
contato dunque il funzionario sorpreso con le mani nel sacco, erano inquinati.
La solita storia di chi viene sorpreso con le mani nel sacco? No, sembra propr
na coincidenza che siano stati trovati con le mani nel sacco nella stessa gior
ui Cesare Lombroso. Fu presa più volte con le mani nel sacco ma, come la volpe
```

Figure 3.8 *Con le mani nel sacco*

verbs.[13] Simply put, it is sometimes difficult to ascertain whether or not one particular verb is part of the idiom, or if instead the idiom merely contains a slot which is filled by any verb that broadly expresses a particular idea, along the lines of lexicogrammatical frame completion.

Less obvious as semantic preferences are the words dotted around the cotext referring to theft (including *estorsori* 'extortionists', *ladro* 'thief', *rubare* 'stealing', *intascata 350 milioni* 'pocketed 350 million', and so on) and to law enforcement (*arrestato* 'arrested', *FBI, marchingegni alla 007* 'Bond-style gadgets', *9.000 sterline di multa* 'a £9000 fine', *polizia* 'police', etc.). The cotext also reveals that not only archetypal criminals but also office workers (*impiegato, funzionario*) and officials (*che la diplomatica lavora per la Cia* 'that the diplomat worked for the CIA') are often the perpetrators of the crime. All in all, the semantic preference tells us a great deal about 'the relevant features of the participants: persons, personalities' (Firth 1950:182), and their (here, non-verbal) action (ibid.); in other words, it fills out the first of the three categories in Firth's context of situation (ibid.; see also §4.3).

3.5   Secondary semantics

3.5.1   The heart of the matter

The abstractions that fall under colligation or semantic preference are little more than a means of grouping the words in the cotext. Going a step further into abstraction requires care, because there is a fine line to be drawn between abstractions that can be inferred from the data (in Gricean meanings as events; see Hanks 1986), and those which are influenced by knowledge of what words mean (meanings as beliefs; ibid.). The element of the unit of meaning that deals with this is semantic prosody.

Semantic prosody is a problematic concept, mainly because it has been used to speak of such different things as connotation, evaluation, appraisal, pragmatic force, irony and insincerity, all of which are kinds of secondary meaning, and none of which is the same as the others. Confusion is rife and different camps seem bent on appropriating the term for their own ends, a matter not entirely helped by the architects themselves:

---

13. The *Oxford-Paravia Italian dictionary*, aimed at both learners and professionals, lists the idiom in five entries: as *essere preso con le mani nel sacco* under both *mano* and *sacco*, as *farsi prendere con le mani nel sacco* under *prendere*, *beccare qn. con le mani nel sacco* under *beccare*, and *pescare qn. con le mani nel sacco* under *pescare*.

Often the use of a word in a particular cotext carries extra meaning of an emotive or attitudinal nature; if you have a sporting victory, and in the presence of your opponents you call it a "walkover", you insult them, although the victory may have been very easy. This kind of meaning is sometimes called "connotation" (as against "denotation" for the ordinary meaning of the word), and sometimes this kind of meaning is called "pragmatic meaning".

We will see in this book that this kind of meaning is structurally important, and essential for the understanding of language text. We will call it SEMANTIC PROSODY – semantic because it deals with meaning, and prosody because it typically ranges over combinations of words in an utterance rather than being attached just to one.   (Sinclair 2003:117)

This piece has been cited in unabridged form to illustrate the important point that 'this kind of meaning' refers back to 'you insult them', insulting being a pragmatic function, not a connotation. The passage is confusing, however, because it starts off by talking about 'extra meaning of an emotive or attitudinal nature', i.e. connotation and evaluation. A reader who does not already know that connotation and pragmatic meaning are as alike as chalk and cheese would be likely to misconstrue them as being one and the same. Yet connotation is subjective, largely private and word-bound while pragmatic meaning is bound to function and intent in communicative situations and very much part of the shared system of language. Unfortunately, semantic prosody has been used to refer to both these types of meaning. §3.5.2 gives an overview of the evolution of the term, which will then be discussed with reference to data in §3.6 and §3.7.

3.5.2   Semantic prosody and related phenomena

Semantic prosody waited rather a long time to be given a name. Although attributed to Sinclair (1991, 1998), the germ of the idea is also apparent in the writings of his colleagues at Birmingham University and within the COBUILD project: consider Hanks' comments on idioms expressing pragmatic intention 'instead of or in addition to the intention to communicate information' (Hanks 1987:134, cf. §3.2).

As with many concepts in academia, semantic prosody was probably noticed and discussed informally amongst colleagues for some time before any of them decided to investigate it in depth. The name was publicly coined by Louw, a close associate of Sinclair's, in his 1993 paper, *Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies*. This paper remains the source of the most commonly-cited definition of semantic prosody: 'a consistent aura of meaning with which a form is imbued by its collocates' (Louw 1993:157).

Louw regrets the ubiquity of this definition on several counts (personal communication). One is that the paper has been interpreted as being an *exposition* of semantic prosody rather than an *application* of it: the exposition is comprehensively set out in later publications, notably Sinclair (1996b) and Louw (2000a); and yet these are often ignored or skirted over in favour of the earliest – and least developed – mentions of the phenomenon (especially Sinclair 1991, 1998; Louw 1993). But Louw's main objection to the eminently quotable "aura of meaning" definition of semantic prosody is that it has bolstered the widespread interpretation of semantic prosody as being essentially connotative rather than pragmatic, and that its primary function is to indicate positive or negative evaluation (see especially Stubbs 1995; Partington 2004; Dilts & Newman 2006; Hunston 2007; Bednarek 2008; Morley & Partington 2009), particularly with reference to irony and insincerity.[14]

Some writers have stuck with the notion that semantic prosody is essentially evaluative, deriving this interpretation from Louw's focus on textual effects as seen in relation to semantic prosody (Louw 1993, 1997). Within this view, semantic prosody is fairly uncomplicated, but could just as readily be called evaluation, as Morley & Partington (2009) rightly point out. Other researchers, however, have embraced a much more wide-ranging and complex conception of semantic prosody 'whose primary function is the expression of the attitude of its speaker or writer towards some pragmatic situation' (Louw 2000a: 57) and which 'expresses something close to the "function" of the item – it shows how the rest of the item is to be interpreted functionally' (Sinclair 1996b: 34). Needless to say, these conflicting interpretations have created a degree of terminological tension as the two sides battle it out for primacy over the term.

It would be fair to say that most scholars who have interested themselves in semantic prosody have noticed that the term seems to be used to refer to (at least) two quite distinct things. For "evaluative" semantic prosody, a dividing line has to be set between inferred evaluation (a product of individual interpretation) and conventionalised evaluation (institutionalised interpretations). Overall, recent years have seen a tendency to add a fifth category to the unit of meaning, in recognition of the fact that semantic prosody can refer to attitudinal and evaluative meanings on the one hand, and to the overall communicative function of the unit of meaning on the other.

---

[14]. Instead, he says, 'irony and insincerity are a binary choice *inside* the exceptions to a semantic prosody! This means roughly 3% of it. So, that means you cannot find the exceptions in large quantities without a computer. So to get 9 lines of exceptions you need 300 lines, and 3000 lines to get 90. Then, of those, 2 out of 3 will be irony and only 1 insincerity' (personal communication, March 2010).

Hoey notes that semantic prosody 'grows out of two different concepts, sometimes confused with each other' (2005: 22), and chooses to split these into "semantic associations" for the connotative and evaluative meanings, and "pragmatic associations" or the discourse function. He avoids entirely the term "semantic prosody" to maintain the difference between the representation of meaning in the mind, lexical priming, and its actualisation in text (ibid.). Hunston (2007) favours what is essentially the same split:

> My suggestion would be that the term "semantic prosody" is best restricted to Sinclair's use of it to refer to the discourse function of a unit of meaning, something that is resistant to precise articulation and that may well not be definable as simply "positive" or "negative". I would suggest that a different term, such as "semantic preference" or perhaps "attitudinal" preference, is used to refer to the frequent co-occurrence of a lexical item with items expressing a particular evaluative meaning.   (Hunston 2007: 266)

Stewart (2009) too suggests a division of this kind. He chooses to label the evaluative type "semantic prosody" by virtue of having been introduced in Louw's paper, which mainly discussed the evaluative aspect of semantic prosody. To the functional-pragmatic type of semantic prosody discussed in Sinclair 1996b he assigns the label "discourse prosody" (after Stubbs 2001: 65), although he adds that 'this solution is far from perfect' and that it 'might have been preferable [...] to do away with the term "prosody" altogether' (Stewart 2009: 162). Philip (2009) proposes that Hoey's term "semantic association" should be introduced as an intermediate stage between semantic preference and semantic prosody. As the name suggests, this additional category covers connotative, expressive and evaluative meanings associated with lexical items, and allows "semantic prosody" to be used for the discourse function as in Sinclair's (1996b) conception and in keeping with the idea that 'semantic [...] deals with meaning, and prosody [...] typically ranges over combinations of words in an utterance rather than being attached just to one' (Sinclair 2003: 117). It is this division and terminology which will be followed for the remainder of this book.

3.6   Semantic prosody (i), a.k.a. semantic association

3.6.1   Idioms and evaluation

Idioms have attitude, and it is often this attitudinal meaning which is said to differentiate them from their literal counterparts. How true that supposition is can be difficult to ascertain because attitude is so intimately tied up with subjectivity,

but at the same time it is difficult to find a conventional idiom that is expressively neutral. Moon (1992) notes that while it is traditional to split fixed phrases by their morphology, a more (lexicographically) sound division is by discourse function. She identifies five such categories: informational, evaluative, situational, modalising, and organisational (Moon 1992: 495). Of these, the evaluative (conveying the speaker's evaluation and attitude), the situational (reflecting and/or responding to the discourse context) and the modalising (conveying truth values) are all primarily expressive rather than informative; and these account for around 40% of all fixed expressions (ibid.: 497). Furthermore, even those whose primary function is informative (propositional) are also evaluative (e.g. *rub shoulders with*, *at a snail's pace*, *rock the boat*; ibid.: 496–497). In contrast, fully compositional language is not tied to the expression of evaluation and attitude, but can choose to convey such meanings through the considered selection of single words.

Even a very small amount of data can be enough to suggest the attitudinal and evaluative meaning associated with an idiom. Although seven concordance lines (Figure 3.9) are far too few for conclusive statements to be made, there is enough attitudinally-loaded lexis in the surrounding cotext (in particular *graft*, *ghastly*, and *obsessions*) for us to say that *bleeding sb white* is viewed as a bad state of affairs by whoever chooses to use the expression. But how do loaded words colour meaning? Although it is a central concern of appraisal systems within systemic functional linguistics (Martin 2000) and a feature of language which speakers are consciously aware of, how a word which is not grammatically negative can be attributed a negative value is rarely addressed adequately. It is usually considered to be inherent in a word's meaning; and this is clearly problematic when working with the notion that words do not have inherent meaning.

To account for the evaluative connotations that certain words can trigger as a matter of course, it is helpful to return to the Gricean distinction between meanings as beliefs about words and meanings as events mentioned at the start of §3.5. Negatively-evaluated words are not negative in terms of their informational, denotative meaning, but refer to real-world phenomena which are viewed negatively within the language community. This evaluation is therefore a form of cultural connotation (Leech 1974; see §3.6.2 below), and boils down to a belief about

```
ort as proof that Mr Clinton has bled the military white. `Over the past several
in, too much graft could bleed any military effort white. Then any kind of enemy
 old Dido and all those ghastly offspring bled him white. As I told you yesterday
 that the twin obsessions were bleeding his estate white. His solution was to att
 for. Course he paid for by er bleeding his people white and so of course did the
 <p> In short, not content with bleeding the public white, they want us to do the
 ed enterprises that are bleeding the rest of China white. Thirdly, in 1997 the re
```

Figure 3.9 *Bleed sb white*

the words' meaning. It can be argued that cultural connotations emerge out of our memory of previous uses of a word, and this is generally how the 'consistent aura of meaning with which a form is imbued by its collocates' (Louw 1993: 157) has been interpreted. In other words, previous encounters with a word, as well as direct experience of the phenomena to which it refers, resonate with the current instantiation in text. This is the essence of Hoey's (2005) theory of lexical priming, and it is echoed in Louw's work on "fractured contexts of situation" (2000, 2005, 2007) as well as in current investigations into collocational resonance (see especially Hanks 2005, 2006; Williams 2008; Williams & Millon 2009).

Resonance is central to metaphorical appreciation (Black 1964, 1993), and although it is largely interpretative in nature, it can be separated into the private (the subjective and personal resonance as interpreted by an individual) and the public (shared by a language community and manifest in its language use). Private interpretations of resonance are too ephemeral to be of interest to us here; public ones, on the other hand, can be investigated as a collocational phenomenon using corpus data. While studies to date have preferred to approach resonance from a diachronic standpoint, contrasting a historical corpus with a synchronic one (Hanks 1998; Williams 2008), there is a case for using synchronic corpora. Resonance is a diachronic phenomenon, but, this is not the same as suggesting that etymology is relevant: it is not. Teubert (2007: 113) reminds us that synchronic discourse necessarily has a diachronic aspect because we acquire language over time and also read texts that were written in the past. Synchronic corpora tell us what kinds of resonance are likely to be current, including those which may only be perceived subliminally (cf. Louw 2005: 161–163).

Resonance pierces through layers of previous usage. *Bleed sb white* seems to have a negative tinge due to its co-occurrence with negatively-loaded terms. As Bednarek (2008) points out, this is not a semantic preference, because the words themselves are not semantically negative, but only connotatively so. Taking one of these collocates, *graft* (line 2), as an example, we can analyse its cotext to see if its perceived negativity is a product of its other contexts of use. Using the BNC as a reference corpus, we find that *graft* is strongly associated with onerous work (*hard graft* occurs as a contiguous collocation in 36 of the 88 occurrences for the "work" sense of *graft*). As well as this very obvious collocate – shared by the more neutral *work*, of course, but not nearly so prominently (only 322 in 3655 occurrences, or about 9%) – we find that (*hard*) *graft* in turn has a semantic preference for quantity measured primarily by time (*fifteen years of hard graft*, *a summer of hard graft*), in reference to getting by financially and professionally. It should be noted that *graft* is not totally negative – there is virtue attached to those who work hard (*He's prepared to work and graft*), but it carries a sense of reluctance and unfairness, both of which are negatively perceived. However one looks at it, *graft* is

something that is resented (Example 2): most would prefer to avoid it in favour of an easy life (Example 3).

(2) A student nowadays has to **graft** for three years on a pittance of a grant, begging the government for a loan and at the same time being excluded from any welfare benefits.

(3) Why **graft** your arse off for a full week to make £130, when you could make £150 for one day's shoplifting?

By back-tracking in this way and analysing the cotexts associated with each of the node's collocates, we arrive at a clearer picture of how evaluative meaning is built up through the patterns of use associated with those words, and can account for it resonating through into the present instantiation. In this way, the bottom-up analysis supports the top-down assumptions made about negativity inhering in word meaning. While *graft* may be associated primarily with a perception of unfairness expressed in relation to working hard for little recognition or financial reward (this being the "discourse function" semantic prosody, see §3.7 below), this is not true of all the loaded lexis in the cotext of *bleed sb white*. In fact, taken as a group, the loaded lexis has little in common, but shares traits that lead to a general suggestion of "negativity", polarity being the absolute essence of evaluation (Morley & Partington 2009: 141).

When attitudinally-loaded lexis co-occurs with the node under study, two observations can be made. The first is that the node's meaning is coloured *in that particular instantiation* due to its co-occurrence with the loaded lexis: this is how persuasive text can suggest that an otherwise neutral word should be interpreted in a particular light. The second point is that when, as in *bleed sb white*, the node consistently appears in the company of loaded words, then it should be viewed as being in "evaluative harmony" with them (Morley & Partington 2009: 139). So, despite Whitsitt's protests that word meaning cannot be "infected" by co-occurrence, that "innocent" words cannot fall 'into a world of "the brutal clash of usage", which leaves them […] forever imbued with negativity' (2005: 292), we find that alas, corpus evidence indicates that this kind of semantic "contagion" (Louw 1993: 159) can indeed occur.

Part of Whitsitt's argument lies in the belief that the words apparently coloured by their collocates would express that evaluation anyway. This might be true of an expression like *bleed sb white* which has a transparent connection to draining the blood out of a living body which *in extremis* leads to death (which is definitely bad), the same cannot be said of a more opaque expression such as *purple prose*. There is nothing inherent in either collocate to suggest that *purple prose* should be perceived negatively, but the concordances in Figure 3.10 confirm

```
1.   sauce. <p> When a menu tends towards purple prose, the cooking rarely lives
2.   gorgeous photographs and the kind of purple prose that has made Oz Clarke su
3.  he waiter who recites the long-winded purple prose of the menu of the day. A
4.  nd Bahama based only on the company's purple prose, while having to find your
5.   now 94 who still churns out acres of purple prose in a miasma of chintz at h
6.  which might well merit a paragraph of purple prose (from his barrister when h
7.  ionship, told in burnished, sometimes purple prose, is interspersed with even
8.  ter and has a tendency to launch into purple prose that should have been curb
9.   treat operators' brochures – and the purple prose within – as divine revelat
10.   tricky. There's a fine line between purple prose and techno-jargon. Seth ad
11. Connolly's incorrigible penchant for purple prose. Also partly set in Maine,
12. ase `gunning down" because it is not purple prose. The four cops fired 41 bu
13.   plea for action (if you can't stand purple prose, don't go to an environmen
14. d to a strange (and often hilarious) purple prose. <hl> Moreover </hl> <subh
15.   anecdotes and colourful – but never purple – prose, Mr Jarman has produced
16. lly journalists and critics to write purple prose. Answer: There is a tiny k
17. meet the expectations aroused by the purple prose. Prices are not cheap but
```

Figure 3.10 *Purple prose*

that it is. While there are no recurring loaded collocates, there is nevertheless considerable evidence that *purple prose* is disapproved of (see words highlighted in bold type).

Closer inspection of the concordance lines reveal that *purple prose* features in menus (lines 1, 2, 3), holiday brochures (lines 4, 17), romantic fiction (5, 7, 8) and art/music criticism (10, 16). These text types are known for their flowery, hyperbolic language, but even though *purple prose* is convoluted and over-long (qualities which tend not to be welcomed by English speakers), to say that these features make it negatively evaluated is missing the point somewhat. Equally important is what is missing, or avoided: there are no mentions of "serious" texts whose *purple prose* would be misleading, with negative consequences. Evaluation does not allow us to say that a speaker/writer describing text as *purple prose* is mocking its excesses, not criticising it. It is a form of overstatement that is not taken too seriously and it is not considered either threatening or manipulative. This, again, is the "discourse function" semantic prosody, which provides detail that the evaluative type does not (see §3.7).

The evaluative function of idioms comes to the fore when they are used to begin and to end texts. In these positions, idioms can also be used as standalone items – particularly in newspaper headlines, where they entice the reader to engage with a story which, as it unfolds, will presumably elaborate on that evaluation and also relate to the meaning of the idiom itself and/or one or more of the lexical items present in the idiom (Examples (4)–(6)).[15] While the examples shown here make use of idioms in their canonical forms, variants are also

---

15. Source: Bank of English 400m version: Ex. 4 and Ex. 5: British Magazines; Ex. 6: Today; Ex. 7 Independent; Ex. 8: Today; Ex. 9: Sun.

common (see Chapter 5). Idioms also occur within text-final sentences, to wrap up the tone and informational content of the preceding text (Examples (7)–(9)). Again here, variants are common (as in Example (9)). In both these textual positions, idioms do not simply carry informational meaning: they also perform an evaluative, situational and/or modalising function.

(4) <h> **SEEING RED** </h> When we first moved into our country cottage the front door was painted in a bright red colour inside and outside-and I hated it. It clashed with the soft, natural tones of the stone walls, and I associate red with aggression

(5) <h> **GREEN WITH ENVY** </h> <p> Green tea seems to be the buzzword in skincare products these days. It's absolutely crammed full of antioxidants which destroy the free radicals that ravage the skin – why else do you think the Chinese always look so good?

(6) <h> **In the pink** </h> <p> Claims that a ban on colour additives in bird feed would turn flamingos from pink to grey were fiercely denied by Eurocrats yesterday.

(7) Their faith has moved on little from the spirit-worship of their ancestors. And yet, only two weeks ago the General Synod of the Church of England approved a new service of exorcism, to be used for casting out demons and evil spirits. **Pot and kettle**, or what? <p> Keith Porteous Wood <p> General Secretary <p> National Secular Society <p> London WC1.

(8) "The last baby born in our cottage was in 1906. Everyone's **tickled pink**." <p> <h

(9) He says the 200-year-old inn could not cope with the crowds in the area, and is planning a private corporate function instead. One thing is for sure-the critics who said he was crazy to buy the pub must be **Green-wich with envy.** <hl>

The features of idioms and other types of figurative language in headlines have been well documented (see, for example, De Knop 1985; Partington 1998; Herrera Soler & White 2007). As far as final sentences are concerned, the picture is less clear.[16] There may be contextually-pertinent modification, as in Example (9), or the idioms may merely be evoked by the presence of their core collocates (Example (7)), but these comments are based only on the observations of the few instances found in the data set. Without in-depth studies to refer to, it is

---

16. Though using journalistic texts in teaching has heightened my own awareness of this function, virtually nothing has been published on figurative language in final sentences. I am only aware of Lama's (1996) study as reported in Partington (1998: 140–141).

difficult to assess how important idiomatic language is as a summative device. That said, it is clear that final-sentence idioms are not restricted to journalism, and readers will undoubtedly have come across instances in academic text: of the heap of books and articles sitting on my desk at the moment of writing, I found Examples (10) and (11).[17]

(10) Or I could **take it to the cleaners**, though others may do that for me.

(11) "**Out of the mouth of babes**", but that is idiom, not resonance.

Although the evaluation expressed by the idiom within its unit of meaning may be perceived as an inherent part of its meaning, it is contextually determined *in praesentia* (i.e. in the cotext of the instantiation) or *in absentia* (i.e. via remembered cotexts). The resonance of remembered instantiations normally reinforces the evaluation suggested in the text. If, however, the instantiation occurs in a markedly different kind of context, a clash occurs between the remembered meaning and the meaning which seems to be being conveyed. This clash can give rise to effects such as irony, sarcasm and humour, or it may simply give the impression of there being something not quite right with the meaning. This latter "effect" can lead to the perceptions of insincerity which Louw discusses (1993). But special effects no longer belong properly within evaluation: they are part of the wider notion of connotation.

3.6.2   A closer look at connotation

Connotation is a fascinating area of study, much of its attraction lying in its intangibility. Connotations flicker on and off in language, and that makes them extremely difficult to pin down. As a general rule, connotations can be split into a number of distinct groupings, the broadest of these being cultural connotations and associative connotations. Although readers may be more familiar with Partington's tripartite division of social/situational, cultural, and expressive connotations (1998: 65–66), I will follow Kerbrat-Orecchioni's (1977) classification as it treats the subject in much more depth, subdividing connotation into a series of classes and types. She identifies five broad types of connotation, of which the first, *connotations dont le signifie est de même nature, mais non de même statut, que le signifie de dénotation*, is brushed aside as being of little interest to the present argument (it concerns onomatopoeia, phonaesthetic meaning [Firth 1935], and suchlike). The remaining classes are *connotations stylistiques* (stylistic connotations), *connotations énonciatives* (expressive connotations),

---

17. Sources: Ex. 10: Stewart (2009: 165); Ex. 11: Williams (2008: 173).

les connotations comme valeurs associées (associative connotations), and *signi-fications implicites comme valeurs connotées* (implicit meaning as connotation). Each of these can be activated *in absentia* or *in praesentia*; in other words, it can be absent yet suggested along the paradigmatic axis, or it can be present along the syntagmatic axis. If connotation is activated *in absentia*, that meaning is superimposed on top of the basic, referential meaning of the word(s) in question (ibid.: 123), leading to semantic enrichment. The appreciation of the ensuing double-meaning is what Partington has called the "smugness effect" (1998: 140), as the reader feels self-satisfied in having successfully understood both the referential meaning and the allusive, connotative one. When, on the other hand, connotative meanings are activated *in praesentia*, they reinforce existing connections between connoted and referential elements and contribute to stylistic or rhetorical effects within text (Kerbrat-Orecchioni 1977: 123; see also examples provided in Table 3.5).

Kerbrat-Orecchioni's stylistic and expressive connotations are close to Partington's (1998) social/situational connotation, whereby words can be said to reveal information about social class, regional origin, age, sex, and speaker relationships (ibid.: 65). What is added, however, is evaluation: Kerbrat-Orecchioni recognises two aspects of expressive meaning, one which is emotional, the other, judgemental. While affective expressive connotations use stress, intonation and word order to communicate additional meaning, or indeed choose lexis on the basis of its emotive value (e.g. *home* rather than *house*), evaluative expressive connotations reveal the speaker's favourable or unfavourable inclination towards the object of expression, and overlap to a degree with affective connotations. The reason for recognising them as separate sub-types is that it is quite possible to have an emotional attitude with respect to an object to which one brings no value judgement, and conversely, regard without emotion an object which one is judging (ibid.: 110).[18] Consider the following extended excerpts from *bleed sb white* (cf. Figure 3.9, above).[19]

(12)　I think poor old Dido and all those ghastly offspring **bled him white.**

(13)　In short, not content with **bleeding the public white**, they want us to do the work.

18.　'…on peut fort bien avoir une attitude émotionnelle vis-à-vis d'un objet sur lequel on ne porte aucun jugement de valeur, et inversement, envisager sans émotion un objet que l'on juge' (Kerbrat-Orecchioni 1977: 104).

19.　Source: Bank of English 450m: Ex. 12: British Books; Ex. 13: Today.

In general terms, the first of these is primarily expressive, the second primarily evaluative. But the distinction is a fuzzy one, especially when we are judging other people's textual output. We can notice explicit lexical choices which have an emotive content (*poor old* Dido, *ghastly*) and compare these with others which are essentially evaluative (*not content with*); but when we imagine the tone of voice, intonation and stress patterns in both examples, we are hard put to find an evaluative connotation devoid of emotion, or indeed an emotional connotation devoid of judgement. So while these are separate in theory, in use they tend to converge, hence their grouping under expressive connotations, which has strong parallels with the features of evaluative meaning covered in §3.6.1.

Kerbrat-Orecchioni's fourth class is associative connotations. What unites all instances of associative connotation is that they do not call upon institutionalised meanings, and while all add extra layers to the information content of as expression, those layers are not always present for the same individual, let alone for the speech community as a whole. This is why connotations are often defined as "occasional" meanings (Bussmann 1996: 96).

Kerbrat-Orecchioni identifies six types of associative connotations. Meaning can be transferred by lexical analogy which involves phonically/graphically identical or similar word forms (homophones and near-homophones, homonyms and near-homonyms). Additionally, paronyms (words with the same root) can be involved: there is a special case of paronym in which the root is taken not to be lexical but cultural and linguistic instead, thus in the example in Table 3.5, *schwartz* denotes Germans and Germanness, not just the colour (see §7.2.1 for further discussion of this variant). The second subtype is one in which semantically-related words have lexically-related forms, e.g. *cherry* and *berry*. Semantic affinity, the third class, encompasses the classic hierarchical relations of lexical semantics: superordinates and their hyponyms, synonyms and near-synonyms, and antonyms. Thus *red in the face* can be downtoned to *pink in the face*, *purple prose* can be contrasted with *grey prose*, and so on (see §6.3).

Combinatorial affinity is an interesting class, and one that Kerbrat-Orecchioni makes little comment about, except to say that it could be called collocational connotation instead, and that it is similar to intertextual connotation (ibid.: 118). The essence is that a word can be connoted when its usual collocates are present, even if the word itself is absent. This is a concept that has already been discussed with reference to evaluative meaning and loaded lexis (§3.6.1), and is at the heart of the identification of irony and humour, as presented in Louw (1993). The habitual collocates lead us to expect a particular lexical item (they *predict* that item), but if another is used in its place – especially if it contrasts semantically with it – what psycholinguists call "dual coding" ensues. Collocations are 'mutually expectant and mutually prehended' (Firth 1950: 181), so when a "false" collocate intrudes

on a pattern, it is interpreted within the string but the absent collocate is also triggered, or connoted (Louw 2000a: 48; see §3.7.1).

Continuing from collocational connotation, the next class is connotations deriving from previous use (in text or in the world). Wider-ranging than collocational connotations, this category covers the wealth of personal and cultural experience of a word or its referent, and includes intertextuality, citations, extra-linguistic experience and knowledge of the world. Being less text-bound than collocational connotations, its influence is more difficult to ascertain. Kerbrat-Orecchioni suggests that the plethora of film titles following the collocational framework *the X, the Y and the Z*, where the slots are filled by nominalised adjectives, recall the "superordinate" or "typical" member of the class, Sergio Leone's *The Good, the Bad, and the Ugly*. Whether this is true or not is hard to say – connotation is, after all, as much a private a matter as it is a public one. However, of the 106 instances of the completed collocational framework in the BNC, eighteen of the twenty-nine that start with *the good…* are followed by *bad* then *ugly* (see Figure 3.11). Yet to suggest that all instances of the string *the X, the Y and the Z* are likely to connote the film title is something of an exaggeration, especially considering that the next most frequent collocates are *old*, *poor* and *sick* (12 each, though their sequencing

```
nister. So we're looking for the good, the average, and the unexceptional as well
formation, watch this space. THE GOOD, THE BAD, AND THE AWFUL Robin McMillan find
        Backstage by Alex Finer THE GOOD, THE BAD AND PETER COOK An East End doctor
ide, brandishing examples of the good, the bad, and the typographically ugly. Row
rs. Election 1992: Campaign: the good, the bad and the seldom seen By Jonathan Pe
mine our policy. John Major. The good, the bad and the sexy By Jaci Stephen WITH
at the Manor were treated to the good, the bad and the ugly side of Oxford United
 scenery" (but no mention of The Good, The Bad and The Ugly and "exotic cuisine"
Wayne dies broke and drunk. THE GOOD, THE BAD AND THE UGLY STRANGELY, THE climax
Few Dollars More (1965) and The Good, The Bad and The Ugly (1966) and immediatel
             news review the good, the bad and the ugly The petrochemical ind
a South American version of The Good, the Bad and the Ugly writ large. From anot
tti western. Which ones were the good, the bad and the ugly? And, of course, ther
  y… ugly? Honestly, Demian, The Good, the Bad and the Ugly … hooray. Well done
, For A Few Dollars More and The Good, The Bad And The Ugly ) and the mellower bu
n autobiographical litany of the good, the bad and the ugly from Lawrence's teena
st punk types. And Iggy Pop. The Good, The Bad And The Ugly "Der ner, ner, ner, n
emocracy – a society full of the good, the bad and the ugly. On the plus side are
 ape. "They could have filmed The Good, The Bad And The Ugly here" I quipped. They
ate is a "classic example of the good, the bad and the ugly", Doug Campbell, depu
cs as "Lawrence of Arabia", "The Good, the Bad and the Ugly", and many more.
is year's shows are entitled The Good, The Bad And The Ugly and take the form of
ground or history. It is for the good, the bad and the ugly; the religious, the p
elist Preston Parish Church) THE GOOD, THE BAD AND THE UGLY A two week Youths Art
nces. One, the good, bad the the good, the bad and the unacceptable. Mhm. One cal
for British ELT publishers – the good, the promising and the problematic, plus Jo
ngers of a movement" Saldaan THE GOOD, THE RAD & THE UGLY Trainers were not made
 a direction of intent upon "the good, the true, and the beautiful" in whatever c
ight, but because he praised the good, the true and the beautiful in useless sple
```

Figure 3.11 *The* good, *the Y and the Z*

Table 3.5 Associative connotations by type

Type	Involving	Examples from data
Lexical analogy	Homophones;	Subtitled *Greene* with Envy, this episode sees the debut of Alan Alda as Dr Gabriel Greene.
	Homonyms;	YOU *BROWN* NOSER. Wee Craig *Broon** has been at his very best this past few weeks…
	Polysemy;	Nor did Bright see *red* when he elbowed Linighan in the face. A yellow card was all he got.
	Paronyms, including cultural roots	Surely a case of the pot calling the kettle *schwarz*.
Signifier-signified	Similar word-similar referent	Red *rag* to a bull /red *flag* before a bull
Semantic affinity	Synonyms and near-synonyms;	Green *fingers* / green *thumb*
	Antonyms;	Mr Dobson's, which arrived a day earlier, was a plain photostat of four sides of *grey* prose.
	Superordinates and hyponyms	I can see her now – *pink* in the face with the exertion
Combinatorial affinity	Habitual collocations	… for the legion of small, mostly American, PC makers struggling to *stem the flow of red ink*.
Previous use	Intertextuality, citations;	*the X, the Y and the Z*: the good, the bad and the ugly
	Extra-linguistic experience and knowledge	… you'll be happy to *swallow* one red herring after another
Referential relationship	Symbolic meanings	*Labour* are *socialists red* in tooth and claw

* Scottish pronunciation of *brown* is *broon*: here, Craig "Broon" is the footballer Craig Brown.

is not fixed). It is even unclear whether the completed collocational framework has to refer to the film: the lack of noun capitalisation indicates that the phrase has evidently become institutionalised. One thing is evident, however: if the string *the X, the Y and the Z* starts with *the good…* the chances are that the next lexical collocate will be *bad*.

The last of the associative connotations is referential relationship, which ties symbolic meaning to language, as when red roses connote romance. This kind of connotative meaning ostensibly lies at the base of most of the colour-word idioms and metaphorical collocations in the data set, but is not usually activated in canonical forms, as will be demonstrated in Chapter 4. The different kinds of associative connotations discussed by Kerbrat-Orecchioni (1977: 112–122) are summarised in Table 3.5 together with examples of how they have been realised in the data set under study. Many of these examples will be discussed in later chapters.

2nd proofs

As can be seen from the examples chosen to illustrate each of these kinds of associative connotation, we are dealing not with conventional language use but rather with creative instances. The meanings inferred from the variants are occasional, usually have a contextualising function, and serve to add layers and nuances to the meanings normally expressed by the idioms and metaphorical collocations. Of course, since connotations like these are not institutionalised, they are interpreted in the light of their context. What must be made absolutely clear is that there is no way of ascertaining if the meanings identified (con)textually are the only connotative meanings implied by the writer and/or inferred by the reader. These are unknown and unknowable quantities. Contextual analysis allows shared meanings to be identified, and it is only these secondary meanings which will be taken into consideration in this study.

Kerbrat-Orecchioni's fifth and final category of connotation is implicit (or inferred) meaning as connotation, a category she herself describes as problematic (ibid.: 163). This category effectively covers conversational implicatures (Grice 1989), making it a feature not of words but of utterances, and therefore related to "discourse function" semantic prosody. As such, it will not be discussed here nor will it be considered a form of connotation proper, following the decision to treat connotation as a word-based phenomenon, not an utterance-based one.

3.6.3   Cultural connotation

To end our discussion of connotation, a final aspect needs to be noted: cultural connotation. In Chapter 2 we saw that colour-word idioms and metaphorical collocations embody real-world knowledge of colours, for example that blood is red, the sea and sky are blue, grass and plant-life are green, and so on. However, while these traits can be said to be universally true, others are culture-bound: *blue* is associated with depression in English (*have the blue devils*, *have the blues*, *feel blue*, *the baby blues*, and *blues* as a musical genre) but not in Italian; and *green* is the colour of jealousy for an English speaker (*the green-eyed monster*, *green with envy*) while *yellow* is preferred in many other European languages (e.g. *giallo d'invidia* [It.], *gelb vor Neid* [Ger.]).

Although colour-word idioms can easily be grouped by connotation (see Appendix; see also Philip 2006), these groupings are only very vague pointers to the connotative meanings that the expressions actually convey, because the linguistic expressions are pegged to symbolic meanings which may no longer be current (cf. §2.3.1) This point cannot be stressed enough: the purported origins of a phrase may have no bearing whatsoever on the perceived meaning it expresses in current use. Unless speakers have had some reason to learn of the phrase's

historical meanings, it is the current meaning and only the current meaning that they will associate with the colour.

Cultural connotation is not discussed explicitly by Kerbrat-Orecchioni, but is mentioned by many scholars of figurative language. It is rare for corpus linguists to venture into the realm of connotation because of its traditional links to meaning in the mind. One of the few scholars who does tells us that cultural connotation concerns 'what a lexical item denotes within a culture' (Partington 1998: 66). However, idiom scholars outside the realm of corpus linguistics have been more forthright in noting the cultural importance that idioms have, in particular because they fix cultural concepts in the language. But culture is a vast area, and cultural connotation perhaps too vague a term.

Cultural connotations in idioms are most typically seen to emerge from their wordings. Idioms capture aspects of culture, whether those be culture-specific concepts, culturally significant events, or value judgements. Gläser (1998) points out that idioms typically express connotations, qualifying her statement by defining connotations as 'additional semantic markers which are associated with the value judgements of a speech community (i.e. a class of social group) or of an individual speaker or writer' (Gläser 1998: 128). In other words, they are stylistic and expressive connotations which have a culture-specific slant. These connotations are institutionalised, and are encoded in dictionaries: expressive connotations are signalled by usage labels including "derogatory", "taboo", "euphemistic" and so on, while stylistic connotations are indicated by style markers such as "formal", "colloquial", and "foreign" (ibid.: 128–129).

Due to their conventionality, these institutionalised connotations might not strike a (native) speaker as conveying particular cultural information. Connotations that signal culture more explicitly are those whose lexis is culturally significant, for instance by featuring allusions to historical events and personages, or to objects and institutions that characterise the culture; and idioms, proverbs, and other stable lexical phrases are viewed as a valuable source of cultural information because cultural connotations can be inferred from lexical collocations (Telija et al. 1998). In this way, idioms are a form of cultural record, linking past and present through shared connotative values. But the cultural information is not merely statement of fact: there is emotion attached. Telija & Doroshenko (1992) discuss the complexity of idiomatic meaning as arising in part from the emotional-expressive dimension of cultural connotation, and note that this is particularly problematic cross-linguistically. Although it is institutionalised, emotional meaning (the emotional effect of the expression on the speaker/hearer) is not explicit in the surface wording of an idiom and can therefore be easily lost in translation.

> Most idioms are deeply rooted in the historical and cultural development of lin-
> guistic creativity, and understanding idioms requires not only the knowledge of
> semantics and connotations, but also the skill of finding an equivalent or analo-
> gous idiom in one's native tongue which would retain the emotive impact.
>
> (Telija & Doroshenko 1992: 438)

Thus idioms do not only carry information but, most significantly, emotive force (i.e. cultural connotation as a type of expressive connotation). Piirainen too finds that a good deal of culture-specific meaning is embodied in idioms and again notes the pragmatic importance of 'the cultural dimension [which] becomes ap-parent at the level of actual meaning' (Piirainen 2008: 213). "Actual meaning" is the communicative intent or discourse function of the idiom rather than its sur-face meaning (Baranov & Dobrovol'skij 1996), so we have the suggestion here that connotative meaning is not likely to emerge by connecting, say, *between the devil and the deep blue sea* with connotations associated with demons and with the ocean in Anglophone culture. Instead we must look to the expressive meaning of the idiom (its emotive force) and its discourse function in order to understand what cultural connotations, if any, are attached to it.

By now it should have become apparent that some types of secondary mean-ings are common to a language community while others are individual and per-sonal. Shared connotations may vary from one social group or culture to another, but, being institutionalised, we can also expect them to be coded in the language. As with any other coded linguistic feature, information regarding shared conno-tations should emerge in the cotext accompanying the words they are convention-ally associated with. As a consequence, although it is true to say that connotative meaning is not inherent in the referential meaning of a lexical item, this does not necessary exclude it from belonging to the semantic and pragmatic meaning of that lexical item in a given language (Garza-Cuarón 1991: 220–221). However, not all connotations are shared, and no claim can ever be made that readers always "get" the allusions that are suggested in the text. Corpus data can reveal patterns and demonstrate the ways in which 'deviation from the norm' consisting of '"ab-normal" meanings or [a] low degree of acceptability' (Garza-Cuarón 1991: 214) can indicate the presence of connotative meanings, but it cannot make any claim about the comprehension of connotation, which is a matter for psychology and is, in any case, highly individual (Steen 1994: 28).

3.6.4   Metaphor within semantic association

In the earlier discussion regarding *graft* and the other loaded lexis in the cotext of *bleed sb white* (§3.6.2), evaluation in positive and negative terms was seen to arise

out of a generalisation of the semantic traits of all the collocates in the set of con-cordances. This kind of generalisation does not emerge out of formal semantics, but is a feature of metaphor, in particular of how metaphors work.

There are many accounts of how metaphorical meaning is understood, but virtually all modern accounts view it as a conceptual process: attributes of the source or vehicle term (the actual word which is being used metaphorically) are "mapped" onto the target or topic (the thing being referred to metaphorically).[20] In this way, metaphors 'characterize a topic of interest in terms of something else' (Glucksberg 1995: 48), but the characterisation need not be made explicit. Some-times a metaphor features lexical items from both target and source domain, other times there is no target term. For example, *he **shot down** all of my **arguments*** and "*fire away!*" are both possible instantiations of the ARGUMENT IS WAR meta-phor. The former can be referred to as a "metaphorical pattern" (Stefanowitsch 2006: 66), because both source and target are explicit, while the latter remains simply a bog-standard metaphor (there is no explicit target).

A word is recognised as being metaphorical because its context (the meta-phor "frame", Black 1962) imposes an interpretation which is different from the most salient or basic meaning that the word is believed to have (cf. §2.1.2). In a metaphorical pattern, the target term makes the frame explicit, but, as with all metaphors, further clues will be present in the surrounding context. This tells us what a metaphor is, but does not tell us how metaphors work.

Broadly speaking, there are two ways of looking at the matter. The first, Aris-totelian in origin, is that metaphors are implicit similes, i.e. they make compari-sons without the presence of a comparative marker in the syntax. Viewed in this light, the difference between an explicit simile (Example (14)) and a metaphor (Example (15)) is determined in relationship to the truthfulness of the proposi-tion (cf. §2.1.4).[21]

(14)   You are just creating a bank that is useless, **like a white elephant**.

(15)   The NHS will be left with an extremely large **white elephant** on its hands if it goes ahead with this network, which is based on the wrong model.

---

20. *Source/target* is the preferred terminology when speaking of conceptual metaphor (Lakoff & Johnson 1980; Lakoff 1987b), i.e. when speaking of metaphor as a conceptual process; *ve-hicle/topic* (after Richards 1936) is preferred when discussing actual instantiations of metaphor in text, often referred to as "linguistic metaphor" (Deignan 1999) to avoid confusion with con-ceptual metaphor.

21.   Source: WordBanks Online collocations sampler.

2nd proofs

The second approach to metaphor interpretation derives from Black's (1962) interaction view, which lies at the heart of recent work on conceptual blends (Fauconnier & Turner 1998; Coulson 2000), and much psychological work on figurative language. In a nutshell, mutually-relevant features of topic and vehicle terms acquire prominence, while unshared features are downplayed. So unlike comparison, where the topic term picks and chooses particular features of the source term, here the activated features allow both terms to interact by highlighting shared features.

This is relevant to evaluation when viewed through Glucksberg & Keysar's (1993) "class-inclusion view" (see also Glucksberg & McGlone 1999). We saw in §3.6.1 that the loaded vocabulary in the cotext of *bleed sb white* expresses different pragmatic functions (e.g. *graft* = disgruntlement and unfairness; *ghastly offspring* = expressive connotation of unpleasantness, and so on). Yet, in spite of their differences, these can be grouped into a class of negatively-evaluated lexis. How this comes about is through the selection of relevant attributes of these loaded terms, in much the same way as metaphorical meaning is created.

In the class inclusion view of metaphor comprehension, no comparison is made between the topic and vehicle; instead, the traits they share unite them in a new, ad hoc class. What is important to note is that the relevant features are not necessarily visual or otherwise related to traditional semantic classes like hyponymy and synonymy, but are usually attributive (Glucksberg & Keysar 1993: 408–409).

> According to our attributive categorization view, a metaphor vehicle, in the context of a specific metaphor topic, acts as a cue for the speaker to infer or construct a relevant category to which both vehicle and topic belong, with the following important constraints. The metaphor vehicle must, to some degree, epitomize or symbolize that category. The metaphor topic, by virtue of being assigned to that category, is characterized along one or more relevant dimensions.
>
> (Glucksberg & McGlone 1999: 1546)

This is similar to the view of meaning in collocations articulated by Sinclair (Sinclair & Teubert 2004: xxi; cf. §3.3.1), and is almost certainly how a negative evaluation eventually comes about. When evaluation is derived from the semantic preference of a node, it tends to be the case that many different word forms are contributing to that evaluation. Each of those words has a specific (discourse function) semantic prosody (see §3.7), but they can be grouped together in an evaluative class on the basis of shared aspects of those semantic prosodies. The more heterogeneous the sum of the semantic prosodies is, the less specific the shared attributes can be, and generalisations are made. Sometimes all that is left is polarity. Thus evaluation, like metaphor in the class-inclusion view, can be said to be an attributive assertion (Glucksberg 2003: 95).

3.7   Semantic prosody (ii): Discourse and pragmatic functions

The previous section unpacked the many subtypes of meaning that are commonly referred to as connotations. Semantic prosody, while often likened to connotation, operates on a different plane. In the first place, speakers are aware of institutionalised connotative meanings, and use this awareness when choosing to use one particular word instead of another one. In this way, expressive nuances can be brought to a text, including evaluative, culturally-pertinent or any of the other kinds of connotative meanings that have just been outlined. Semantic prosody, on the other hand, while perceived subliminally, is not available to our intuitions (Louw 2000a), and therefore is not chosen consciously as a means of conveying shades of meaning.

One of the reasons why semantic prosody is not accessible to introspection but must rather be identified in (large amounts of) corpus data is that it is not attached to discrete words, but to discourse units (units of meaning). In fact, the same word can be involved in a range of distinct semantic prosodies, depending on its collocates, colligates, and semantic preferences. This will become very clear in later chapters, where colour words like *red*, *blue* and *black* are seen to attract totally unrelated prosodies depending on their contextual environments.

The second reason why semantic prosodies are largely uncontrollable is that they come before linguistic choices are made. The semantic prosody of an utterance is something like a pre-lexical meaning, a gut reaction or reflex, which is then verbalised.

> The semantic prosody of an item is the reason why it is chosen, over and above the semantic preferences that also characterize it.   (Sinclair 1998: 19)

The best way of controlling an impulse is simply to suppress it, not to modify it in mid-stream, hence the widely held maxim that it is better to say nothing at all than to open one's mouth and regret it: it is infelicitous or inappropriate comments that *slip out* (see Figure 3.12),[22] inadvertently reflecting what the speaker is feeling. Not to put too fine a point on it, we could risk saying that while semantic associations come from the head, semantic prosody essentially comes from the heart, which is why it so difficult to manipulate.

---

22. Source: British National Corpus: 1 in 2 relevant concordance lines shown.

```
He hadn't meant to say that. It just slipped out. Eleanor's warm, inviting voice
    On those rare occasions when it slipped out accidentally, her rage was a
  was remarkable how easily the lies slipped out once you got started. Rose was
 guessed yet, darling? (Sorry! that slipped out.) Really, it's rather exciting.
  not yet. The last words had simply slipped out. Then, recovering her composure,
          I asked. "Did I? Must have slipped out. When I was a kid, in Hull, my
     " Galvone said hastily. "It just slipped out. I was so stunned you'd managed
 Sorry about the language. Inept. It slipped out." "What do you mean –
      and Timothy Gedge letting the facts slip out because he didn't care, because
  She couldn't prevent the question slipping out. "What's in Athens that's so
      meant to ask, the words had just slipped out. He shook his head gently. "Not
 on with Portia Forbes?" The question slipped out before I gave myself time to
        "Why?" The suspicious query slipped out before she could stop it. Isabel
            Clive…" The name seemed to slip out accidentally, and before Doreen
 next door," Lucy quipped, the words slipping out before she could give them
  "Sorry. It just slipped out…" "Just slipped out, that's a good'n … ah" Rex
         The observation had simply slipped out. "This way," he was instruct
    and obsession, and the words just slipped out. But it was absolutely the wrong
```

Figure 3.12 *Slip out* (words, emotions)

3.7.1   Semantic prosody and the context of situation

In the earlier discussion of colligation and semantic preference, it was seen that the collocational patterns clustering around a node reveal elements of the context of situation, in particular the participants, their actions, and the relevant objects. The third element of Firth's (1950) context of situation is 'the effect of the verbal action' (ibid.: 182), and this is largely accounted for by semantic prosody.

> A context of situation for linguistic work brings into relation the following cat-
> egories:
> A. The relevant features of participants: persons, personalities.
>   (i)   The verbal action of the participants.
>   (ii) The non-verbal action of the participants.
> B. The relevant objects.
> C. The effect of the verbal action.   (Firth 1950: 182)

The semantic prosody of an utterance has to be identified relative to its collocates, colligates and semantic preference, not in isolation. Unlike semantic associations such as connotation and evaluation, it is integral to the context of situation, only becoming identifiable in actual contexts of use: *slip out* in itself has no semantic prosody, though it may connote smooth movement of small things, for instance. Its semantic prosody is revealed after the other aspects of the context of situation are in place. So when it is used of a person moving voluntarily away from a place or group of people, and seeking to do so without attracting attention (the more common meaning of *slip out*), the semantic prosody can be said to express something like "resolution of conflict between obligation to be present and wish

to be elsewhere" in relation to a range of social norms (that one should be in bed at night-time; that one should sit through meetings attentively, etc.), and is connected to non-verbal action (through prepositions of movement) and an absence of verbal action. This sense is probably evaluated positively. When used of words which "escape" into the real world where they are noticed and may even cause damage, *slip out* may be evaluated negatively (or carry negative connotations, or have negative semantic associations), but this of course depends on the relationship between the speaker of the words, the words said, and their effect as viewed by the person who refers to the event with *slip out*. Yet the semantic prosody need not bring any judgement to bear: it expresses the "inadvertent revelation of feelings or information, in contexts where doing do may be harmful or damaging", and is more often used of verbal action in the form of the words uttered, than of non-verbal action.

Semantic prosody is not something that is normally noticed within any single instantiation of a pattern. It is 'a collocational phenomenon and one which is preferably to be regarded as recoverable computationally from large language corpora rather than intuitively' (Louw 2000a: 48). Intuitions come in when the semantic prosody is violated, making it come to our attention. This happens when its preferred lexical realisation is altered, as for example in *There is a beach in Vietnam where the sand is whiter than driven cocaine* (cf. Table 3.5 above). The appearance of a word (*cocaine*) which differs dramatically from the expected collocate (*snow*) creates what Louw (ibid. 2007) calls a "fractured context of situation" which forces reinterpretation of the utterance. Sometimes the meaning is understood to be ironic, sometimes humorous, sometimes insincere, depending on how the rest of the discourse context moulds our interpretation. The "fracture" within the context of situation occurs because an intruder is mingling with the normal participants and/or objects (as realised in the collocates). In the example above, the expected form *whiter than (the) (driven) snow* with its associations of spiritual and moral purity, and semantic prosody of willing submission to authority (usually God) is incompatible with the associations surrounding *cocaine* and indeed associations regarding beach tourism in the far East.[23] As a result, the lexical form clashes with the semantic prosody and the speaker/writer's communicative intentions have to be reassessed.

---

23. The semantic prosody can only be ascertained in combination with *as white as snow* since there are only two occurrences of the canonical form *whiter than (driven) snow* in the data consulted (Bank of English 450m version). The semantic associations are strong: both are biblical in origin and both refer to spiritual purification.

3.7.2   Semantic prosody and idioms

That idioms and other types of figurative language have a pragmatic function is often remarked upon both within and outside corpus linguistics (see for instance Telija & Doroshenko 1992; Čermák 2001; Deignan 2010). It is that pragmatic function, the semantic prosody, which underlies the decision to use an idiom in the first place:

> Stylistically marked idioms … are also pragmatically loaded. Emotiveness along with rational evaluation and stylistic markedness create the pragmatic bloc of the idiom meaning. To choose a certain idiom is to perform a speech act. The awareness of the idiom motivation is at the basis of such a choice.
>
> (Telija & Doroshenko 1992: 438)

Idiom motivation refers to the complex conceptual knowledge commonly referred to as image schemata (Lakoff 1987a) or metaphorical scenarios (Musolff 2006). These bring together the presumed literal meaning of the idiomatic string in its supposed original context of use, the implications that it apparently had, and a mental representation of that context of situation (the participants, their actions, the relevant objects, and effect of the verbal action). The hedging here is intentional. Motivation is a personal matter, and etymological accounts of the origins of idioms, however well-researched, can do little more than record the expression's earliest appearance in print, which is no guarantee of its original meaning in speech. Unlike motivation, the pragmatic meaning of an idiom is fully institutionalised. The choice of an idiom reflects that pragmatic choice, and although it is probably fair to say that the surface wording can also make a contribution, this is as likely to be in the form of avoidance of a particular idiom (when the words are contextually inappropriate or irrelevant) than its actual use.

In recognition of the pragmatic force of figurative language, particularly as identified through corpus data, metaphor scholars are now talking of "metaphoremes" (Cameron & Deignan 2006). These are essentially units of meaning whose core is a metaphorically-used word: they feature 'specific lexical and grammatical form' combined with 'specific conceptual content and with specific affective value and pragmatics' (ibid.: 274). The parallels with collocation, colligation, semantic preference, semantic associations and semantic prosody are evident. The only important point to note is that the semantic associations of a metaphoreme, ascertained by grouping together semantically-related collocates, feed into the conceptual metaphor to which the metaphor belongs, rather than including evaluations and other connotations.

Within this framework, it is not the metaphor itself but the metaphoreme which is said to express both affective and pragmatic meanings. Only when used in context is it possible to determine which sense of the core lexical item is intended, and the pragmatic meanings relate to that particular sense. Although they do not use terms familiar to corpus linguists, Cameron & Deignan explain that of the two distinct senses of the metaphor *walk away from*, one carries a semantic prosody (for wont of another term) "unwillingness to confront a moral obligation" (ibid.: 683), and another "surprise and relief" (at having survived an accident unscathed) (ibid.: 685). Neither of these pragmatic meanings is inherent in the purely semantic meaning of *walk away from*; they are actualised through the particular cotextual patternings which accompany (and determine) one or other of the two senses.

3.8   Idioms and the idiom principle

This chapter has treated the idiom principle in considerable detail because it is of fundamental importance to the arguments put forward in the remainder of this book. Analysis of idioms all too often concentrates on decontextualised word-strings, often with the implication that their meaning exists independently of context and even that context *interferes* with the meaning they convey (see Temple & Honeck 1999). Nothing could be further from the truth: idioms are the paragon of the idiom principle, and the features of the unit of meaning are all absolutely essential elements of idiomatic meaning. The same may be less true of other phrases. For example, Philip (2009) argues that semantic prosodies need not add an additional layer to the meaning of a phrase if that phrase is fully analysable semantically: in such cases 'the words alone are enough, so the semantic prosody is redundant' (Philip 2009: 1).

Collocations, colligations, semantic preferences and semantic prosodies contribute in subtle ways to the meanings that can be expressed in text. Institutionalised phraseology fixes meaning to word forms, but we have seen that idioms have a peculiar relationship to fixity. While typically identified by a core collocation, their actual realisation in text may vary considerably, and canonical forms of phrases are rather elusive. Yet much of idioms' semantic integrity comes from their external phraseology. Idioms are phrases within phrases, thus a core collocation itself attracts collocates in the cotext, assumes particular syntactic positioning and/or function, and is restricted in use to particular semantic fields and the expression of 'an attitudinal, evaluative or emotional stance with

2nd proofs

regard to a particular context or scenario and its outcome (anticipated or actual)' (Philip 2009: 3).

The next chapter investigates the natural consequence of phrasal fixity, delexicalisation, and how it relates to beliefs regarding the meaning of the idioms and their component words.

Words in usual collocations

Delexicalisation

This chapter explores delexicalisation, which has been referred to as the 'necessary correlate of co-selection' (Sinclair 1992: 16). As words crystallise into collocations and those collocations attract stable cotextual patternings, the meaning potential of each of the component words is severely restricted. Just as 'a word in a usual collocation stares you in the face just as it is' (Firth 1957: 182), so too the established meaning of the combination overrides the other possible meanings that could be intended, but are not. Because they are highly salient and have an unambiguous "basic" meaning-referent, colour words are ideal for testing the strength of the "meanings have words" approach. The case studies and discussion in this chapter should be of particular interest to translators and foreign language teachers. It is demonstrated not only that the meaning of a phrase is fixed by its conventional lexical realisation, but also – and counterintuitively – that the salient meanings of its components are largely dormant.

4.1   Meanings that come out in the wash

> Normal text is largely delexicalized, and appears to be formed by the exercise of the idiom principle, with occasional switching to the open-choice principle.
>
> (Sinclair 1991: 113)

Delexicalisation refers to the gradual weakening of meaning that occurs when words are habitually used in combination. It is much the same as "semantic bleaching", the main difference being that delexicalisation covers the loosening of syntactic relations as well as the "washing out" of meaning.[1]

Strictly speaking, delexicalisation is a diachronic phenomenon and it manifests itself at particular stage in its development in synchronic corpus data.

---

1.   John Sinclair used to explain delexicalisation as doing to words what repeated washing does to a favourite pair of blue jeans: the form is unaffected by wear and washing over time, indeed it becomes more "comfortable" and familiar, but the intensity (of meaning/colour) gradually diminishes. (Workshop on *How to Use Text Corpora in Literary Text Analysis*, Tuscan Word Centre, June 2000.)

Through repeated co-selection, words coalesce into collocations and units of meaning. Once a unit of meaning has taken root, the potential meanings of its constituents are restricted and are subservient to the pragmatic function of the whole unit. Louw (2000b) observes that delexicalisation operates along a cline, with semantically empty words at one end, and maximally salient words at the other. What makes words delexical rather than salient, he argues, is their collocational behaviour. A word occurring with very normal collocates is less able to exert its independent meaning potentials than it would be if its collocates were less normal. This happens because the overall meaning of an unusual word combination has to be pieced together from the sum of all the words' potential meanings, because a novel meaning, by nature of its not having previously been encountered, does not have a familiar collocational environment to refer back to. Novel word combinations, especially metaphors, illustrate how 'each word when used in a new context is a new word' (Firth 1951: 190). In order to understand this "new word", the reader is forced to resort to his or her knowledge of its possible meanings, and as Giora and co-workers have demonstrated through the Graded Salience Hypothesis (Giora 1997, 1999; Peleg et al. 2001), this comprehension process starts with the most salient meanings first and only addresses less salient meanings if the salient ones cannot provide a suitable interpretation.

In the case of contextual environments which are not new, the word and the meaning being conveyed cannot be new either. Collocational stability creates the ideal conditions for delexicalisation to thrive. Familiarity breeds contempt, it is said, and in delexicalisation the "contempt" is aimed squarely at salient meaning. A word relinquishes part of its autonomous identity when it occurs alongside its normal collocates, its meaning merging with theirs. What emerges is a phrasal piece, which can be associated with a unit of meaning. Since the defining feature of a unit of meaning is its semantic prosody, or pragmatic function, the inevitable consequence of delexicalisation is that word meanings take a back seat: words in their habitual combinations do not and need not realise their full meaning potential because the pragmatic function takes precedence.

Before returning to the examination of colour-word idioms and metaphorical collocations, delexicalisation will be briefly illustrated with reference to *categorically*, an adverb which is often encountered in both print and broadcast media, especially in the context of public statements.[2]

Corpus data reveals that *categorically* collocates significantly with three verbs: *deny*, *state*, and *say*. When *categorically* collocates with all these verbal processes, its primary function is to convey emphasis. Its full, lexical meaning (stating views

2.   Source: WordBanks Online collocations sampler.

clearly and firmly), on the other hand, is only partially active, if present at all. *Categorically + deny* (with variants such as *refuse*, *rule out*, *reject*) is associated with accusations of a defamatory nature, criminal or amoral acts allegedly committed in the past (Examples (1)–(3)). In this configuration (*categorically* with a semantic preference for denial), the collocation is read semi-compositionally: meaning is extracted from both words, but *categorically*'s main role in the collocation is for its intensifying function rather than its semantic contribution.

(1)   After the Time story was published, Depardieu **categorically denied** that he had been a rapist at nine, or at any other age.

(2)   The government has **categorically denied** reports thought to have emanated from radicals opposed to President Rafsanjani's attempts to bring Iran back into the world…

(3)   Both of them **deny categorically** knowing about or giving to The Sun or any other newspaper the names, addresses or telephone numbers of women alleged to have had affairs with Mr Ashdown.

Categorically + state (or *say*), on the other hand, is not associated with disproving allegations through the definiteness and firmness of fact. Instead it is associated with epistemic modality: it colligates with modals (*can*, *could*, *would*), hedges (*whether*, *if*), and negatives, and expresses future directions and intent – but nothing concrete – and no verifiable facts (Examples (4)–(6)).

(4)   We are *unable to* **state categorically** *whether there has or has not been* bullying by senior management.

(5)   One of the bankers **stated categorically** that his bank *would not be* associated with any action that *could* lead to a change in the political balance of the British press.

(6)   We **categorically guarantee** that no salesmen *will* call on you.

Compositionally, one can only *state* something *categorically* if it is known to be true, not if it is uncertain, hypothetical or yet to occur, so *categorically*, in this configuration (with a semantic preference for announcement of intent) does not convey the meaning "stating views clearly and firmly". It is delexicalised, its presence having come 'to fulfil a particular function but [having] no meaning apart from this to contribute to the phrase in which it occurs' (Partington 1993: 183). *Categorically stating* something therefore means "stating with intent to convince or placate", and does so because the words are 'mutually expected and mutually prehended' (Firth 1957: 181) in that particular collocation, expressing that particular meaning.

4.2   From origin to use

That colour words are highly salient is undeniable: they have an unambiguous referent and do not lend themselves at all well to polysemy except via motivated metaphor, metonymy and symbolism (see Niemeier 1998). In fact, colour words have a rather particular status in figurative language, because their semantic transparency is so great that they are almost obliged to conform to truth conditions. As a result, they regularly form metonymic relations (*see red*, *in the pink of health*, *green with envy*) but colour metaphors displaying no metonymical motivation whatsoever are very rare, if indeed possible at all in the normal run of language (ibid.). In fact, it is only in surrealist writing that we can find colour words being used to refer to the "wrong" referent, e.g. *La terre est bleue comme une orange* 'the earth is blue like an orange' (Élouard 1929). Although the oceans and the atmosphere make the earth appear blue from space, and the earth is like an orange in that it is spherical, oranges are never blue, the sea and atmosphere never orange. The colours *blue* and *orange* are opposites, or complementary, meaning that they stand at opposite ends on the colour spectrum. It is impossible for them to merge into one another, there is no such shade as "bluish orange", so the two colours cannot be alike in the way that orange and yellow, or blue and green are. Élouard's comparison goes further than a mere violation of truth conditions in that it introduces an impossible condition, and it is of course this feature which marks it as surrealist.

Figurative expressions which include colour words may however over-ride the limitations that salience and truth conditions place on the permissible extensions to colour meaning. While it is possible to trace the origins of *be green* meaning "immature", *caught red-handed* meaning "be discovered while committing a crime", and so on, it is not necessary to know of the supposedly literal or original meaning to use and understand expressions like these.

Salient meanings are much-debated in psycholinguist accounts of figurative language comprehension (see Gibbs 2001 for an overview). Some scholars believe that conventional expressions, like *caught red-handed* or *bolt from the blue* are so salient as chunks that the salient meanings of the component words, e.g. *red* and *blue*, can remain dormant when the expression is encountered in context (Gibbs 1994, 2002; Gibbs et al. 1989; Gibbs & Nyak 1991; Gibbs, Nyak & Cutting 1989). This Direct Access View, as its name suggests, advances the hypothesis that phrasal meanings are accessed directly, not broken down word for word first. The more familiar and apt the expression in the context provided, the more likely it is that only the phrasal meaning is activated (Blasko & Connine 1993). The Direct Access View would argue that *bolt from the blue* (Example (7)) is understood as meaning 'completely unexpected and surprising' with no activation of the colour *blue*.

(7)   "Actually to be remembered by the public! So good for the old ego, darling. I really am glad I decided to come."
– "Was it a surprise to be asked?"
"*Bolt from the blue*, darling. Hadn't heard from him for years, and didn't expect to. Knew he was standing, that was all."[3]

Other psycholinguists would argue that that the colour is (also) understood in its own right.

> According to the graded salience hypothesis, highly salient meanings should be processed initially. Among other things, this means that alternative interpretations of similar salience should be processed in parallel, whereas less salient meanings should be evoked after more salient meanings have been activated.
>
> (Giora & Fein 1999: 1602)

Many psycholinguists are unhappy about what context does to word meaning, intended in the sense of words and meanings as represented in the mind. It is clear that contextualised test items (words, phrases or complete utterances) are understood in contextually-relevant ways, while decontextualised test items are far more readily understood compositionally, and this can be construed as context interfering with meaning. It is also known that unfamiliar phrases can be understood when found in context even if their meaning is opaque (Keysar & Bly 1999; Giora 2003: 115), and this sometimes provides further justification for taking a sceptical view of context, the reasoning being that it is the context, not the test item, which is being interpreted (Temple & Honeck 1999).

The fact is, however, that words normally do appear in context, and the meanings they convey are usually relevant to that context. Additionally, we saw in Chapter 3 that conventional expressions attract stable cotextual patterings. These not only fix meanings onto phrases but also set up the expectation for particular meanings to occur (cf. *seventh heaven*, §3.3.3).

> What happens in everyday language is that the meaning of the word and parts of its immediate context become inseparable. Repeated usage actually fuses the sense and the expression.   (Sinclair 1996a: 115)

To return briefly to psycholinguistics, it is interesting to note that non-canonical forms of familiar phrases take longer to read than their canonical counterparts, even when they are contextualised (McGlone et al. 1994), which suggests that patterings are anticipated by the preceding cotext and whenever the instantiation

---

3.   Source: Bank of English 450m version. British Books subcorpus.

```
's decision six years ago came as a bolt from the blue for Corbett. <p> We
ings. <p> The whole thing came as a bolt out of the blue.” <p> It has been
s situation. <p> This has come as a bolt from the blue for him,” he said. <
 base rates. `It wouldn't come as a bolt out of the blue if they came down
edon manager, said: `It came like a bolt out of the blue.” <p> The alleged
ou mean?” <p> Yes, that came like a bolt out of the blue. But with Jock mak
sthood did not come suddenly like a bolt from the blue but as a result of t
oment and then it came to me like a bolt out of the blue. As a town dog, Ed
everal countries away. Then, like a bolt from the blue, Depardieu will come
usic lessons. Then suddenly, like a bolt from the blue, I said I wanted to
 <p> Describing Granada's bid as a `bolt from the blue”, Mr Astor defended
  Conservative Central Office like a bolt from the blue” and an investigatio
ly will be terribly worried. It's a bolt from the blue. <p> I don't think m
 pregnant by Ken Barlow really is a bolt from the blue.” Doesn't she think
 why he was appointed. <p> It was a bolt out of the blue, but I shall endea
ms of ING of the Netherlands, was a bolt from the blue while the absorption
t entirely unexpected, it was not a bolt from the blue. It was prepared by
vedly second best until Lee Dixon's bolt from the blue left Rioch to reflec
ped with just a broken arm. <p> The bolt from the blue came as Tony Knight
with any bereavement, but the cruel bolt from the blue that took Terry's he
sed,” he said. `It wasn't exactly a bolt from the blue because I've been on
the 4th and 15th. In the event of a bolt from the blue, don't overreact or
missing a big performance-enhancing bolt from the blue. <p> Yet at one stag
as or anybody who becomes a major a bolt from the blue and I had no one spo
s in no time at all, literally in a bolt from the blue. She has given us a
```

Figure 4.1 *Bolt from/out of the blue*: 25 of 101 concordance lines

differs from the norm, the flow of comprehension is held up while the reader/lis-tener works out how the new element fits with what was expected.

The set of concordances for *bolt from the blue* in Figure 4.1 makes it imme-diately apparent that the pragmatic, non-compositional meaning is the one sup-ported by the context, with any activation of *blue* as a colour being secondary and superfluous. The phrase's principal collocates include letters, phone calls and oth-er information conveyed or discovered; these typically arrive with no forewarning and often contain bad or otherwise unwelcome news, which is unexpected. The evidence of usage utterly fails to support the notion that the phrase might refer to objects falling out of the sky, even though this image is said to *motivate* the idiom, i.e. it can explain the meaning on the basis of what the literal meaning would be (see §5.4–5). What really seems to happen is that it is not the image as such, but its imagined consequences that are relevant to the meaning of *bolt from the blue*.

> The reference here is to a bolt of lightning coming from a cloudless blue sky. If atmospheric conditions had not led one to suspect that it might happen, such an event would be shocking and unexpected indeed.   (Flavell & Flavell 1992: 33)

The discussion of the data in Figure 4.1 is intentionally brief, and serves as a taster of what is to come in later sections. The rest of this chapter focuses on canonical forms in their typical contexts and examines the way in which cotext enables the correct meaning to be selected, in spite of salience. It will be seen that idiomatic

(non-compositional) meanings are set up by and then activated in the cotext. While it is impossible to state whether or not the salient colour word meanings are also activated, it is clear that they make no significant contribution to the meaning expressed by the phrase. In a similar vein, the motivation for an idiom, while potentially enhancing its meaning for an individual, has no effect on its institutionalised meaning.

Using a corpus to analyse canonical forms of phrases illustrates the extent to which intuitive beliefs about the meaning of idiomatic phrases contrasts with the reality of their use. In fact, as soon as a canonical phrase is contextualised and viewed in a concordance program with other contextualised occurrences of the same phrase, it becomes apparent that there is nearly as much phrasal fixity outside the boundaries of the canonical form as there is within it. For this reason, it can be useful to observe near-synonyms in context too, to see what contribution the context makes in delimiting the meaning of each, and therefore in dictating when they are selected in text production. The next three sections present case studies which examine phraseologically similar near-synonyms to illustrate how the cotext and the pragmatic function of each expression deter-mines its specific meaning.

4.3   *Caught red-handed*

Is *caught red-handed* used because it is a colourful metaphor (in more ways than one), or is its idiomaticity incidental to the meaning being expressed? Taking as a point of departure the Hallidayan notion that language should be based on a theory of meaning as choice (1992: 15), it can be postulated that the expression is chosen from other options in the paradigmatic slot including *caught in flagrante delicto* and *caught in the act*. Whether the selection is made out of a conscious de-sire to use one particular wording over another cannot be answered by the corpus data (see Teubert 1999 on this point), but what can be addressed are the factors that make *caught red-handed* more appropriate than *caught in flagrante delicto* in a particular context.

The *Oxford English Dictionary*, often the first port of call for native speakers seeking the "proper" meaning of a word or expression, provides the following definition for *red-handed*.

Red-handed
a. 1805. **1.** In the very act of crime, having the evidences of guilt still upon the person, esp in phr. *to take*, or *be taken*, red-handed.
b. Having the hands red with blood.
c. That sheds or has shed blood; bloody, sanguinary, violent.

Being based on historical principles, the OED often tells us what an expression should mean, or is believed to have originally meant, rather than what it means in contemporary use (see Considine 1996).[4] However, this feature allows us to establish that *red handed* is believed to be associated with blood and violence (hence *red*), committed directly or at close quarters (hence *handed*), and that the evidence of the crime is visible on the person of the perpetrator. In other words, it provides us with the motivation underlying the phrase.

Because this is a definition of meaning rather than an explanation of usage (an Aristotelian *definiens* rather than an *explicans*, see Hanks 1987), we can only surmise what relationship might hold between the expression *red-handed* and the events it describes. As *red* has a metonymical relation to blood, murder would presumably feature amongst the likely crimes appearing in the expression's cotext. How *caught red-handed* differs from its more neutral-sounding synonym *caught in the act* may therefore lie in the nature of the crime being committed. Yet the belief that a crime is indeed being committed stems from the expression's metaphorical motivation.

> If someone **is caught red-handed**, they are caught while they are in the act of doing something wrong. (COBUILD)

The above explanation (*explicans*) is based on the observation of data and tells us about meaning in use. Instead of committing a "crime", the person *caught red-handed* is merely doing 'something wrong'. The remainder of this case study investigates the "crimes" and "criminals" associated with *red-handed* and its near-synonyms *in flagrante delicto* and *in the act*. By identifying the participants (criminals) and verbal and non-verbal actions (crimes) in the cotexts of each expression, it will be seen that each phrase operates within a different context of situation. The near-synonymy lies at the pragmatic level: the semantic prosody expressed through these near-synonyms is effectively the same, but it is distinguished and refined by reference to its collocational preferences.

When we look closely at *caught red-handed,* the corpus reveals that it is the verb *catch* that occurs as the most frequent left-hand collocate (in 175 of the 191 examples, almost always as *caught*), while *apprehend* (2), *nab* (2) and *find* (1) are also present. *Take*, suggested in the OED definition, is not attested, though it fits the semantic preference. The verb is usually found in the agentless passive, but where an agent is specified it is associated with crime prevention (*police officers*, *crime squad officers*, etc.), and security cameras feature prominently (in three of the twelve instances where the agent is mentioned). Those who are *caught*

---

4. Consider that the first attested use is listed as 1805, which intuitively seems rather late for an expression alluding to violent crimes which leave the hands stained with blood.

```
that `25 motorists who were caught red-handed by traffic-men in the Corniche whi
nt Chart Thai supporter was caught red-handed in a house stacked high with small
 ". <p> The gang was finally caught red-handed in an armed police ambush in Septe
 Stephen Margiotta, 41, was caught red-handed on a secret video camera installed
rent. `Before we had to catch them red-handed raiding the nest," he said. `Other
lotment shed to catch the culprits red-handed. <p> But instead of finding vandal
nce who time and again were caught red-handed as they were driven back by a more
ky - unless the thieves are caught red-handed. <p> Even if they are caught, they
ster Glen) came out and caught him red-handed with his brush and tin. <p> A spel
 ended up catching the future king red-handed with his mistress. <p> The house w
on aimed at catching whale-killers red-handed, Australians and New Zealanders ha
time and money, when you're caught red-handed with the chicken in your hands, so
ning after allegedly being caught `red handed" with 200g of marijuana. <p> Boral
7-year-old youths have been caught red-handed trying to place1000 illegal bets i
ad to resign because he was caught red-handed trying to fit up a Minister for so
b> VINE T </b> A bandit was caught red-handed, ran way, was captured, and then t
an an embezzler and a thief caught red-handed. <p> Modern banking was legitimise
<p> He ran out and caught the boy `red-handed". <p> I didn't hit him hard," Mr S
nd believed she'd caught a culprit red-handed. <p> As a head came out of the pip
<p> Page # 23 Caught painting town red-handed By PETER HANSEN HIDDEN cameras pla
all to steal apples and was caught red-handed by the owner, George Bernard Shaw.
ion. <p> The Americans were caught red-handed and five were expelled. <p> As a r
 d seized from PRI officials caught red-handed. <p> Or is sticky-fingered? The sw
ually catch drug syndicate members red-handed. <p> The man has produced document
ing six days. <p> You've caught us red-handed," a spokesman said. <p> Our typist
```

Figure 4.2 *Caught red-handed* (25 of 191)

red-handed are usually criminals (*thief* (6), *culprit* (2), *bandit*, *poacher*, *drug syndicate member*, etc.) and members of extremist political groups (*a prominent Chart Thai supporter*, *PRI officials*, *senior Inkatha leaders*), though a separate group emerges featuring high-status individuals (*a high-echelon public servant*, *leaders* and the proper names of various prominent figures).

As far as the 'non-verbal action of the participants' (Firth 1950:182) is concerned, *caught red-handed* typically involves being caught engaging in one of four types of activities: dealing or using drugs, smuggling arms, stealing, and committing fraud. These are all easily identifiable as acts of criminal behaviour, although interestingly enough there is not one example mentioning murder, knife crime, or any other violent act that would be likely to draw blood, despite the fact that the expression seems to be metaphorically motivated by murder. Such a context is plausible and possible given the semantic preference for criminal activity, yet it does not appear to be used. One reason for this is that the perpetrators of the various crimes are caught while engaging in the activity, not *à fait accompli*, which again belies the literal reading and presumed origin of the expression, since blood on the hands is a sign of a crime committed, not a crime being planned or about to be committed. There is one more thing to be said about this expression: there is a degree of triumphalism expressed in the fact of catching perpetrators, glee that the baddies have been thwarted. This marks the semantic prosody.

The metaphorical motivation of *caught red-handed* can be seen to be only marginally relevant to its patterns of use despite its close connections with criminal activity. That the phrase has undergone delexicalisation is also apparent: *red* does not refer to anything of that colour, *handed* only metaphorically refers to the hands, through its associations with things held or touched by the hand, gestures, as well as control (MacArthur 2005). In fact, *handed* could just as easily be construed as referring to the capturer as to the person captured, unlike in compounds such *left-handed* or *bare-handed*. There is also evidence of *caught* being used in an extended sense rather than the more basic "entrapment", *capture on film*, *witness*, or *apprehend* being more specific accounts of what action has really taken place.

Moving on to another member of the *caught in the act* paradigm, a similar dislocation of salient and pragmatic senses can be found. There are only 23 unique occurrences of *caught in flagrante delicto* in the data consulted, which makes it more difficult to draw detailed conclusions about its typical use, but one characteristic is inescapably obvious: is used when reporting the discovery of sexual activity, which the transliteration from the Latin, 'in the act of the crime', could only account for in very oblique terms, e.g. if *the crime* were used as an institutionalised euphemism for adultery.

Only 3 of the 191 occurrences of *red-handed* relate to sexual behaviour. Only two of the 23 occurrences of *in flagrante delicto* do not – one is signalled by scare quotes to draw attention to the atypical usage (top line), the other is ironic

```
ds of America" showed three jays `in flagrante delicto", stolen egg in beak. Yet
 <p> News 1 <h> The List </h> <p> IN FLAGRANTE DELICTO: Howard Hughes and Bette
about the headmaster being caught in flagrante delicto with, how shall I put it,
 better than catching celebrities in flagrante delicto just ask EastEnders star
ous businessman found by his wife in flagrante delicto. <p> What can I do to mak
 y shown to a senior officer found in flagrante delicto in the car park at Roma S
ame this year when she was caught in flagrante delicto with British actor Hugh G
he time. My woman has often said, in flagrante delicto, `Tell me what it's like.
to have seen Caroline and Pergami in flagrante delicto on a sofa was too frighte
aths is that Mrs Nobel was caught in flagrante delicto with a mathematician. It'
 he wrongly thought he caught her in flagrante delicto with an old boyfriend, Le
 German shepherd. She caught them in flagrante delicto." <p> What, the record gu
e Carl there and then, discovered in flagrante delicto as we were, and go off an
, the man had caught young Gordon in flagrante delicto with one of the mill-girl
at his wife had been photographed in flagrante delicto. In other words, she thou
 Lambert Plaidy; being discovered in flagrante delicto by Sir Edwin, and in the
aught your father and brother <f> in flagrante delicto, <f> shall we say, it mak
ets who killed his wife and lover in flagrante delicto. <p> Mr Bletschacher was
nces. Ruby's father is discovered in flagrante delicto in the backyard while fam
 n Alicia finds Vanessa and Fergus in flagrante delicto on the college high table
 if Ms March and Mr Leung are not in flagrante delicto then it's as near the knu
s of the upper class being caught in flagrante delicto (on the job) with guardsm
s, he has not yet caught a circle in flagrante delicto. <p> Such failure is hard
```

Figure 4.3 *Caught in flagrante delicto* (all)

(bottom line).[5] Two distinct aspects of sexual behaviour are associated with *in flagrante delicto*: one is when the people involved are married (but not to each other); the other refers to the act being performed in public places. The two categories are not mutually exclusive. Although neither is a crime as such, adultery is a violation of the laws of matrimony, and both adultery and public sexual congress are considered to be an affront to public decency.

There is more detail associated with the Firthian participants of *in flagrante delicto*. Just as the difference between murder and assassination is that only important people are assassinated, *in flagrante delicto* is normally used when at least one of the two persons involved – usually the male – has high social standing or is a celebrity, while the other has a lower standing in the social hierarchy (*a woman of ill-repute*, *one of the mill-girls*, *guardsmen*). This introduces a further peculiarity, because when the people involved are not famous or otherwise important, the otherwise optional location slot becomes obligatory, as if it were necessary to justify the use of the phrase by specifying its newsworthiness (all 23 occurrences are found in journalism).

The phrase shares with *red-handed* a semantic prosody which expresses self-satisfaction: whoever chooses to use one of these phrases does so because they feel smug that the perpetrators were caught and have received their just deserts. That said, both the expressions examined so far could probably be substituted by *caught in the act*.

Caught in the act is the most general of the three phrases, and appears not to have the prosody of smugness that can be detected in the cotext of the others. In fact, it is difficult to say what the prosody of *caught in the act* is, for two reasons: the first is that the phrase is not complete, because it projects the "crime" in an extension of the syntax. *Caught in the act* is therefore either sentence- or clause-final with the "crime" mentioned in the preceding text, or *caught in the act of -ing* (the "crime"; see Figure 4.4). The second reason for an apparent lack of semantic prosody stems from this: the phrase, when not clause-final, is followed by an open slot whose 'only restraint is grammaticalness' (Sinclair 1991: 109): there must be a present participle, but there is no discernable semantic preference. When the phrase is clause final, the open slot is even more variable. Simply put, although it is recognisable as a phrase, it is not actually idiomatic and it does not participate in a complete unit of meaning since no semantic preference can be identified. The word string conveys a delexical meaning which is simply a washed-out version of the compositional one. A distinct pragmatic meaning is not attached to the

---

5.   The text in this concordance line is discussing crop circles. These were presented through the media as evidence of alien landings in rural England, but were in fact man-made.

```
occur, do not scold him unless he is caught in the act. Your puppy will not be a
night meeting of their theater club, [p] Uh, sure, Garrett,"
as Barrymore, You've Been Framed and Caught In The Act.  [p] All rely on partici
    and New Zealand.  [p] [h] Click! Caught in the act;Elton John and Madonna [/
      to catch violators of the court order in the act. [p] Catherine Spiller (F
  his hand, that is, unless it were caught in the act. Sitting in his apartment
  woodwork." [p] Are you a member of Caught in the Act?" Marsha asked the newcom
  Shane Ritchie;People Today [/h] [p] CAUGHT In the Act TV host Shane Ritchie, 28
broken the law without having to catch them in the act [p] The European Communit
situation that allows you to catch your dog in the act. Do your workday morning
        please call 0530 273873.  [h] Caught In The Act (bbc1 Television;Last nig
   more difficult [p] [h] I caught Taylforth in the act [/h] [p] A POLICEMAN told
and caught fellow gardener Derek Fothergill in the act.  [p] Alan, 60, collected
culture since `Anarchy In The UK [p] Caught in the act: Damon spreads The Word [
    said impulsively. `We're called Caught in the Act and we meet every Wednesd
leapt on a passing punt after being caught in the act at Cambridge University.
  a Romanian statue and flag and are caught in the act by a military police patr
police is you know if you don't catch them in the act it's very unlikely [ZF1]
Awful, because A.J., Robin, and the Caught in the Act kids had ignored her tota
gotten a new girlfriend. But to catch them in the act major humiliation. [p] Th
which the former England Manager is caught in the act of murdering the language
life, a dedicated reform politician caught in the act of being human in a milie
   the Iraqis apparently caught red-handed in the act of hiding nuclear materia
      himself. However, having been caught in the act of stealing he had to ret
 of gas dimly lit by about 50 stars caught in the act of winking on.  [p] The b
 still flushed with guilt over being caught in the act of searching Luke's shack
 he had seen to himself: he had caught her in the act of rolling down the littl
 look cooler than our Neon Cat Lamp caught in the act of slinking across your b
        [p] Since you have to be caught `in the act' to be prosecuted for fl
  flame. [p] Of course, when he was caught in the act originally, he vowed to r
  sort of cameras just to try and catch it in the act sort of thing how it's fo
```

Figure 4.4 *Caught in the act*

phrase because there is no need for one: the compositional meaning, albeit delexi-
calised, expresses everything that it the phrase means (Philip 2009).

Although *caught in the act* has no identifiable semantic preference, the present
participles following the phrase are interpreted loosely as being "crimes", because
if you are *caught* or *found* doing something, as opposed to simply doing it, you are
doing something that others wish to reprimand you for. This is a connotation as-
sociated with *catch*, due to its typical environments, which lends cohesiveness to
the various uses of *caught in the act*. As can be seen in Table 4.4, it is possible to list
"perpetrators" and "crimes", just as it is for *red-handed* and *in flagrante delicto*, but
these resist classification into groups. The range of "crimes" is very wide, encom-
passing stealing and adultery, as well as *talking*, and *enjoying ourselves*), and the
"perpetrators" are equally varied: *criminals*, *vandals*, and *thieves*; *lovers* and *hus-
band* – which correspond respectively to the phrases already examined – as well
as an interesting group involving *scientists*, *students*, *galaxies*, *stars* and *planets*. In
a sense, then, given its contextual heterogeneity, *caught in the act* can be seen to
be the superordinate expression which counts *red-handed* and *in flagrante delicto*
amongst its hyponyms. It is more flexible and neutral than the others because

Table 4.1 *Caught…* and contexts of situation

Caught…	Participants	Action
red-handed	thieves	drug dealing
	criminals	arms smuggling
	political activists	theft
	businessmen	fraud
	civil servants	
in flagrante delicto	celebrities	adultery
	authority figures	public sex
	lovers and spouses	
in the act	criminals	theft
	thieves	adultery
	lovers and spouses	giving/receiving information seeking
	astronomical phenomena	information
	academics	formation (of stars/planets)

being incomplete (i.e. having an open slot in the collocational framework which
has no semantic preference) the expression does not latch onto any single context
of situation and grouping of semantic preferences.

Table 4.1 summarises the cotextual features of the three expressions that have
been discussed in this section, and the similarities and differences are immediate-
ly apparent. Both *red-handed* and *in flagrante delicto* have very distinct cotextual
environments, and it can be seen clearly that they are not synonyms, but rather
co-hyponyms of the superordinate *in the act*. Their function and use cannot be
determined from a compositional reading of the phrases, which are idiomatic and
delexicalised. This is especially evident with the salient term *red*. Delexicalisation
of a colour word goes counter to our intuitions because in semantic terms, co-
lours are extremely transparent and have an unambiguous referent. Yet there is no
evidence whatsoever in the examples analysed that *red* in *caught red-handed* has
any current relation to the colour or to metaphorical and/or symbolic meanings
traditionally associated with it. The phrase as a whole is delexicalised, but kept in
place semantically by the formal features of its cotext which can be seen to limit
and fix its sphere of meaning and its communicative function.

4.4   *Once in a blue moon*

The delexicalisation of salient words is not limited to those which occur within
formally separate synonyms such as *red handed* and *in flagrante delicto*. The same
occurs in lexicogrammatical frames (Moon 1998: 145–146, see §3.1), structures
which have a variable slot which must be filled with an element from a particular

2nd proofs

semantic group. *Once in a blue moon* fits into a lexicogrammatical frame, as we have already noticed (§3.4.2). Unlike *caught red-handed*, which is a fairly transparent metaphor (even if its use belies the imagery it evokes), *blue moon* is opaque. But in fact this expression is also said to be metonymically motivated.

> Blue moons really do occur but only under extremely rare atmospheric conditions. Collins lists the occurrences of recent blue moons and explains them as dust particles (the eruption of Krakatoa in 1883, or a forest fire in September 1920).
> The allusion to the moon being blue goes back at least to a 1528 rhyme:
>> If they saye the mone is belewe,
>> We muct beleve that it is true.   (Flavell & Flavell 1992: 33–34)

The dates cited in the above description indicate that the phenomenon of blue moons is extremely rare, so we have a full explanation for the collocation and the fact that it is used to refer to (in)frequency. However, within the corpus examples themselves, we find a different account, although it does not reference any relevant literature.

> **Blue moon** is the name given to the second full moon within a calendar month, and that happens about every 32 months. (Bank of English)

However, it may well be that the second full moon within a calendar month is called a *blue moon* because of the existence of real blue moons.[6] Infrequency and a shared referent (moon) suffice as motivating factors for the original term to be extended to cover this second phenomenon, but there is no motivation provided for the colour. For this reason, the earlier account is more likely to be the correct one.

The variable slot in the *once in a lifetime* paradigm is most often filled by *once in a while* and *once in a lifetime*, occurring respectively in 57% and 33% of occurrences of the string *once in a …* in the data examined (see Figure 3.6 above, §3.4.2). In contrast, *blue moon*, although it is the third most frequent choice, accounts for just under 5% of occurrences. Is there a reason for this disparity in frequency? If so, it must be because each of the recurrent collocates in the slot expresses a different nuance, not because they are truly interchangeable.

Once in a while, being the most frequent member of the paradigm, displays the most variation, much in the same way as *to catch somebody in the act* does. Detailed analysis of this phrase will not be entered into here, as the focus is primarily on *blue moon*, but a distinction can be made, for example, between *every once in a while* and the bare form *once in a while*, and there is a further difference to be

---

6. An alternative explanation is that a *blue moon* is the occurrence of a fourteenth full moon in a solar year, instead of the usual thirteen. The motivation for *blue* in this account is unclear.

made when *only* is present (as *only once in a while* or *only every once in a while*). Because *once in a while* occurs so often at clause and sentence boundaries, this prevents it from attracting strong cotextual patternings. The consequence of this is that any identifiable semantic prosodies are weak, and really no different to the meaning expressed compositionally by the phrase (cf. Stewart 2009: 69–72). Yet two underlying moods can be traced in the use of *once in a while*, suggesting the presence of two (complementary) semantic prosodies. One of these is associated with pleasurable events and activities which only occur rarely, and can be glossed, perhaps unsatisfactorily, as "regret that the desired event/activity only happens infrequently" (Examples (8)–(9)). The other is associated with undesirable and/or unpleasant events and activities, and excess, and can be expressed as "inevitability that the unpleasant event will sometimes happen", sometimes coupled with "relief" or "thankfulness" that it only does so occasionally (Examples (10)–(11)).[7]

(8)   We had no radio. We had virtually nothing to read. We'd get some, you know, newspapers or a few magazines **every once in a while**.

(9)   I enjoy going out **once in a while** and would like it if we could set aside an evening next week, say Friday, to see a film or eat out.

(10)   Well, I think in general, people are getting just hearings and just trials, but every **once in a while**, something will slip through system.

(11)   There is a certain risk with our style of play, and **once in a while** we will pay for it.

The wide variety of collocates and the difficulties incurred in trying to identify distinct semantic groupings which do not spill over into the others, indicates that *once in a while* lies near the open choice end of the open choice–idiomaticity continuum. This is supported by its colligation with punctuation/sentence positioning. Of the 864 occurrences, 207 (24%) feature a full stop after *while* and a further 256 (30%) have a comma in the same position. *Once* starts a new sentence 282 times, while *every once* starts the new sentence 88 times; over a third of these are also followed by punctuation, therefore isolated from the surrounding syntax. In other words, in nearly three-quarters of occurrences of *once in a while* the phrase is isolated from its surrounding syntax either initially or finally, or indeed both (see selection in Figure 4.5).[8] Being an adverbial, this should not be particularly remarkable, but it does demonstrate that the boundaries of this lexical item are

---

7. Source: Bank of English 450m version. Ex. 8: British Spoken; Ex. 9: British Books; Ex. 10: British Magazines; Ex. 11: Guardian.

8. Source WordBanks Online collocation sampler. The selection of 20 concordance lines shows the proportions of each type of punctuation relative to the full data set.

```
eeds to be taken out and cleaned once in a while. With a pair of clean fingers
`t feel so terrible if I saw you once in a while." You could reply, `I'm sorry t
spapers or a few magazines every once in a while.  But there was virtually nothi
ly have one very often just like once in a while [F04] [ZGY] [F02] When you erm
y, everyone should act this cool once in a while! [p] If these nights have a the
 locals might be glad to see her once in a while, and she's bound to find someth
u need. [h] Energy audi [/h] [p] Once in a while, take the time to look at your
ly and clean out the bingo hall. Once in a while, I substitute bartend too. My p
ley von Strunckel [/b] [p] EVERY once in a while, even in the barrage of unsavou
 bus?" The soldier joked: `Every once in a while, but last time they only took a
ley von Strunckel [/b] [p] EVERY once in a while, even in the barrage of unsavou
vehicles went groaningly upward. Once in a while they could hear Ellel's voice
 [p] from Jane's Attic [p] Every once in a while I find a forgotten piece of clo
irst words of her article: Every once in a while Cro-Magnon man comes out of his
 frenzied if Senser decided-just once in a while-that democracy needn't be taken
bed. Well, maybe most of them do once in a while-but they don't as a way of life
 a while.  I admitted that every once in a while-especially when I hear about a
nswers for everything, it's good once in a while to ponder over the answers to l
he television. I enjoy going out once in a while arid would like it if we could
the banking system up there. OK, once in a while they'll put the money in a bunc
```

Figure 4.5 *Once in a while*

very well-defined, suggesting that it should be considered as more of a slot-filler than a full lexical item. It participates in no extended units of meaning to speak of and, as a result, the meaning it expresses does not differ significantly from the combined meanings of its component parts: it 'just "means"' (Sinclair 1996b: 34).

Once in a lifetime, the second most frequent member of the paradigm, is distinct from *once in a while* in that it functions as a noun modifier in 400 of the 498 occurrences, 339 of which appearing as the hyphenated form *once-in-a-lifetime*. As a noun modifier, it attracts most of its collocates from its semantic preference for chances and lucky breaks, with *opportunity* (99), *experience(s)* (57), *chance* (49), and *event(s)* (26) the most common, in addition to *trip* (20) and *holiday* (5). Despite these characteristics, *once in a lifetime* is associated with hype, co-occurring with the dream more often than with the reality of the experience. It is a preferred means for enticing members of the public to enter competitions whose prizes promise to take them away from their ordinary lives, or to buy things they don't need at (apparently) give-away prices. In short, if something is described as *once-in-a-lifetime*, the likelihood of it ever occurring is slim or (misleadingly) presented as being so (see Figure 4.6).

The explanation of the origins of *once in a blue moon* (above) tells us that the phenomenon is rare and unusual, and that for both these reasons it is also surprising. Given the preceding discussion of *bolt from the blue* and *caught red-handed* we can expect that the meaning of the idiom is not likely to have anything to do with moons or the colour blue, merely drawing on the rarity value of this event to describe other rare and unusual events. Yet since motivation has already been

```
 That's [M17] No more than once in a lifetime [M01] Save the cost of elections w
UDY, FRANK & DEAN [/h] [p] ONCE IN A LIFETIME [p] This extraordinary video featu
aid. `But the World Cup is once in a lifetime [p] But President Kim said: `Both
. This shot is only needed once in a lifetime, and protects against 87 percent o
blic dowry, available just once in a lifetime, and to be called Getting Off To A
to murder someone at least once in a lifetime. Our socialisation, our better nat
n to make the Haj at least once in a lifetime, Rome and Chartres, not being forb
the [f] Nan-shan [f] was a once-in-a-lifetime blow. Conrad, with his experience
Alan, 31, said: `This is a once in a lifetime catch. Landing two like this withi
gency, and knew this was a once-in-a-lifetime chance.  [p] We were desperate to
o England. It was magic, a once-in-a-lifetime connection to end extension of the
aid last year. `This was a once in a lifetime connection for a school like ours
 ard to in this magnificent once in a lifetime Edition. [p] PHOTO In Volumes 1 an
e staid family business of once-in-a-lifetime engagement and wedding rings.  [p]
g their equipment to get a once-in-a-lifetime glimpse of that celestial find.  J
 the upheavals caused by a once-in-a-lifetime meeting of Uranus and Neptune whic
resent list gives you the `once in a lifetime' opportunity to choose wonderful g
rial ethos.  [p] This is a once-in-a-lifetime opportunity, not just to set the r
then, when I was offered a once-in-a-lifetime opportunity to wear the beard of m
   [p] Poulson saw it as a once-in-a-lifetime opportunity to create the world's
es Oh, you know, this is a once-in-a-lifetime opportunity.  I'm not going to go
acle from heaven.  It is a once in a lifetime opportunity to get museums through
 in the park, presenting a once-in-a-lifetime opportunity to study their effect
nd I said, `It has to be a once-in-a-lifetime opportunity for me, too [p] I aske
ew and passers-by take the once-in-a-lifetime opportunity to get something for t
 gives a bride and groom a once-in-a-lifetime opportunity to ask for a diverse s
 country, this is almost a once in a lifetime opportunity for them and they're t
only trouble with having a once-in-a-lifetime opportunity is that once is never
 country, this is almost a once in a lifetime opportunity for them and they're t
ed to come.' It would be a once-in-a-lifetime opportunity to spread her message
there is of me taking up a once in a lifetime opportunity to buy tableware from
 shopping, and dining-plus once-in-a-lifetime special events and sports competit
d mean that the disastrous once-in-a-lifetime storms could return once every few
 is very exciting. It is a once-in-a-lifetime thing to excavate something as imp
Ford Mondeo 1800 LX car; a once-in-a-lifetime trip to the music capitals of the
-watching in Scotland to a once-in-a-lifetime trip to track pandas in the mounta
 [p] Mrs Ham, 52, was on a once in a lifetime trip with her husband John to cele
more than £ 4,000 for the `once in a lifetime" trip and Cunard has admitted that
[p] Panic [p] Talk about a once-in-a-lifetime vacation. We'll never go anywhere
lebrity family make their `once in a lifetime visit to Father Christmas at home
```

Figure 4.6 *Once in a lifetime*

seen to contribute to the pragmatic meaning of idioms, we will now investigate whether or not *once in a blue moon* is used to express surprise or astonishment.

Apart from explanations of the origins of the phrase, which feature in 9 of the 68 occurrences, *once in a blue moon* has a marked semantic preference for "sex, drugs, and rock and roll". These do not form a single grouping, but three distinct ones.

One semantic preference is for the expression of exasperation that something only ever happens very occasionally. Sex features prominently in this group (Example (12)), as does giving to charity, being a regular client or church-goer,

keeping in touch with friends and family, and tidying the house or garden (Examples (13)–(14)).[9]

A second group of activities centres around habits which are generally considered to be vices, such as *smoking*, *drinking alcohol*, and *taking drugs* (Examples (15)–(16)): annoyance is expressed either because the perpetrator's behaviour is anti-social, or (more usually) because the perpetrator feels annoyed that others are telling him/her what to do, and feels the need to justify his/her behaviour. This "vice" group is extended to other indulgences such as pampering oneself and spending to excess (Examples (17)–(18)). These petty vices are perfectly legitimate activities, but involve spending hard-earned cash on frivolities, which is something that society does not altogether approve of. Such "excesses" therefore induce guilt in the person engaging in the activities, much in the same way as smoking or drinking to excess do.

The third group refers to theatre, cinema and literature, and does so to praise their unusually high quality, but again, the unstated expression is one of dissatisfaction that such high quality is not the norm (Examples (19)–(20)). So here we have three distinct meaning groups in the semantic preference, but common to all is a sense of annoyance or frustration. The few instances which do not express exasperation are used to reassure by dismissing others' worries and preoccupations about violence, disasters, and other dangerous events (Example (21)): they are unimportant because the undesirable is extremely unlikely to happen.

(12)   The lucky ones have it all the time. Most of us, however, experience it only **once in a blue moon.**

(13)   You don't give a damn about the Church. You only come **once in a blue moon**.

(14)   Both our children had left home long ago. Now and then a letter came from one of them. And **once in a blue moon**, on a holiday, say, one of them might telephone – a collect call, naturally, my wife being only too happy to accept the charges.

(15)   Thank you, we know that nicotine and alcohol are nasty, dangerous substances. We also know that **once in a blue moon**, and in small amounts, they do no damage.

(16)   It's just like I basically always fancied coke. But <ZF1> it's it's <ZF0> it's **once in a blue moon** thing sort of thing like. I'm not smoking it every day of the week like.

---

9.   Source: Bank of English 450m version. Ex. 12: British Magazines; Ex. 13: British Books; Ex. 14: US Books; Ex. 15: British Magazines; Ex. 16: British Spoken; Ex. 17: British Magazines; Ex. 18: British Magazines; Ex. 19: British Magazines; Ex. 20: Guardian; Ex. 21: Independent.

(17)   Instead of going to a salon **once in a blue moon**, you could have a treatment once a week at home.

(18)   **Once in a blue moon** he would spend a vast sum to have them laundered, while his brand-new washing machine still had a sticky label on its pristine drum. Such a waste of money, apart from anything else.

(19)   Rock, rap and soul all have the same roots and **once in a blue moon** someone comes along who embodies this fact.

(20)   In fact, you wish Ms Christie's stage-appearances were not such **once-in-a-blue moon** events.

(21)   Mark you, it's only **once in a blue moon** that you ever get things escalating into violence.

Having looked at the most frequent recurrent collocates of *once in a…*, we can see that the most frequent form is the most general, while decreasing frequency brings specificity. The difference between *once in a blue moon* and *once in a while* lies in the observation that the idiomatic phrase clearly expresses an attitude towards events which occur at distant intervals, whereas the generic form does not. *Once in a lifetime* conveys a very different attitude – one of excitement and anticipation – which contrasts starkly with *once in a blue moon*'s disgruntlement. This pragmatic meaning associated with *blue moon* could never be guessed from the origins of the idiom alone, nor by examining it in isolation as a member of the *once in a [timescale]* lexicogrammatical frame. As we found to be the case with *caught red-handed*, the colour word in the idiom is completely delexicalised, no relationship whatsoever being detected with the colour blue, or, indeed, with moons. Furthermore the motivation, apparently indicating surprise, does not seem to be relevant to *once in a blue moon* as it is actually used.

4.5   *Passando la notte in bianco*

This third case study draws on Italian data, mainly to show that the same phraseological and contextualising parameters operate across languages, but also, in particular, to demonstrate that delexicalisation is not limited to English alone. The idiom to be examined is *notte in bianco* 'a night in white', which is broadly considered to be synonymous with *notte insonne* 'sleepless night'. In this section the differences in patterning and use of these apparently synonymous terms are identified and followed up with a brief comparison of *notte in bianco* and *notte in* + noun.

Notte in bianco as a phenomenon, if not as a conventional expression, has its origins in Mediaeval times.

> **Notte in bianco, notte bianca**
> insonne (i cavalieri medioevali nella notte precedente l'investitura vagliavano vestiti di bianco, in segno di purezza).   (Zingarelli 2001: 1189)
> [without sleep (Mediaeval knights held vigil on the night before their investiture dressed in white, as a sign of purity).]

There are 33 occurrences of *notte/i in bianco* in the CorIS data plus a further ten instances of the variant *notte bianca*, though only twice does this form refer to the night when no sleep was had (it usually refers to a cold and frosty/ snowy night). The most frequent left-hand collocate is the verb *passare* 'spend' which, along with its near synonym *trascorrere* account for 17 (51%) of the 33 occurrences. These verbs, when collocating with other time periods (*giorno* 'day', *periodo* 'spell', etc.) usually co-occur with activities, indicating that the time is spend engaged in some kind of activity or undertaking, so the same should be expected to be true of (*passare una*) *notte in bianco*. A further seven instances speak of tiredness experienced the morning after (*…dopo una notte in bianco*), which indicates that suffering is involved, if only as a consequence.

The reasons that an individual might have for not sleeping are many and varied. Those that collocate with *notte in bianco* are on the whole work-related, and this seems to be an important factor in differentiating *notte in bianco* from *notte insonne*. The most prevalent activity that keeps people up all night is *work* (Examples (22)–(23)), e.g. *programming* (Example (23)), although there are also some more familiar candidates, such as going clubbing (Example (24)), and several instances where generic stress (Examples (25)–(26)), caused by a range of factors, is responsible.

(22)   Le tue indagini ti hanno molto stancata, anche se non sembra. Troppe **notti in bianco**, troppe emozioni …
[Your investigations have tired you out, even if they don't seem to have. Too many nights without sleep, too much excitement …]

(23)   Non è solo un problema dei programmatori e delle loro **notti in bianco**?
[Isn't it just a problem for the programmers and their all-night sessions?]

(24)   Stefania, 25 anni, grafica di computer, ha smesso con la discoteca perché non ce la faceva a lavorare **passando le notti in bianco**.
[Stefania, 25, a computer graphics artist, has given up clubbing because she couldn't work if she stayed up all night.]

(25)   REPORTAGE IL CONTO ALLA ROVESCIA **La notte in bianco** di Hong Kong. Timori e paure, Borsa in fibrillazione.
[REPORT COUNTDOWN Hong Kong's all-night vigil. Fear and dread. Markets in panic.]

(26)   La principale causa delle **notti in bianco** è lo stress.
[The main cause of sleepless nights is stress.][10]

As can be seen from these examples, when people spend *una notte in bianco* it is when they have stayed up all night: they are not in bed, and are not trying to sleep. The exception to this is a stylistic preference expressed in newspaper and magazine health supplements, where *notte in bianco* appears to be the favoured term for sleepless nights of an involuntary nature (caused by insomnia). This accounts for most of the instances where stress is cited as the cause of the sleeplessness. This journalistic use may be a kind of damage-limitation technique, whereby the negative causes of insomnia are described with a term that carries associations of voluntary sleeplessness, but it also blurs the meaning distinction between the two expressions in the data.

The near-synonym, *notte insonne*, indicates true insomnia, when individuals are tossing and turning in bed, trying to sleep but unable to do so. Unlike *notte in bianco*, emphasis tends not to be placed on activities carried out during the sleepless night (*passare* or *trascorrere* are only found in 20% of the 159 examples analysed, compared with 51% of *notte in bianco*). Even when they do, the focus is on mental states such as worrying, thinking and wondering which prevent the individual from sleeping: the day's events and problems continue to occupy the mind, making it impossible to switch off and rest (Examples (27)–(29)). Not surprisingly, then, *notti insonni* forms combinations with agitation (5), anxiety, torment and tension (2), none of which appear in the *notte in bianco* data; the cumulative effect is to emphasise the involuntary nature of the sleeplessness and the distress that it causes.[11]

(27)   Era possibile che la **notte insonne** fosse stata provocata dalla conversazione con il governatore, e dalla minaccia della guerra.
[There was a chance that the sleepless night had been caused by the conversation with the governor, and by the threat of war.]

---

10. Source: CorIS. Ex. 22: Fiction; Ex. 23: Magazines; Ex. 24: Daily newspapers; Ex. 25: Daily newspapers; Ex. 26: Magazines.

11. Source: CorIS. Ex. 27: Fiction; Ex. 28: Fiction; Ex. 29: Non-Fiction; Ex. 30: Daily newspapers; Ex. 31: Daily newspapers.

(28)   Il giudice passava le **notti insonni** a lambiccarsi il cervello per capire che cos'era che non funzionava.
[The judge spent his nights awake, racking his brains to work out what it was that wasn't working.]

(29)   Vede, i miei eserciti sono schierati in Russia e passo **notti insonni** a chiedermi come andrà a finire.
[You see, my armies are marshalled in Russia and I spend many a sleepless night wondering how things will turn out.]

(30)   Poi la festa, splendida e insperata, e una **notte insonne** per la tensione accumulata.
[Then the party, splendid and unhoped-for, and a sleepless night because of the accumulated tension.]

(31)   È stata, per i Kennedy, una **notte insonne**, di pianti e preghiere, di disperazione e qualche bicchiere di whisky.
[It was, for the Kennedys, a sleepless night, of tears and prayers, of desperation, and the odd glass of whisky.]

We have seen that *notte in bianco* and *notte insonne* are distinguishable because of the different reasons for sleeplessness, and because of the location of the person who is not sleeping. Both are evidently recurring situations, as the plural form accounts for around half of the instances of both expressions, and neither is considered in a positive light, especially if recurrent. If it is a *notte in bianco*, then the urge to sleep is being defied (or at least postponed), and in a *notte insonne*, it is sleep that is defying the individual, causing a vicious cycle whereby anxiety causes insomnia and insomnia in turn becomes a source of anxiety.

Because sleep is essential for health and well-being, and usually done at night, the noun collocates of *passare/trascorrere una notte in…* might also be expected to share this sense of mental and physical strain. This is indeed the case. If people talk about what they spent the night doing, they were not sleeping (this being too normal a night-time activity to merit comment). In the Italian data, people typically spend their nights in hospital (6), in prison (2), in a cell (4), in a police station, on the road (3), in hotels (11) or in tents (2) or other makeshift shelters (2). Some of these places are downright inhospitable, others simply lack the comfort of one's own bed, but all represent places and situations where sleep is unlikely to be found. The ostensibly more enjoyable activities appearing in the variable noun slot – on the town (3), in the nightclub (4) – may be conceived as enjoyable activities, but not only do they too imply significant sleep deprivation, but they also collocate with uncomfortable frequency with *morte* 'death', reflecting the so-called *strage del sabato sera*, the 'Saturday night massacre' that claims the lives of

young people every week in Italy as they mix alcohol, drugs and lack of sleep, then drive considerable distances home in the early hours.

It seems, then, that any night-time activity that is spoken of generates anxiety. This has very little to do with word meaning, and everything to do with pragmatic reality, and may help account for the fact that even *notte in bianco*, despite its noble and worthy origins, is best avoided if at all possible.

In the expressions examined in this third case study, we have found yet again that the salient meanings of the constituents of the idiom, together with the motivation, have all but disappeared. Although *notte in bianco* does refer to a night spent awake, this is shared by all the expressions, both formulaic and non-formulaic, that have the structure *notte in* + noun. The vigil alluded to in the idiom, while appearing in my translation of Example (15), is not a salient feature of the idiom's meaning in Italian, as can be seen by the fact that meditative activities, including prayer (Example (21)), co-occur with *notte insonne*, not *notte in bianco*. *Bianco* also means *blank* in Italian, and this contributes to the obfuscation of the salient meaning of the colour within the idiomatic string, but it is evident that the colour *white* has been delexicalised just like *red* and *blue* in the preceding case studies.

4.6   Salient meaning, phraseology and delexicalisation

In each of the three case studies presented in this chapter, we have noted that the meaning that idioms convey is unaffected by the salience of their constituents or by awareness of their etymology. Motivation does seem to play some part, however. If the idiom's meaning seems transparent enough for a language user to motivate it without recourse to a dictionary of idiom origins (e.g. *bolt from the blue*, *caught red-handed*), then that motivation seems to harmonise with the semantic prosody. Thus a *bolt from the blue* is used of unpleasant and unwelcome surprises, and people who are *caught red-handed* are unable to conceal the evidence of their misdeeds. If, on the other hand, the expression is considered opaque (*once in a blue moon*, *notte in bianco*), the pragmatic meaning of the idiom and any apparent motivation do not seem to correspond. Blue moons are not considered normal occurrences, so any motivation for *once in a blue moon* may well relate to infrequency, but it is difficult to find either semantic associations or a pragmatic meaning related to moons and rarity that tallies with the actual meaning of *once in a blue moon*. Similarly, the virtue and willing sacrifice of sleep for purification

and the wearing of white at night-time or for prayer have no relevance to what *notte in bianco* expresses.[12]

The compositional, salient meanings of *bolt from the blue*, *caught red handed*, *once in a blue moon* and *(passare una) notte in bianco* are cast in a very different light when examined in context. Although they can be said to be "imageable idioms" (Lakoff 1987b), 'figurative expressions that tend to call up a conventional scene in the native speaker's mind' (Boers & Demecheleer 2001: 255), the reality of language use is that they function as non-decomposable units. The colour words, along with the other components, become delexicalised: *bolt from the blue* does not conjure up images of lightning and summer storms, *red-handed* does not refer to blood, *blue moons* have no functional relationship to the colour blue nor to the moon, and *bianco*, if interpreted at all, is certainly not interpreted as a colour (i.e. "blank", not "white").

Comparing near-synonyms provides further evidence that language choice is governed not by the compositional meaning of a phrase's constituents but instead by particular features in the surrounding cotext. In the case of *red handed*, a context of situation regarding crime is the deciding factor in selecting the phrase from among the other members of the paradigm (Table 4.1). *Once in a blue moon* is located on a scale of frequency of occurrence, and acquires its semantic identity through comparison with other timescale indicators, but is differentiated from these in terms of its semantic prosody. *Passare una notte in bianco* has taken on a meaning that is quite distinct from its etymological one, and it is differentiated from its apparent synonym *passare una notte insonne* not in terms of its origin or its constituents, but in terms of the different kinds of actions and events found in the context of situation.

4.7   Out of the frying pan…

This chapter has looked at two important characteristics of the relationship between form and meaning. One is that language does not combine nearly as freely

as many people would like to believe: choices have both motivating factors and consequences, as is evident when we consider the regularity of patterns which coalesce as phrases are built up into full units of meaning. The other is that when patterns start to crystallise around a phrase, the pragmatic meaning that the phrase expresses takes precedence over the compositional word meaning, depriving words of their fully salient meanings.

The case studies have demonstrated that there is a fundamental mismatch between compositional, salient meaning and contextual meaning. The connotative meanings of the colour words discussed here appeal to compositional salient meanings, but are not activated when the phrases are in use. The special status of colour words as highly salient and mono-referential is also put into doubt when we observe that they are not immune to delexicalisation. This demonstrates the idiom principle at work, and it is fascinating precisely because it goes against so much of the linguistic research that has been carried out on figurative language. Salient words are no more protected from delexicalisation than non-salient words; decomposable idioms do not undergo decomposition; metaphorical meanings of colour words are not activated despite our conscious knowledge of their connotative and symbolic meanings; motivation seems to emerge from meaning in use, not origin.

But language is not always delexical, or at least, not fully so. In the next chapters, we will look at what happens to meaning when idioms are used in unusual ways. Chapter 5 addresses variation in colligational and semantic preferences, treated here as unmarked variation, following Philip (2008). Chapters 6 and 7 address the fascinating area of marked lexical variation, in an attempt to uncover where salient, and even connotative, meanings are to be found.

---

12. Those native speakers (non-linguists) I have asked to motivate this idiom have invariably related it to a night to forget or cancel out (from the "blank" meaning of *bianco*), probably because this harmonises with the meaning it conveys. "Expert" native speakers – teachers, translators and linguists – might motivate it differently, especially since they are more likely to have come across an account of its origin. However 'experts' intuitions often differ from those of ordinary individuals who have no preconceived notions about the phenomenon of interest' (Gibbs & Matlock 1999: 263), so judgements regarding metaphoricity and metaphorical motivation tend to be asked of non experts, most often university students, but also children (e.g. Cameron 2003).

Phrases in context

Relexicalisation

Despite the regularities apparent in the cotext of phrasal chunks, it should be remembered that even 'so-called "fixed" phrases are not in fact fixed' (Sinclair 1996b: 30). One aspect of idiom fixity which is often overlooked is the idiom's colligational preference in terms of syntactic role and positioning. Most of the idioms included in this study have a marked preference for one particular syntactic role and one particular syntactic position, rarely (if ever) deviating from that preference. Yet in those cases when they are found in unusual forms or positions, that novelty attracts the reader's attention and interferes with the delexicalisation process typical of canonical instances. The first part of this chapter is dedicated to an investigation of this type of variation.

Another way in which delexical meanings can be relexicalised is by making them more semantically transparent. By far the most common means of doing this is to pair an idiom or metaphorical collocation with unusual collocates, which blocks delexicalisation and *re*lexicalises the phrase. Instead of co-occurring with one or two recurrent collocates, a much wider semantic preference is brought into play, and this can activate salient meanings within the idiom or collocation's constituents, or indeed allude to an image schema. How this works will be illustrated in a series of case studies; for now consider the difference of meaning conjured up by the usual collocates of *red tape*, i.e. *cut* and *tie (up)*, compared to alternative collocates within the same semantic preferences, including *hack*, *slash*, *hogtie*, and *strangle*. The chapter ends with a case study illustrating how salient meanings can be resuscitated, sometimes through the smallest of changes.

5.1 Variation inside the unit of meaning

The previous chapters have made repeated reference to the "meanings have words" and "words have meanings" views of language. By now it should be apparent that this study leans in favour of the former, but the latter is not to be dismissed out of hand. Instead, it is one of the ultimate aims of this analysis to discover where that belief comes from and how it can in fact be explained through corpus data. One of the reasons why the "words have meanings" view is so pervasive is that

salience is closely tied up with unusual, marked forms – the ones which attract our attention – while delexical, phrasal meanings tend to operate beneath the radar of salience and are hence inaccessible to introspection. When forms which are normally delexical undergo creative exploitation, the process of delexicalisation is halted and occasionally reversed. Words in their habitual conformations are unremarkable, but when those conformations are modified, they become visible and noteworthy.

There is a common, folk-linguistic belief about idioms that their variant forms are always highly marked, and usually result in word play or puns. This sort of variation is well documented, especially as a feature of newspaper headlines (see especially De Knop 1985; Partington 1998; Herrera Soler & White 2007), but the reality is that variant forms of idioms come in all shapes and sizes, many passing virtually unnoticed because they are not different enough from the canonical form to merit particular notice. These *unmarked* variants are less eye-catching and less interesting than those found in word play, and often involve minor syntactic or morphological changes which allow the idiom to fit seamlessly into the syntax in which it appears. The changes made do not change the meaning conveyed by the canonical form, but often block the delexicalisation process and may serve an emphatic and refocusing function. It is principally this type of variation which will be examined in this chapter.

5.2   Colligational preference and modification

Unmarked variation is typically structural, manifesting itself in the grammatical or morphological presentation of the phrase in question. It can involve, for instance, the use of the plural form of a noun when the canonical form prefers the singular, a change of tense, aspect or mood, a change of preposition, and so on. Some of these changes – especially those which affect aspect – are cited in the idioms literature as delimiting the boundaries of idiomaticity. Passivisation, in particular, is usually believed to force an idiomatic meaning to be read compositionally (Gläser 1988), but really even tiny variations made to the canonical form are known to affect idiom processing (McGlone et al. 1994). Idiomatic meaning is utterly dependent on formal features, and any deviation from these puts its non-compositional meaning in jeopardy.

Most expressions resist alteration to their canonical syntactic position, which provides an initial indication that the syntactic flexibility of idioms is indeed limited and lends support to claims made in the theoretical literature regarding their limited transformation potential. Of the 72 English idioms and metaphorical collocations analysed in this study, only seventeen showed any evidence of

syntactic role ever being changed. The most common tendency found in syntactic transformation is for a phrase or collocation to be converted into a noun modifier.[1]

Table 5.1 shows which expressions underwent a change of syntactic role, what kind(s) of changes occurred and how often, and an example of each type of change as found in the data. The two expressions which seem to resist variation the least, *white elephant* and *brown-nose*, will now be examined in turn.

5.2.1   *White elephant*

One of the more interesting types of variation has already been discussed within the *once in a [timescale]* lexicogrammatical frame (Table 3.6), and it is this: while most members of the frame function as adverbials, one of them, *once in a lifetime*, is usually a noun modifier instead. This same type of change of syntactic role will now be examined with the idiom *white elephant*.

> The devious kings of Siam invented an ingenious way of ridding themselves of any courtier who irked them. They would present the hapless fellow with a white elephant, a rare and sacred beast. The cost of maintaining the creature, which was not permitted to earn its keep as a working animal, was excessive and gradually ruined its new owner.
> A *white elephant* has taken on much more diminutive proportions in modern use, where the reference is often to unwanted items, encumbering bric-à-brac. These are sold off at *the white elephant stall* at the local church bazaar or school fete.   (Flavell & Flavell 1992: 199)

White elephant occurs 190 times in the Bank of English in the idiomatic sense specified in the definition above, and three times in the jumble sale sense.[2] The phrase is usually pre-modified by adjectives relating to size (*biggest*, *huge*, *great*, *massive*), expense (*expensive*, *costly*, *£/$ million/billion*), and attitudinally negative epithets (*embarrassing*, *controversial*, *over-ambitious*, *self-indulgent*), and overall expresses contempt both for the object described as a *white elephant*, and for the person(s) responsible for bringing it into existence. Its preferred colligates

---

1. Gustawsson's study of 300 verbal idioms also found this to be the most common type of derived form, despite the fact that verbs can (and do) also convert to gerundial nouns (Gustawsson 2006: 171–174).

2. While the jumble sale *white elephant* provides conceptual knowledge of the meaning of *white elephant* in general, it is unlikely to feature heavily in written texts for a national audience. It is also true that jumble sales are becoming less common, so the use of the term in that context may be less familiar now than it was when the definition was worded.

Table 5.1 Complete listing of POS variations found in data

Expression	Syntactic role	Exception type/freq.	Example
Black box (48)	compound noun	noun modifier (13)	That *black-box* approach to science has been largely abandoned…
The pot calling the kettle black (48)	phrase	binominal (7) / salient collocation (5)	There is a bit of *pot-and-kettle* about its outrage. / …a woman from 'Tatler magazine said on Channel 4 News last week that regional accents made the speaker "sound stupid" (*pot, meet kettle*).
Paint sth black (15)	vb-adj collocation	noun modifier (1)	…more vulnerable, much warmer than his *paint-it-black* reputation.
White elephant (193)	compound noun	noun modifier (21) / quantifier (7)	Civil servants are to move to London's £3bn *white elephant* office block Canary Wharf. / It recently opened a *white elephant* of an airport…
Red card (123)	compound noun	verb (3)	Motherwell ace Rab Shannon was *red-carded* at Fir Park for a foul…
Red ink (249)	compound noun	noun modifier (2)	Their target is about $50 billion in *red-ink* reduction.
Red tape (1464)	compound noun	noun modifier (42)	Joan Murfett, 68, got a *red-tape* demand from East Cambridgeshire District Council
Red letter day (86)	compound noun	noun modifier (1)	…rounding off a *red-letter* foreign-policy week for President Bill Clinton…
Red in tooth and claw (53)	postpositive noun modifier	noun (pre)modifier (2)	…green feminism meeting *red-in-tooth-and-claw* masculinism…
Red herring (193)	compound noun	noun modifier (9) / compound adj. (1)	… several *red herring* clues and several delays in solving the mystery. / … despite a couple of *red-herring-esque* mentions of 'suede' and its razor sharp wit,
Red rag to a bull (54)	manner adverbial	compound noun (2)	It was the *red rag* that the Spanish *bull* needed.
Once in a blue moon (57)	time adverbial	noun modifier (5) / compound noun (1)	… I was offered a *once-in-a-blue-moon* opportunity. / This *once-in-a-blue-mooner* among modern footie flicks …
Green light (1360)	compound noun	verb (8)	…in an attempt to see if they would green-light the plans.
Bolt from the blue (104)	manner adverbial	noun modifier (1)	…in the form of a *bolt-from-the-blue* brainwave.
Purple patch (149)	compound noun	noun modifier (1)	…created by John Buchan and, perhaps, more recently, by the *purple-patch* race-mysticism of Laurens van der Post.
Tickled pink (70)	compound adj.	(active) verb (4)	He was made an honorary Yorkshireman. That *tickled him pink*.
Brown nose (34)	phrasal verb	noun modifier (10) / compound noun (8) / adj-noun coll. (2)	Ditto the Blessed Margaret, as some *brown-nosing* cabinet colleague once dubbed her. / If you look on participating enthusiastically in meetings as *brown-nosing*,… / …apart from a brass neck and a *brown nose*, you will need up to date information …

---

are superlatives, often complemented by a locative (*in the world*, *in history*, etc.). Needless to say, neither *white* nor *elephant* are used in their salient, referential meanings.

The habitual and typical use of the phrase can be schematised as *BE* + *the/a* + (modifier) + *white elephant* + (locative), as illustrated in Examples (1)–(3).[3]

(1)   But political pressure to continue was impossible to resist, and Concorde became the greatest "**white elephant**" in aviation history.

(2)   Charles Arthur suggests that the International Space Station is a costly **white elephant**, and merely an exercise in public relations.

(3)   READING your article about Prescott deciding to wage war on wasted land brings to mind how many houses could have been built on the land taken by the biggest **White Elephant** ever built in Britain – the Dome.

In 21 of the concordance lines studied, *white elephant* appears in noun modifier position, as *a white elephant [noun]* (Examples (4)–(5)) and there are a further seven instances of the quantifying structure *a white elephant of a [noun]* (Examples (6)–(7)).[4]

(4)   The Tyneside MetroCentre was initially derided as a **white elephant** project.

(5)   …the betrayal on student tuition fees, the executive's economic impotence as Gordon Brown cracks the whip over them, their meek acceptance of Westminster's block grant, and the fiasco of the new Scottish parliament **white elephant** building – a catalogue of failures that can be pinned on the Government.

(6)   The building allocated to the project was a **white elephant** of a theatre, built in the heart of Bradford, left empty for 17 years, and completely the wrong shape for conventional use.

(7)   Long before that he had headed a design team to build the giant Brabazon, an over-ambitious **white elephant** of an aircraft, five times bigger and heavier than anything preceding it.

The function of this variation in syntactic role is primarily for focusing. In shifting the idiom from its habitual post-modifying position to pre-modifier instead, the author's contempt is foregrounded (especially in *a white elephant of a…*), and the precise referent of *white elephant* is specified in unequivocal terms. The

---

3.   Source: Bank of English 450m version. Ex. 1: Independent; Ex. 2: Independent; Ex. 3: Sun.

4.   Source: Bank of English 450m version. Ex. 4: Guardian; Ex. 5: Sun; Ex. 6: British Magazines; Ex. 7: Guardian.

MetroCentre project (Example (4)) is not just any old project, but is described first and foremost as a *white elephant*; it is the project rather than the centre that is being criticised. Similarly, in Example (5) it is not the *Scottish parliament*, but only the process of its building that is the *white elephant*, although the term comes at the end of a barrage of criticisms that are then summarised as *a catalogue of failures*. We see that the pre-modified nouns are in fact very generic – *project*, *building*, *airport*, *stadium* – and a connection is being made between the particular case specified in the cotext being a *white elephant* and nouns of that class in general, which presumably are not normally *white elephants*. Thus the specific event or structure is connected to events and structures in general, the shift to noun modifier position highlighting its uselessness and expense – the essential *white elephant* characteristics which, however, remain unaltered despite the deviation from the expression's typical patterning.

5.2.2   *Brown-nose*

Another illustration of syntactic role variation is provided by *brown-nose*. This idiom is unusual in that different dictionaries list it in different forms, and understandably, because it is difficult to determine which is the canonical form. COBUILD, for example, lists it as *brown-nosing*, which is found in 16 of the 34 occurrences (not quite 50%), half as gerundial nouns, half as present participles; the base form of the verb *brown nose* occurs 4 times, the adjective-noun collocate twice. Table 5.1 treats the verb form as canonical, and nominalisations as derivatives.

The range of possible forms (verb, gerundial noun, compound noun, noun-adjective collocation) does not seem to affect the meaning that the expression conveys, suggesting that it may be defined more accurately as a restricted collocation than as an idiom as such. However, its meaning is idiomatic and metonymically motivated: it hinges on *brown* being the colour of excrement, and that the obsequious are accused of getting too close to its source.

> If you accuse someone of **brown-nosing**, you are saying in a rather offensive way that they are agreeing with someone important in order to get their support. (COBUILD)

Being found in the gerundial form gives it flexibility in use as noun, noun modifier and verb, so one would not necessarily expect it to be found in many variant forms, but this simplistic observation belies the pragmatic function of the expression. The action involves two participants, and it is usually the actor, not the beneficiary, whose actions are at the heart of the pragmatic meaning. Obsequiousness is not approved of, and those who engage in it are held in thinly-veiled contempt. Thus there is an important difference in focus between speaking of the action, *brown nosing*, and the actor, the *brown-noser* (Example (8)). The *brown-noser* is in fact more likely to *brown-nose* somebody (Example (9)) than engage in the practice of *brown-nosing* (Example (11)); only rarely does the individual on the receiving end ever get a mention, being passively *brown-nosed* (Example (12)). Examples (8)–(12) illustrate each of the forms present in the data.[5]

(8)   It would certainly be true if they were like the Tory **Brown-Nosers**, those furry, servile little creatures.

(9)   "I understand. I'm a cop myself," I said. I didn't add that I wasn't about **to brown-nos***e* anybody, lawyer or layman.

(10)   Ditto the Blessed Margaret, as some **brown-nosing** cabinet colleague once dubbed her.

(11)   He'd pushed his way to his present moderate eminence by a mixture of hard work, **brown-nosing**, and ruthless opportunism.

(12)   While Saddam Hussein was allegedly **being brown-nosed** by a British MP this week, his people were out on the streets to mark the third anniversary of the start of the Gulf War.

The variation in syntactic role found for this expression can be put down to a matter of standard language flexibility whereby the different nominal, verbal and adjectival forms convey different shades of meaning. Such flexibility allows the term to be directed precisely onto the individual, the action, or the result of the action, as the accuser sees fit, increasing the strength of the already insulting accusation. Rather than accusing someone of *brown-nosing* (a nominalisation which allows for the possibility of the action being a one-off), it is possible to accuse them of being a *brown-noser*, which is a slur on their character, as it implies repeated and habitual action. The expression is found mainly in journalism, and we find that the people who are most likely to be accused of *brown nosing*, or being *brown-nosers*, are active in politics, as well as in the cut-throat worlds of business and popular entertainment.

---

5.   Source: Bank of English 450m version. Ex. 8: Guardian; Ex. 9: British Books; Ex. 10: Today; Ex. 11: British Books; Ex. 12: Today.

5.2.3   Colligational tendencies and their exceptions

The examples *white elephant* and *brown-nose* show that some syntactic variation is possible with idioms, although it is also fair to say that these are both metaphorical collocations rather than phrasal idioms, which may well account for their increased flexibility. Variation of syntactic position is, as we have seen in Table 5.1, a relatively uncommon phenomenon, and when compared with the other types of variation that are to be examined in this study it can barely be said to have any significance (but see Gustawsson 2006). This does not mean, however, that it should be disregarded, as it so often is. Variation of any sort needs to be compared with the canonical form, but extreme variation cannot reasonably be assessed without making reference to unmarked variation. All types of variation follow tendencies, and although marked forms are the ones that catch a reader's attention, they will be seen to fit the variation types that are present within the range of unmarked variants.

A second point to be made is that the absence of particular types of variation in any particular idiom cannot be taken as conclusive evidence that no variation of that type *can* occur. As phrasal strings are built up, the number of occurrences of that string diminish in proportion to the number of words it contains (cf. Zipf 1935) and the result is that idioms and other fixed multi-word sequences do not occur frequently in corpus data. The smaller the corpus, the fewer the instances of canonical phrases – let alone variations – attested. Conversely, the more times a canonical expression occurs in a corpus, the greater the chances are that variants of that phrase will also be present. This has become evident in the two corpora examined: the smaller Italian corpus of 80 million words yielded a negligible number of variants overall.

Once variation tendencies have been identified, it becomes possible to predict what general forms further variants are likely to take, but the frequency with which such variants are likely to occur cannot be predicted. There is no hard and fast rule determining the proportion of occurrences that will be canonical or variants. Related to this point is the fact that each idiom generates its own variation tendencies. Of the phrases studied, those which showed evidence of syntactic role variation were most commonly noun phrases in the canonical form shifting to noun modifier position in variants. Taking the idioms listed in Table 5.1 as a case in point, it can be seen that *white elephant* undergoes syntactic role variation in around 20% of cases, *once in a blue moon* in about 10% of occurrences, *tickled pink* in about 5%, *bolt from the blue* in less than 1%, and there are no attested examples at all for syntactic role variation in *caught red-handed* or *blue in the face*. Instances of noun phrases becoming verb phrases, or adverbials becoming nouns are rare, probably because they involve a much more forceful

kind of manipulation: because English permits the use of compound nouns and nouns in pre-modifier position, no morphological alterations have to be applied for this kind of syntactic role change, whereas creation of verbs from nouns involves changes to the word morphology as well as the position and grammatical function in the phrase. Such changes risk upsetting the delicate phraseology of idioms: too many alterations, or too drastic an alteration, can interfere with the idiomatic reading of the string and cause the meaning to fall apart. Verbs can be created from nouns, but only within idioms which are verb phrases (see also Gustawsson 2006), and when this happens it reverses nominalisation and blocks passivisation. It seems to be an extremely rare form of variation, being found in only two idioms in the data set: *(give sb.) the green light*, and *(give sb. a) red card* (Example (13)). To give an idea of how rare this is, *green light* is used as a verb (*to green light sb.*) in 0.5% of the 1360 occurrences retrieved. Adverbials can also become nouns, using hyphenation to hold the phrase together (Example (14)).[6]

(13)   Motherwell ace Rab Shannon **was red-carded** at Fir Park for a foul on him in last year's UEFA Cup clash – and picked up a six-game ban, which was later reduced.

(14)   This **once-in-a-blue-mooner** among modern footie flicks contains no swearing, sex, drugs or nudity.

The absence of any single type of variation does not necessarily preclude its existence, because the number of attested variants is always constrained by the content of the corpus. Each change to the Bank of English has introduced new non-canonical examples of fixed phrases, and removed some of the previously existing ones. However, the basic schematic patterns essentially remain the same, with new examples merely enriching the description of the tendencies already identified.

Syntactic variation and changes in the structure of multi-word expressions generally shift emphasis onto an otherwise unmarked part of the expression or onto its immediate cotext. Emphasis can be achieved by means of shifting syntactic patterns as we saw with *white elephant* and by foregrounding actors rather than their actions (*brown-nose*). The reformulation of a phrase, which occurs when the collocates are removed from their normal positions in the phraseological patterning, results in a different kind of emphasis. Here, it is possible to introduce, substitute or eliminate some elements, provided the resulting whole remains true to the canonical form's image schema.

The reason why idioms do not undergo much syntactic variation may be because they are memorised as non-compositional wholes, meaning that they are

---

6.   Source: Bank of English 450m version. Ex. 13: Sun; Ex. 14: Oznews.

not analysed and subjected to grammatical re-processing. Certainly it is more common for the cotext to bend to the will of the idiom than for the idiom to adapt to the surrounding syntax, as will be seen in the next section. A second possibility is that idiomatic meaning is felt to depend too strongly on the canonical phrasal configuration for it to withstand much variation. If this is the case, however, it is difficult to account for lexical and semantic idiom variation, which is widespread and extremely varied.

The rest of this chapter deals with variation within the semantic preference, which results in semantic reprocessing of the (otherwise delexical) core idiom or metaphorical collocation. At this point in the study, the intention is to identify and analyse general trends of semantic variation occurring in the immediate cotext of the idiomatic core, to gain an understanding of how such variation can affect the meaning of the underlying canonical form. As most idiom variation occurs at the semantic level, the next section deals only with those changes which affect the semantic preference, leaving other kinds of substitution and variation to later chapters.

5.3   Semantic preference and re-metaphorisation

When idiomatic strings are encountered in their normal contexts, i.e. together with their habitual collocates and colligates, in their usual syntactic position, and performing their usual pragmatic function, they are at least partially delexicalised. If, on the other hand, the cotext displays some unusual or less usual features, those draw attention to the idiomatic core and reactivate the semantic values of its constituents. This section investigates the kind of semantic resuscitation that can occur when the semantic preference is exploited. As we are dealing with metaphorical and idiomatic language, this involves not only relexicalisation, but also re-metaphorisation.

Before examining the data, it is worth taking time to look at how collocation affects the interpretation of metaphorical meaning. One of the best accounts to date, though little known, can be found in Louw (2000b). His "cline of progressive delexicalisation" is reproduced in full in Figure 5.1.

Figure 5.1 schematises the kind of meaning that is perceived by the reader/ listener in relation to the presence or absence of a metaphor's habitual, recurrent collocates. We learned in Chapter 2 that metaphors are identified as such in context, because they attract a different set of collocates to those preferred by other meanings of the same word. Very conventional metaphors, commonly known as dead metaphors, have highly conventionalised contexts, and this makes them prone to delexicalisation. As Louw explains, 'the presence of normal collocates

Involvement of collocation in metaphorical meaning						
Grammarian's citational examples	Symbol	Private symbolism	Poetic imagery	Full and delexical metaphors in text	Delexical	Learned expressions
Types of perceived meaning						
(No collocates)	Word "shorn" of collocates	Normal collocates replaced systematically by author	Related metaphors act as collocates to keep meaning "full, intuitive"	For humorous effect	Normal collocates	Set collocates

Left axis: FEWER NORMAL COLLOCATES — Full, Intuitive, Citational
Right axis: MORE NORMAL COLLOCATES — Delexical, Phraseological, Textual
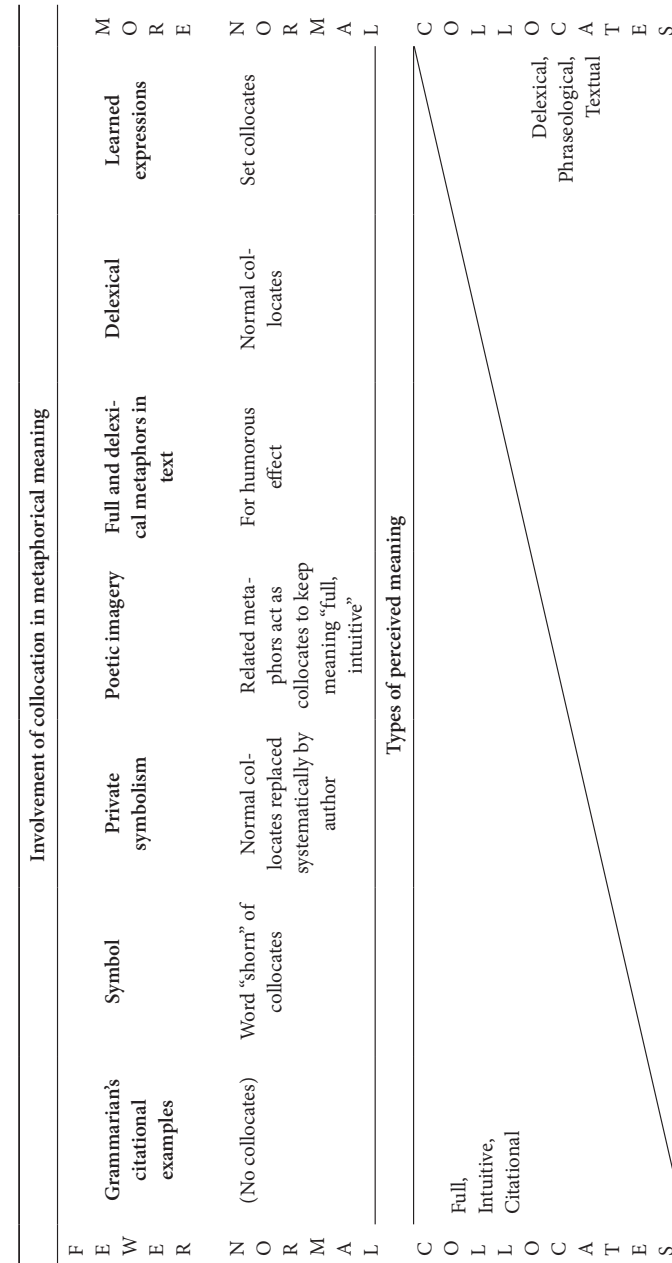
Figure 5.1 Cline of lexical and delexical meaning (Louw 2000: 15)

tends to characterise delexical forms whose metaphorical aspect is long forgotten and glossed over by readers' (Louw 2000b: 15–16). Moving through the cline from right to left, we find that the progressive appearance of unusual collocates, or conversely, the progressive disappearance of normal collocates, increases the likelihood of a word being interpreted metaphorically: 'the absence of normal collocates tends to appeal to human intuition's desire for attesting word meaning as quasi-symbolic or metaphorical' (ibid.: 15).

Louw's cline emphasises the gradable nature of metaphor, also discussed at length by Hanks (2004, 2006). Metaphor is not a clear-cut linguistic category, and the harder scholars try to locate the cutting off point between what is literal and what is not, the more impossible the task seems. All that we *can* say is that some metaphors are perceived as being much more metaphorical than others, and one of the things that makes words appear particularly metaphorical is unusual collocation. Established meanings are fixed by stable collocational patternings, and this is true of "literal" meanings and dead metaphors alike. A very metaphorical word or expression avoids those familiar collocates. As a result, its intended meaning has to be constructed from the context of its instantiation rather than retrieved from memory. A "good" metaphor is one which has a high level of aesthetic appreciation (cf. Giora et al. 2004; Steen 1994) as will discussed in Chapters 7 and 8, and a vital contributing factor to aesthetic pleasure ratings is cognitive effort.

Variation in the semantic preference of an idiomatic expression does not create a new metaphor. Instead, it relexicalises the component words and in so doing renews the connection between the words and their metaphorical meaning. The motivation of the metaphor is enhanced, and the expression is foregrounded against the rest of the cotext. Again, this kind of enhancement operates along a cline: §5.4 focuses on relexicalisation which does not lead to re-metaphorisation, while §5.5 investigates how the semantic preference can be exploited to re-metaphorise and re-motivate dead metaphors.

5.4   Relexicalisation

Relexicalisation occurs when a word or phrase becomes more noticeable than usual, and as we read in the previous section, collocation plays a major role in the process. In this section we will explore how a semantic preference is identified, starting with the most frequent recurrent collocates. These act as a focal point for the semantic preference, and determine the range of variation that can be permitted before any change becomes particularly noticeable. The semantic preference(s) for three idioms, *blue in the face*, *red tape*, and *bleed red ink*, will be investigated, illustrating the directions that variation of a preferred collocate

tends to take, and how those variations affect the overall meaning expressed. This section will also address why semantic preferences sometimes do not fit traditional semantic or lexical sets, introducing the relevance of metaphorical reasoning and the class inclusion hypothesis (Glucksberg & Keysar 1993) in the creation of semantic groupings. §5.5 will continue the discussion of this data, investigating how re-metaphorisation is caused by particular kinds of variation within the semantic preference.

5.4.1   *Blue in the face*

One of the problems facing lexicographers when including idiomatic expressions in dictionaries is to decide which words are core, and which variable or optional. In Chapter 3, we saw that many idioms are effectively fixed collocations which attract regular collocations along with the other kinds of regular patternings within the unit of meaning. One such example is *blue in the face*. This idiom is a manner adverbial typically collocating to the left with *till* or *until*, and to the right with *and* or *but*. Its verb collocate is not fixed enough to merit inclusion in the canonical form: it is "merely" a collocate.

Table 5.2 lists the verb collocates which contribute to *blue in the face*'s semantic preference for verbal processes in the 80 occurrences of *blue in the face* in the data. By separating the collocates into typical, recurrent, and non-recurrent, it is possible to appreciate how the core of the semantic preference hinges on the recurrent collocates and is refined through the variation observed. The most common form is the neutral and general verb *talk* (13). This is made more specific by the recurrent collocates: how loud? (*shout*), what purpose? (*argue*, *tell*, *deny*); and

Table 5.2 Semantic preference for *blue in the face*

Typical collocate(s)	Recurrent collocates	Non-recurrent collocates
talk (13)	shout (5)	admit
	argue (4)	answer
	tell (3)	ask
	call (2)	cite
	deny (2)	criticise
	read (2)	discuss
	speak (2)	demonstrate
	try (2)	lecture
		offer advice
		tut-tut
until (49)	'til (2)	
till (15)		

the non-recurrent collocates are more specific still. In the end, the picture that emerges is that the main type of verbal process associated with *blue in the face* is persuasive (*admit*, *cite*, *offer advice*), often contentiously so (*criticise*, *lecture*, *tut-tut*). These more detailed aspects of meaning contribute to the semantic associations of *blue in the face*.[7]

Instantiations of *blue in the face* which feature the most common collocate, *talk*, are likely to be delexical, following the argument presented in the previous chapter. Those which feature a recurrent collocate will be less completely delexical, but the extent of delexicalisation is connected to the comprehender's familiarity with a collocate, and by extension, its frequency. An instantiation with a recurrent, low-frequency collocate is more likely to be relexicalised if that collocate also has a low frequency in the language as a whole (downward collocation) than if it is very frequent (upward collocation); see Philip (2010b). The remainder of this section will look at this notion in more detail.

The semantic preference identified for *blue in the face* is clearly one where verbal processes prevail. A typical use is shown in Example (15); less central and typical but still firmly grounded in the existing schema is Example (16). Finally, Example (17) takes us into marginal territory, because although *sing* is a verbal process, *booze* and *celebrate* are most certainly not.[8]

(15)   My big concern is that we're going to say we did reform. We're going to *use the word reform* **till we're blue in the face**. And we may not have reform when you get right down to it in the end.

(16)   Writers have been *waxing lyrical* about the country house hotels for years, *extolling the virtues* of family heirlooms, clay pigeon shooting and inevitable excess **until they're blue in the face**.

(17)   NOW that everyone has *sung, boozed and celebrated* Waltzing **Matilda until 're blue in the face**.

If speaking or shouting for long periods of time without stopping for breath are a fairly transparent motivation for being *blue in the face*, also justifying its semantic preference for verbal process verbs, how can the presence of verbs not belonging to this semantic area be accounted for? In the case of *booze* and *celebrate*, a connection can be found: drinking without coming up for air; exuberance and

---

exertion causing breathlessness. However, since *sing* is already present it is not necessary to seek motivation for these "anomalous" verbs. The same cannot be said for Examples (18)–(21).[9]

(18)   Michael Howard *can go on setting up agencies* **until he is blue in the face** or any other part of his anatomy. But what happens in the prisons remains his responsibility.

(19)   I've *soundproofed* **till I'm blue in the face**, but the noise of a piano you cannot guard against. It's as though it was playing in the same room!

(20)   "You can *prepare all you want and putt* **until you are blue in the face**, but if you get on the wrong side of the hole that's it here," Sherry said.

(21)   And frankly, he can *wear a dinner jacket, write his thank you letters by hand, get up when women come into the room and wear Gieves and Hawkes shirts* **till he is blue in the face**, but nothing he does can make his behaviour that of a gentleman.

Setting up agencies, *soundproofing*, *putting* and acting the gentleman in various ways all seem to violate *blue in the face*'s semantic preference, although it should be noted that they are perfectly compatible with the semantic prosody of "exasperation caused by futile effort" which goes hand in hand with *blue in the face*. So how do these anomalous collocates fit with the expected pattern?

There is more than one way to look at this problem, and more than one possible answer. One answer is that the phrase is so delexicalised that it is chosen for its pragmatic value alone, making it suitable in these examples because the semantic prosody is appropriate. This would imply that the semantic prosody can be realised through the presence of the multi-word core collocation and particular collocates (particularly *until/till/'til*, which is crucial if *blue in the face* is to express "exasperation" rather than, say, "cold") in the absence of the semantic preference. There is no reason to believe that this is not indeed the case.

Another way of viewing the matter is to look at the similarities which can be identified between the expected semantic grouping and the separate instantiations. In other words, the incongruity is resolved through metaphorical reasoning, in particular through class-inclusion (Glucksberg & Keysar 1993; cf. 3.6.4). The verbs in Examples (18)–(21) are all actions, and verbal processes are also a form of action. Speaking until you are *blue in the face* is a kind of exhortation; the other actions share this sense of purpose. Through class inclusion these incongruous verbs can be made to fit the semantic preference, though it must be stressed

---

that they form an ad hoc and temporary classification: under normal circumstances – in any other cotext – none of these verbs would be seen as relating to verbal processes.

Independently of how we explain the acceptability of these unusual collocates, however, *blue in the face* is not predicted by the preceding cotext of Examples (18)–(21). This means that it becomes marked (more salient than normal), making relexicalisation inevitable.

5.4.2   *Red tape*

As a second illustration of semantic preferences, let us turn to *red tape*. This collocation is by far the most frequent expression in the data under study (1464 occurrences), making it also the most complex and interesting. It would be an oversimplification to suggest that there is only one semantic preference associated with the collocation, so the main areas have been separated out in Tables 5.3a and 5.3b, which treat noun and verb collocates respectively. In general, the noun collocates produce two broadly-defined semantic preferences, one stipulating the *type* of red tape (which bureaucratic entity is responsible for its creation), the other its quantity (typically excessive, which is the element that feeds a negative semantic association). The verb-based semantic preferences inform us that *red tape* collocates significantly with *cut* and a very wide range of synonyms, which again supports the feeling that it is excessive, and also with an additional preference for impediments, i.e. that *red tape* prevents people from doing the things they would like to. These verb-derived semantic preferences also contribute to the negative evaluation that is associated with *red tape*.

Due to its frequency and resultant complexity, only four broad groupings have been shown here, and only a selection of the non-recurrent collocates has been provided due to limited space. The four main semantic preferences are identified as those which can be pinned to at least one frequent collocate, and subsume more specific sub-groups which emerged in the analysis but consisted of non-recurrent collocates, or collocates which occurred at most three or four times.

"Origin", the first semantic preference illustrated in Table 5.3a, can be split into three subgroups: nations, government departments, and business sectors; while the second noun-based semantic preference, "quantity", includes a clear subgroup featuring quantity expressed in terms of length (which appeals to the image schema suggested by *tape*). Of the verb-based semantic preferences (Table 5.3b), there are further recognisable groupings for "reduction": one in which *burning* is preferred to *cutting*, another in which war is declared as a means to eradicate it, as well as a neutral grouping expressing avoidance. Finally, "hindrance" produces

Table 5.3a Semantic preferences for *red tape* (i) nouns

Typical collocate(s)	Recurrent collocates	Non-recurrent collocates (selection)
bureaucratic (41) government (23)	legal (9) regulatory (5) British (4) EC (4) Civil Service (3) council (3) United Nations (3)	Brussels diplomatic domestic educational European federal Foreign Indian industrial Korea medical political procedural protectionist
lot (10)	amount (7) bit (3) excessive (3) burden (3)	additional bushel extra great deal growing amount heaps lots weight

Table 5.3b Semantic preferences for *red tape* (ii) verbs

Typical collocate(s)*	Recurrent collocates	Non-recurrent collocates
CUT (140) SLASH (21) REDUCE (21)	ELIMINATE (6); FIGHT (5) remove (3) scrap (3)	banish beat conquering halve tackle waive
TIE (UP) (18)	STRANGLE (9) WRAP (UP) (5) BIND (UP) (4) delayed (4) hampered (4) held up (4) BOG DOWN WITH (3) enmeshed (3) ENTANGLE (3)	bundled held back hung up slowed snared tangle thwarted tied down trapped wrapped up

* Capitalised form indicates lemma.

2nd proofs

the most aesthetically pleasing of the exploitations, with metaphorical mappings involving forests, mazes and bogs – all of which create physical obstacles to progress – plus elaboration into a kind of kidnap theme involving *tying up*, *stifling* and *strangling*. Many of these metaphorical exploitations will be discussed in more detail in §5.4.2.

In order to sort the data in Tables 5.3a and 5.3b, the following rule of thumb was followed: the "frequent collocates" had to occur at least ten times (one per 150 occurrences of the node), but preferably fifteen or more (one percent of occurrences of the node); "recurrent collocates" are those occurring less than ten but more than three times, and the many "non-recurrent collocates" were supplemented by those which occurred twice. This decision was made because the increasing degree of specificity which emerges from the non-recurrent collocates is more accurately described in proportion to the total data set than with raw figures. The cut-off point for "recurrent collocates" in this case is really somewhere between three and four (i.e. more than one token per four hundred instances of the node), as can be observed in Tables 5.3a and 5.3b, though there seems to be no significant difference between recurrent word forms and recurrent lemmas.

The typical collocates for the first of the semantic preferences identified, "origin", are *bureaucratic* (41) and *government* (23). These are obviously very general nouns, so it should come as no surprise to find that individual nations (*British* (4), *Chinese* (2), *Indian*, *Israeli* (2), *Korea*, *New Zealand*) and government departments (*Civil Service* (3), *Home Office* (2), *NHS* (4), *social welfare* (2), *Whitehall* (2)) feature prominently in the recurrent and non-recurrent collocates, as do particular sectors which might be associated with bureaucracy: *educational*, *environmental* (2), *industrial*, *legal* (9), *medical*, *regulatory* (5).

The "quantity" semantic preference is expressed in terms of mass, length and weight, mass having the greatest number of both tokens and types of collocates. The generic terms *lot* (10) and *amount* (7) are supplemented by more information-rich alternatives such as *burden* (3), *bushel*, *great deal*, *growing amount*, *heaps*, and *weight*. There is also a sub-group denoting length which recalls the literal meaning of *red tape*: *line*, *mass*, *maze* (2), *thick cords*, *yards*.

If the noun-based semantic preference for quantity suggested excessive amounts, the most frequently-featured of the semantic preferences identified, populated mainly by verbs, is "reduction". *Cut* is by far and away the most common collocate of *red tape*, occurring 140 times in the data, or one in every ten occurrences of *red tape*. In comparison, its most frequent near-synonyms, *slash* and *reduce*, pale into insignificance with only 21 occurrences each. A strong collocational preference such as that for *red tape* and *cut* yields interesting possibilities for creative manipulation; not simply because there are many available synonyms for *cut*, but because there are many different ways of cutting, with

different instruments, and cutting can have a range of outcomes. These factors affect the choice of recurrent collocates, which can be seen to prefer two particular senses (cf. Table 5.3b): a relatively neutral one pulling together *eliminate* (6), *remove* (3) and *scrap* (3), and a more active and aggressive one, *fight* (5). This latter subsense is the one which is picked up on most often in the non-recurrent collocates, *banish*, *beat*, *conquering*, *overcome* (2), and *tackle*, each of which adds its own particular nuances to the interpretation of the expression as a whole.

The final semantic preference, "hindrance", is the most semantically productive area of all. In one sense, the recurrent collocate *tie* (21), often as part of the phrasal verb *tie up* (14), simply causes *red tape* to be demetaphorised, pulling the interpretation back to the literal sense of legal documents being bound in red coloured ribbon. Yet the wide range of recurrent and non-recurrent collocates suggest that something else is going on. Rather than harking back to the etymological meaning (which language users are not necessarily expected to be familiar with), it can be seen quite clearly that *tie (up)* is being creatively exploited in its "restriction of movement" sense, with all but two of the recurrent collocate types (nine of the 39 recurrent collocate tokens) and just over half of the 32 non-recurrent collocates supporting this sense rather than the one which would be literal if viewed from an etymological perspective (i.e. holding a bundle of paper together): these include *hogtied*, *hung up*, *trussed*, and *snared*. The remaining collocates support a general meaning of hindrance, and do not conjure up images of red tape as a concrete object (e.g. *embroiled*, *frustrated*, *hindered*, *thwarted*).

From this overview of *red tape*'s semantic preferences, it can once again be noted that variation to the most typical collocate enhances the semantic associations which are already in place within the unit of meaning. The variant collocates, whether recurrent or not, highlight aspects of attitudinal and connotative meaning which are often overlooked when the most frequent collocate is in use. The priming patterns (Hoey 2005) associated with the unit of meaning determine the choice of alternative lexical realisations of the meaning: there is clear evidence in the data that only one particular, preferred subsense of the typical collocate is exploited. It seems therefore that there is a binary choice to be made within each semantic preference identified. That choice is either (i) to select a neutral synonym which does not alter the meaning in any tangible way, or (ii) to choose instead to enhance and refocus the meaning by exploiting the relevant subsense of the most frequent, typical collocate. Scanning the lists of recurrent collocates, it is possible to build up a picture of the attitudinal stance taken vis-à-vis *red tape*, or *blue in the face* or, as we shall see shortly, *red ink*. In this way, the unit of meaning's semantic association emerges in corpus data as an abstraction of the relevant, shared features of the collocates appearing within each semantic preference.

2nd proofs

5.4.3   *Red ink*

The idiomaticity of *red ink*, typically within the expression *bleed red ink*, is highly marked in the data. All other colours of ink – even *black*, which is also associated with finance and so like *red ink* has a figurative meaning in addition to its literal one – collocate principally with writing: words, paper, pens and pen nibs, etc. *Red ink*, even in the absence of *bleed*, only does so to a very limited extent: only in 48 of the 249 occurrences of *red ink* does it refer to writing with red coloured ink.

Although *red ink* is included in the study because it is the core collocation involved in the idiom *bleed red ink*, it is also a metaphorical collocation in its own right, conventionally used in economics and financial contexts to mean "debt" (cf. *into the red*, *in the red*), primarily as a terminological item (136 of 201 occurrences of the metaphorical collocation). The terminological meaning of *red ink* co-occurs with a semantic preference for very large amounts of money, usually expressed in *billions*, and with businesses. However, it is important to note that this sense has no typical collocates as such, only a semantic preference. In other words, it does not coalesce into a unit of meaning but is used within the open-choice paradigm, which marks it as a terminological item (Sinclair 1991: 113). For this reason, it will not be discussed further.

Both the metaphorical collocation and the idiom which it has generated are present in the data, and they contribute to complementary, related units of meaning. The first of these has *bleed red ink* as its node, and explicitly forms a metaphorical connection between two forms of red liquid: ink and blood. The second unit of meaning has an obligatory slot immediately to the left of *red ink* in which its quantity is defined. Again, two forms of liquid are involved in the metaphor: ink and expanses of water.

Although two units have been identified here, this has been done in acknowledgement of the recognised status of *bleed red ink* as an idiom. In reality, it seems more accurate to say that there is one unit of meaning which permits two related metaphorical exploitations. It is a single unit of meaning because the semantic prosody is the same for both, glossed as "helplessness and loss of control in the face of overwhelming debt"; the semantic preferences for businesses and very large sums of money are also shared (see Table 5.4), as they are for the terminological item too. The main difference lies in the choice of typical/recurrent collocate and the metaphorical activity which this sets off, either alluding to blood trickling out of the body, leaving it lifeless, or being swept away or drowned in *a sea of red ink*. Both scenarios reinforce the prosody of helplessness; both contribute to the formation of negative semantic associations, which are once again honed in the range of low-frequency collocates in each semantic preference. Yet

Table 5.4 Semantic preferences for *bleed red ink*

Typical collocate(s)	Recurrent collocates	Non-recurrent collocates
billion (19)	million (2)	$195.1 million
$ (20)	£ (3)	$1 billion
		$1½ billion
		$2.9 billion
		$30 billion
		$500 billion
		¥26.6 billion
BLEED (7)	*FLOW* (vb) (5)*	cauterize
	stem (vb) (4)	gush
	flow (nn) (5)	hemorrhage
	staunch (2)	stanch
		trickle
SEA (9)	*FLOW* (vb) (5)	*gush*
	tide (4)	rainstorm
	river (3)	stream
	flood (2)	tidal wave
		trickle
	DROWN (3)	bathed
	swimming (2)	knee-deep
		wade
		wallowing
[none identified]	COMPANY (4)	AT&T
	Ford (3)	Australis Media
	Chrysler (2)	Banks
		fund managers
		General Motors
		News International
		Nikko
		Red Rooster
		subsidiary
		wholesaler
		World 4 Kids

* Collocates in italics can be interpreted as belonging to either (or both) of the groups.

the metaphorical aspect also seems to be important and interesting, as it distinguishes two distinct outlooks.

Perhaps it is here that the term "metaphoreme" (Cameron & Deignan 2006; cf. §3.7.2) is most usefully applied, because it makes it possible to differentiate two related and complementary metaphors which share the same pragmatic value and contextual setting. In other words, rather than being essentially a synonym of "unit of meaning", a metaphoreme would be operating *within* or *together with*

a unit of meaning, specifically as a spin-off from the semantic preference, thus contributing to the formation of more elaborate semantic associations including typical features of metaphor such as image schemata and metaphorical entailments (i.e. when the implications suggested by the metaphor are allowed to develop into scenarios). However, although it may enrich the semantic associations, the metaphoreme is not in itself a form of semantic association. The elaboration of a metaphor takes place in the mind of the language user, often on an ad hoc basis, and need not be visible in text; semantic associations, on the other hand, can be traced back to previous uses of words (cf. §3.7) and therefore motivated and supported by corpus data. Associations are held in the memory, while metaphors stimulate creativity.

To round off this section, a brief note needs to be made about the relationship of variant collocates in the semantic preference and connotation. The semantic preference has been seen to feed semantic associations, which were said to be broadly connotative in nature (§3.6). There are three types of associative connotation which are relevant here, following Kerbrat-Orecchioni's classification as illustrated in Chapter 3, Table 3.5:

i.   Signifier-signified (similar words denoting similar referents), as found in *flow of red ink / flood of red ink*;
ii. Semantic affinity, which is of course a general feature of semantic preference. This involves synonyms and near-synonyms but not antonyms: *talk / speak / tell*; *tie up / wrap up / bind up*; *bleed / hemorrhage* (sic.); and superordinates and hyponyms: *government / EC / United Nations*; *talk / cite / discuss / lecture*, *sea / tide* (*of red ink*);
iii. Previous use, in the form of extra-linguistic experience and knowledge: *thick cords / yards* (of *red tape*); *staunch the flow of red ink* (blood).

5.5   Metaphor and re-metaphorisation

The three case studies discussed in §5.4 described in detail what goes on within the semantic preference, but while we mentioned those collocates which went beyond merely enhancing meaning, we stopped short of explaining them in detail. As we have seen, sometimes variation does not limit itself to blocking delexicalisation, but also makes a real contribution to the meaning which is conveyed. Different aspects of metaphor come into play, including class inclusion as a means of justifying the presence of otherwise anomalous collocates, as well as metaphoremes and the activation of metaphorical image schemata.

Metaphorical collocates such as those just discussed, even when delexicalised, retain their metaphorical potential. They are, in Goatly's terminology, 'tired' or 'sleeping', rather than 'dead' (Goatly 1997:32), meaning that even if their metaphorical nature is not normally perceived, under particular conditions it can be revived. This process is known as re-metaphorisation, and will now be investigated.

In §5.4, we saw that the semantic preference for each of the expressions under examination is populated by two broad types of synonyms, one which enhances and refocuses the meaning of the common or typical collocate(s), the other which behaves to all extents and purposes as a neutral, unmarked alternant. In the data examined, the enhancing and refocusing collocates may well be emphatic (consider *shout*, *argue* and *deny* as synonyms for *talk* collocating with *blue in the face*, cf. Table 5.2), contributing to the attitude conveyed by the expressions as a whole and thus contributing to the formation of its semantic associations, but they do not usually re-metaphorise the core expression. Re-metaphorisation is always a possible direction for collocates to take, but it is one which is not frequently taken. *Red tape* and *red ink* provide enough data for it to be possible to illustrate the phenomenon; it should be borne in mind that some of the metaphorical collocations in the data showed no sign of being exploited in this way at all.

Red tape, it must be remembered, was originally precisely that: red-coloured ribbon. Whether language users know this fact or not, it can be safely assumed that they can form a mental picture of what red tape looks like. It is not particularly important for language users to know that the tape originally bound papers together, nor to include this in their mental picture. It is enough to know that a lot of bureaucracy finds an equivalent in large quantities of *red tape*. The variant collocates found bear testimony to these observations regarding mental images. In fact, no evidence was found at all of an image of papers bound by red tape, nor of the more abstract notion that large numbers of papers would require correspondingly large quantities of red tape. The tape has, as it were, taken on a life of its own.

Red tape may be particularly prone to re-metaphorisation because of its recurrent collocation with *cut*. Used in a conventional sense to mean "reduce", there is nothing untoward about this word; yet when it occurs in close proximity with *red tape*, something remarkable happens: the literal meanings of both *cut* and *red tape* are activated in tandem with their metaphorical meanings. So *cut* means both "reduce" and "chop/slice with a sharp instrument", and *red tape* means both "bureaucracy" and "ribbon". Since the combination is very frequent (about 10% of occurrences of *red tape*), any re-metaphorisation that occurs is tempered by the natural process of delexicalisation, but once synonyms of *cut* start to appear, the metaphor is revived.

Partly because of its concrete literal counterpart, and partly because of the semantic prosody of "powerlessness caused by excessive bureaucracy hindering one's movements", *red tape* attracts a colourful range of low-frequency collocates in its semantic preference. Tables 5.3a and 5.3b only showed those which could be pegged to a frequent collocate, but identifiable groups include: mazes, traps, jungles, swamps (all reflecting impediments); mounds and length (reflecting quantity); war (suggesting a backlash); and a range of synonyms for *cut* which tell us how the cutting is done. We will look first at this last group, which sets the re-metaphorisation wheel in motion.

Table 5.3b shows a number of the synonyms which occur for *cut* in the data examined, but the full range is: *chopping*, *hack* (2), *hacking* (2), *slash* (14), *slashed* (2), *slashes*, *slashing* (4), *slice*, *trim*, *trimmed*. These synonyms in themselves suggest that the reduction of bureaucracy is not a neutral activity. *Chop*, *hack* and *slash* are associated with particular methods of cutting, with particular cutting instruments, and with particular objects being cut, as previous encounters with the words in other contexts have primed us to believe (cf. Hoey 2005). *Slice* and *trim* also have their own semantic frames, but these contrast with the more pervasive group in terms of vigour of the action, amount being cut, object being cut, and cutting instrument used.

From this first step into re-metaphorisation, there is evidence of particular image schemata being set up. One is a metaphorical conceptualisation of quantity represented as height (vertical space), which is seen to be in operation in the likes of Examples (22)–(24).[10]

(22)   Mr Lamont *slashed* through *heaps* of **red tape**.

(23)   There is a *maze* of **red tape** to *hack down*.

(24)   … "*hacking back* the *jungle* of **red tape**" in 1992, when he was President of the Board of Trade.

Even in the absence of cutting verbs, the same conceptualisation prevails (Examples (25)–(27)).[11]

(25)   "All! *Mountains* of **red tape**, that's what all!"

(26)   …Sao Paulo had to descend through a *thicket* of **red tape**.

---

10. Source: Bank of English 450m version. Ex. 22: Economist; Ex. 23: Today; Ex. 24: Independent.

11. Source: Bank of English 450m version. Ex. 25: British Books; Ex. 26: Economist; Ex. 27: Times; Ex. 27: Guardian.

(27)   …such as late payment of bills and the *forest* of **red tape** that impedes company growth.

Red tape is confusing or disorienting, so it comes as little surprise to find recurrent correspondences between *red tape* and forests (*jungle* (4), *forest*, *thicket*), and mazes (*maze* (5), *labyrinthine, web*); picking up on an association common to mazes is the sense of entrapment which is expressed as *TANGLE* (9), and also with a recurrent "swamp" metaphor (*BOG DOWN* (5), *mired*, *swamped*, *swamps*, *wading*), which in turn corresponds to "water" associations (*DROWN* (3), *sinking*; *rivers*, *sea*; Example (27)).

(27)   Germany *casts refugees adrift* in a *sea* of **red tape**.

What do these separate metaphorical extensions have in common, and how do they fit with *red tape*? On the surface, there is little to connect metaphors of hacking through jungles and drowning. Firstly, and in common to the enhancing and refocusing collocates which were discussed in §5.4, the metaphors here all conform to the strictures of the unit of meaning of which *red tape* is core, particularly to the semantic prosody. The connection is therefore an abstract one in which, theoretically, any scenario which adequately expresses the struggle against dominion by a greater force can justifiably be used to express the struggle against *red tape*. The mapping process can follow relatively conventional lines, as in the recurrent metaphors examined, or it can be completely novel. The data examined shows that there is a preference to evoke struggles against the natural world, but other scenarios are possible: all that is essential is that certain core features, or attributes, are echoed in the metaphor.

Red tape, unfortunately, is only *cut* in about 12% of occurrences. Relevant to another common collocate, *tie*, is the "kidnap" metaphor which was briefly mentioned in §5.4.2. The motivation for using a kidnap metaphor in the context of *red tape* is, again, visual. Kidnapping involves the prevention of freedom, typically by binding and restraining with tape, cords or chains; excessive bureaucracy restricts one's freedom to act as one wishes. The parallels here should be quite obvious. Relevant verbs found in the data are *WRAP (UP)* (5), *BIND (UP)* (4), *enmeshed* (3), *ENTANGLE* (3), *shackled*, *swathed* (2), *tangled, trussed, hogtied*, in addition to the main collocate, *TIE (UP)* (18); see Examples (28)–(29).

An unfortunate corollary of kidnapping is all too often death. In particular, death by suffocation is frequently evoked in the context of *red tape,* an aspect which brings together the features of excessive quantity, hindrance/impediment and helplessness, all mentioned above as relating to the semantic prosody:

STRANGLE (8), *choked, smothering, stifled, struggle*, and *throttle* (Examples (30)–(31)).[12]

(28)   The police blame the CPS for *binding* them in **red tape** and preventing them from getting on with their job.

(29)   The right to buy left *hogtied* by **red tape**.

(30)   **Red tape** *strangles* teddies.

(31)   They pointed out that those who argue that **red-tape** is *throttling* British industry frequently muddle **red tape** (i.e. admin) costs with the cost of providing new employee benefits (e.g. cost of the minimum wage or of paid leave entitlements).

What we find from this brief foray into re-metaphorisation is that the surface wording of a conventional form can be exploited in order to draw out relevant features of the context of situation, in particular the effect that *red tape* has on its "victims". The directions that metaphorical exploitations can take are many and varied, but again it seems that some themes are preferred over others. The analysis of other data and other expressions may lead to different conclusions, but the tendencies described above were noted throughout the colour-word expression data set, making it unlikely that this is an isolated example.

Finally we come to the blood/water metaphors already identified as being typically exploited in the context of *red ink*. This collocation is slightly problematic, because although the idiom *bleed red ink* has been lexicographically documented, it could be that *bleed* is just the preferred collocate within a semantic preference for "flow of liquid". Clearly, *red ink* is a liquid, and *bleeding* also refers to a red liquid; this makes *bleed* an enhancing and refocusing collocate, but not necessarily an essential constituent of the idiom.

Bleed, in such close proximity to *red*, re-metaphorises *red ink*, making the conceptual switch from *ink* to *blood* extremely easy. A handful of instances do precisely this through the use of the verbs *stem* (4), *staunch* (2), *stanch* and *cauterise* (Examples (32)–(34)), all of which typically collocate with *flow* (of blood).[13] In other words, we expect *stem the flow of…* to be completed by *blood*, so when in its place we find *red ink*, the connection is made to blood through the 'combinatorial affinity' that habitual collocation patterns provide (Kerbrat-Orecchioni 1977; cf. §3.6.2).

---

12. Source: Bank of English 450m version. Ex. 28: Today; Ex. 29: Times; Ex. 30: Today; Ex. 31: Guardian.

13. Source: Bank of English 450m version. Ex. 32: Economist; Ex. 33: Oznews; Ex. 34: Economist.

(32)   That is good news for consumers, who are likely to see the price of PCs tumble by at least a third in 1993, but dreadful for the legion of small, mostly American, PC makers struggling to *stem the flow of* **red ink**.

(33)   Hollis Harris's first job with his One Big Airline must be to make the cuts to people, planes and facilities that will end service duplication and *cauterize the flow of* **red ink**.

(34)   The deteriorating US trade performance comes despite the Clinton administration's determination to *staunch the flow of* **red ink**.

The liquid nature of *red ink* unsurprisingly attracts verbs which describe the movement of liquids. Of these, only *bleed* and *hemorrhage* (sic.), are more appropriately related to blood than any other particular liquid. This point is important precisely because of the ambiguous status of *bleed red ink* as an idiom. The remaining verbs have no exclusive association with blood (*FLOW* (5), *SWIM* (2), *DROWN* (3), *bathed, dipped, dripping, gush*) and, additionally, the quantities of liquid mentioned are rather larger than we would comfortably associate with blood: *trickle, flow* (5), *rainstorm, tide* (4), *flood* (2), and *sea* (9) (Examples (35)–(37)).[14]

(35)   The profit comes after a *tidal wave* of **red ink**.

(36)   Like its namesake, it disappeared beneath a *sea* of **red ink**.

(37)   Balance sheets *swimming* in **red ink** made them vulnerable to takeover on the cheap.

The re-metaphorisation of *red ink* seems to occur simply by enhancing its liquid state, whereas the connection of blood leads into the creation of one metaphoreme, *bleed red ink*, which places particular emphasis on helplessness. The presence of extents of water (*a sea of, a river of*, etc.) signals the other metaphoreme, whose emphasis is on quantity. This division into metaphoremes makes it possible to highlight features of the unit of meaning, developed further through metaphorical reasoning, where new connections, metaphorical entailments and connotations are forged. Importantly, although it may conjure up an image in the mind of the language user, a metaphoreme enhances but does not add new meanings or nuances to the unit of meaning. In the examples mentioned here, we have seen metaphorical connections which consist of the mapping of the pragmatic inferences of a novel scenario onto the semantic prosody, which is a more complex kind of class-inclusion procedure than the one which we mentioned with reference to *blue in the face*, §5.4.1. As far as associative connotations are concerned,

---

14. Source: Bank of English 450m version. Ex. 35: Guardian; Ex. 36: National Public Radio; Ex. 37: US Books.

re-metaphorisation can involve (in addition to those already mentioned at the end of §5.4):

i.   lexical analogy in the form of polysemy, found in the dual processing (simultaneous interpretation) of the literal and metaphorical meanings of *bind up/wrap up in red tape*; and

ii.   combinatorial affinity through habitual collocations, as illustrated in *stem the flow of blood/red ink*.

5.6   Semantic preference, variation, and "imageable" idioms

Studying the variation patterns in an expression's semantic preference has been seen to yield insights into the finer points of meaning that it conveys. In particular, it has emerged that whenever an idiom conjures up a cognitive representation in the mind, this representation can be exploited in the coining of new variant forms. For simplicity's sake, this representation has been referred to so far as an image schema (after Lakoff 1987b), although it should be noted that there is considerable individual variation in the ways in which people conceptualise. No single conceptual image is shared by all language users, for instance, and a proportion of the population is unable to visualise mentally at all (see especially Finke 1989:27–28 on this point). So in addition to the visual, an image schema may also feature sound, smell, and other so-called 'secondary perceptual simulators' (Ritchie & Dyhouse 2008:94; see also Ritchie 2006). These relate to the cognitive representation of meaning, generally based on previous experience of the meaning in the real world or in the imagination (reading, watching films, and so on). Because of this individual variability, no single researcher can legitimately suppose that his or her image schema is shared by other users, and it is here that the analysis of attested variation proves its worth.

This chapter now concludes with a case study examining the "imageable" idiom *(like a) red rag to a bull* which is rather interesting because although not particularly frequent, it undergoes a wide range of variation (see Philip 2000): the canonical form accounts for only 14 of the 54 occurrences located. This enormous variability comes as a result of its semantic transparency which makes it very easy to form a mental image of its compositional meaning. It is a commonplace that bulls react angrily to red, especially in the form of a red cloth being waved in their faces, and this image is available for creative exploitation along with knowledge of what the implications of provoking bulls are. In other words, even without knowing the idiom it is quite possible to use the relevant image schema when referring to a context of situation in which the implications of provoking bulls is relevant.

> Spanish bull fights are well-known throughout the world. The brave matador shakes his cape in front of the bull's nose, enraging it. It was believed that the red lining of the cape excited the bulls and made them even more fearful opponents. Sad to say, it appears that bulls are colour blind and react to the movement of the cape, not to its colour. However, that was not widely known when the Spanish bull-fighting practice found its way into our simile *like a red rag to a bull*.
>   The phrase is used to mean "likely to cause great annoyance or anger" […] It is often used in connection with people who get angry very quickly.
>   (Flavell & Flavell 1992:43)

In terms of the idiom and its variants, a wide degree of flexibility can be observed – far more so than with non-imageable (opaque) idioms. It can withstand most of the variation types identified in Moon's (1998) exhaustive study, including truncation (shortening of the phrase), substitution of constituents, ellipsis, and expansion; and it can also be completely paraphrased, the link to the idiom being preserved by the presence of the key words *red*, *rag* and *bull* and the pragmatic inference within a relevant context of situation. The examples that are to be discussed here are the British English *red rag to a bull*; the institutionalised variant in American English, *red flag before a bull*, is discussed elsewhere (Philip 2000).

The variability of *red rag to a bull* is not only caused by its semantic transparency and "imageability", but also by the simple notion that the longer the phrase, the more potential it has for variation. This idiom is calculated as being between five and seven words in length, making the potential for variation considerable. There are three key collocates: the core collocation, *red rag*, and its preferred collocate, *bull*, to the right (n+3). In the canonical form, both nouns are accompanied by the indefinite article (*a red rag*, *a bull*), and are connected by a preposition, usually *to*. The resulting noun phrase has a colligational preference for comparative structures (left of the node), usually expressed in simile form (*BE like/as a red rag*), though, as is common with figurative language, the comparison may be rendered implicit by the absence of *like/as*, transforming the phrase into a metaphor. The preferred syntactic position is sentence-final.

Although it has already been mentioned that the phrase is particularly prone to variation, this does not seem to operate within the colligational preferences. Only two instantiations of the idiom show the "noun phrase to noun modifier" type of variation (Example (38)), and there is no attested conversion to verb (i.e. **to red rag somebody*) although in place of *be a red rag* we can find *wave a red rag*, which conforms to the image schema, six times (Example (39)).[15]

---

15. Source: Bank of English 450m version. Ex. 38: Independent; Ex. 39: Oznews.

(38)   If he was being taken to task for plain poor taste rather than an imagined departure from the jazz purist's path of righteousness, Branford proved he was more than fair game, wearing his baseball cap back to front with pride, a badge of manhood *waved* **like a red rag to** *the* **bull** *of his leery lad's club of a band*.

(39)   After South Africa confirmed their place in Sunday's final with a 75–45 win over Antigua and Barbuda, van Dyk *waved* **a red rag to a bull** with her announcement she was glad not to be facing New Zealand in the final.

This addition of *wave* makes other changes necessary. From a description of a state (*be like…*) to an active process (*wave…*), elements of the image schema have to be made more explicit. So instead of *like a red rag to a bull*, we find examples such as those illustrated in Examples (40)–(42).[16]

(40)   The thought of making it harder on criminals and easier for the police is, to the council, like *waving* **a red rag to a bull**.

(41)   Good on you for *waving the* **red rag** *in front of the* **bull**, writes Mr Doyle.

(42)   Our presence may have the effect of **a red flag** *held permanently* **before a bull**.

This illustrates a more general observation which can be made regarding idiom variation: it is common for more than one type of variation to be in operation in a single non-canonical instantiation. Variation types seem to work together to maintain some kind of equilibrium, preserving the image schema on the one hand, and ensuring that the pragmatic meaning is not violated by inappropriate additions (or removals). The interlinking of structure and meaning often makes it nigh-impossible to change just one small element of a phrase and keep the original meaning unaltered. It is more usual that one change will lead to a series of re-workings so that the phrase can keep to its original meaning while obeying fully the grammatical and lexical restrictions that bind it together.

Variation operating within the semantic preference is particularly interesting in this idiom. Intuitively, the key words are the salient content words, *red*, *rag*, and *bull*, the image being pinned down by the proximity of these collocates, which would explain why variations to the less salient elements can occur with little effect on the meaning. But *bull* can occur at quite a distance from the core

---

Table 5.5 Semantic preference in *red rag to a bull*

Typical collocate	Modified forms	Other non-recurrent collocates
bull (22)	a bullish Beijing Diane's bull the ramblers' bull the Spanish bull the Unionist bull those Bulls	a slightly parochial regulatory board an earthy border side like Melrose B[ritish] T[elecom] businessmen some people the Americans the Euro-sceptics the purists the red rose the red-baiters the true blues Turkey Turner

collocation *red rag*, admitting modification (Examples (43)–(44)), inversion (Example (45)), and it can also – surprisingly perhaps – be absent (Example (46)).[17]

(43)   But the damage had been done and it was **a red rag to** *the Unionist* **bull**.

(44)   Mention of the Forest of Bowland is a **red rag to** *the ramblers'* **bull**.

(45)   "People who say we sound like The Faces should keep their big mouths shut", spits Rich, ignoring my question and charging towards his pet hate-topic **like a bull to a red rag.**

(46)   The idea of cross-company unions is **a red rag to** *businessmen*.

Bull thus appears to be an optional element, in the sense that the slot it occupies in the canonical form is subject to considerable variation (Table 5.5); yet unlike the semantic preferences we looked at earlier, there are no recurrent collocates occurring in place of *bull*, and no distinct metaphoremes emerging at this point. Instead, there is a trend to modify the referent of *bull* through possessive structures, the "modified form" (cf. Table 5.5) representing an intermediate stage in which *bull* is still present, while the non-recurrent collocates omit the word entirely.

In the second column, it can be appreciated that, e.g. *Diane's bull* is not a bull at all, but Diane herself; similarly the *Spanish bull* is the Spaniard. In these cases, *bull* metaphorically stands for the bull-like qualities expressed by those people. This trend is continued through the non-recurrent collocates, which are

---

connected through class inclusion because they share, at some level, relevant attributes of *bull*.

It should be remembered that class inclusion is not determined from the tabular representation provided, but within each individual instantiation. The preceding cotext sets up an expectation for *bull* which is not fulfilled, but instead sets off connotations associated with *bull* (connotation by combinatorial affinity; §3.6.2). What are those connotations likely to be? Uncontrolled anger, stubbornness, strength, stolidity, and any manner of other bull-like attributes (Philip 2000: 226–227). These are then transferred onto the non-recurrent collocate, albeit temporarily for the purposes of local interpretation. Viewed together as in Table 5.5, their only shared semantic trait is that they are all human (this is sometimes expressed metonymically, e.g. *Turkey*, or metaphorically, e.g. *the true blues*, the *red rose*). They coalesce into a semantic set only by virtue of the shared attributes which are relevant to the context of use. Some of these attributes are more obvious than others (Example (47)).

(47)   It is **like a red rag to a bull** and, *being a Taurean*, I react accordingly.[18]

In the range of variation observed for this idiom we can start to appreciate where evidence from psycholinguistics and cognitive linguistics converges with evidence from corpus phraseology. Although canonical forms are delexicalised and betray no adherence to any underlying image schema, once variation starts to occur, that schema is activated, and metaphorical activity can start to take place. But we are not dealing with a creative free-for-all. Idiomatic meaning is constrained by the image schema and also by wording. The variation that occurs must not be so extreme as to interfere with the image schema nor may it fracture the link between that schema and the pragmatic meaning conventionally expressed by the canonical form. This is why the precise constraints vary from one idiom to another: different combinations of image schemata, wordings, and meanings lead to different lexical and syntactic limitations.

5.7   Toeing the line

While Chapter 4 introduced the role of delexicalisation into idiomatic and phraseological meaning, this chapter has shown that delexicalisation is not a one-way street. Apparently minimal changes to wording can block or reverse the process, and sometimes cause dormant metaphorical associations to spring to life. Yet

---

18. Source: Bank of English 450m version. Today.

variation seems to be primarily a lexical event. Non-compositional phrases on the whole resist variation to their syntactic position and morphology, supporting the restrictions of idiom creativity and variation that are put forward from a theoretical perspective (e.g. Gläser 1985; Makkai 1972; Fernando & Flavell 1981).

Variation to fixed phrases seems to occur within particular parameters, respecting the image schema (if present) and the pragmatic and attitudinal inferences of the context of situation which are initially set out by the canonical form within its extended unit of meaning. Variations must preserve or enhance the schema, otherwise the link between the phraseological meaning and the words used to express it is at risk. Variation must also conform to the semantic prosody, and this point will be expanded upon further in the chapters to follow.

So far we have only scratched the surface of semantic and lexical variation, and only in the final case study is any suggestion made that core elements of an idiom can also be changed. This is in fact the most common type of variation found in the data, and the one that provides the most information about the activation and suppression of connotative meanings. The next chapter investigates substitution, variation and addition of key words, and intends to shed light on the interplay between literal, metaphorical and connotative meaning of (colour) words, and their instantiated meanings in text.

Variation, metaphor and semantic association

The variant forms of idioms that have so far been encountered represent a broad class of relatively unmarked variation types in that despite the variety of their instantiations, none except *red rag to a bull* has involved the substitution of an idiomatic key word. This opens up marked variation, which will be discussed from this chapter forward.

While idioms and their unmarked variants can be quite easily located in corpus data, the same cannot be said for marked variants. A study of variations made to the most salient constituents of an idiom, its key words or core collocation, necessarily involves the retrieval of versions which do not contain those words. This is exceptionally problematic in a corpus study: corpus query software is designed to retrieve specified words, but if those words cannot be identified before the data is analysed, how can they be extracted in the first place?

This chapter starts with an overview of the paradox of finding the unfindable in corpus data, followed by an explanation of how it can be overcome. Later sections then deal with the variation tendencies that can be observed when colour words are substituted by replacement colour words, finding connections between referential, metaphorical and propositional meaning.

6.1 Don't count your chickens…

6.1.1 Variants and corpus data

As every corpus linguist knows, corpus query software is designed to retrieve specified words. The default option is for those words to be invariable, thus a search for "good" will retrieve every occurrence of that four-letter string in the data. Regular expressions can almost always be used in addition to this basic operation, thus alternant forms may be retrieved, e.g. "book/books", to retrieve all occurrences of *book* and all occurrences of *books*, "book*" which uses the wildcard (*) to retrieve all occurrences of every word string which starts with the four-letter sequence b-o-o-k, including *bookish*, *bookshop*, *bookworm*, etc. As well as these options, more sophisticated corpora, typically those available commercially, are also POS-tagged and lemmatised. A part-of-speech (POS) tagger makes it possible to search for e.g. *book* used only as a verb. If the corpus has also been

lemmatised, it will be possible to retrieve all inflections of the verb in a single search; and while this may not seem important for a regular form like *book*, it is of invaluable help when a verb is irregular (and therefore cannot be retrieved using a wildcard), e.g. *go-goes-going-went-gone*, and in languages which inflect more than English does.

As well as single-word searches, it is possible to combine words. A word pair, e.g. *red rag*, can be retrieved as a multi-word node (the query syntax will vary from application to application), and, often, also as a non-contiguous collocation. In this second case, the user can specify a maximum number of intervening words (five is typical), and usually the results returned show the words in the sequence indicated in the regular expression. Thus *red* + 0–5 intervening words + *rag* would appear only in that sequence (*red … rag*), never inverted (i.e. *rag…red*). There are some exceptions to these general rules, and these will be mentioned shortly.

The kinds of search queries just described start from the (not unreasonable) premise that a researcher knows what he or she wants to study, and have been designed to facilitate access to recurrent patternings. Yet, as we have already begun to realise, sometimes recurrent patternings operate at a level that is higher than the word form: the variants already examined all functioned at the level of semantic preference, so were related in terms of meaning – but not surface form – to the canonical version.

Semantic relatedness in itself is not a problem. The word sense disambiguation literature abounds with reports on the effectiveness of natural language processing applications to aid semantically-based queries (e.g. Lin & Pantel 2001; Edmonds & Hirst 2002; Budanitsky & Hirst 2004; Piao et al. 2005), and Rayson's (2007) *Wmatrix* has also been proved effective for the identification of metaphorically-used lexis (e.g. Hidalgo Tenorio 2009). But semantics is a tricky business. Most formal semantics works with the hierarchical relations of synonymy, hyperonymy and hyponymy ("a *bull* is a *bovine quadruped* is an *animal* is a *living organism*") which are essentially literal (i.e. basic, referential) in nature. The variations that we have been observing display similarities at an attributive, metaphorical level. Budanitsky & Hirst (2004), in fact, lament the difficulties to be encountered in assessing semantic relatedness, because 'lexical semantic relatedness is sometimes *constructed* in context and cannot always be determined from an a priori lexical resource' (2004: 45, emphasis in original). This takes us to the heart of the matter. Attributively-determined semantic groupings constitute open-ended semantic sets whose members are neither fixed nor quantifiable. Akin to Hanks & Ježek's (2008) "shimmering" semantic sets, membership waxes and wanes depending on which attribute, or even which facet of any given attribute, is deemed to be relevant in the context of use. It is therefore quite easy to see why 'from a lexicographical viewpoint, they are simply nightmares' (Moon 1996: 252).

In the preceding chapters, it has been possible to identify the idiom from its core collocation; yet with *red rag to a bull* there came the realisation that this strategy might not always work, because *bull* has been seen to undergo considerable (and unpredictable) variation. So what if *red* or *rag* can also be varied? How could such variants be found, or, alternatively, the hypothesis disproved? One way to find variant forms is to focus on the core or key words of the idiom, which are likely to be its salient lexical items, and to conduct loosely-defined searches in which these words are both present, but not necessarily contiguous nor in the same sequence as they occur in the canonical form. This is possible on small corpora run through the *ConcGram* corpus query software (Greaves 2009) but not possible on corpora which are stored remotely and accessed by proprietary query tools, as is the case with the corpora consulted in this study. *ConcGram* builds on natural language processing work into n-gram and "skip-gram" (Guthrie et al. 2006) extraction, but allows a single search to be made for non-contiguous collocations (up to 50 characters / circa 12 words apart) and automatically retrieves all combinations, regardless of sequencing in the initial query (Cheng et al. 2006, Greaves & Warren 2007). Yet what it cannot do is recognise that one (or more) of the input words has been substituted by a synonym, nor can it output results for searches for lexicogrammatical frames, semi-prepackaged phrases (cf. §3.1) or idiom schemas (cf. §3.4.2) which have no fixed lexical form but rather require each slot to be filled by a member of a particular semantic set.

There are further complications to the retrieval of variant forms. In §3.3.3 the concept of the idiom key (Cacciari & Tabossi 1988) was introduced; it was also noted that although many psycholinguists are quite sure that this element exists, it remains unclear whether it is lexical or syntactic, or indeed a combination of the two, and whether it is a single word or a collocation. Variation, therefore, can potentially affect any constituent of the idiom. In the examples examined in previous chapters, it has been assumed that a core collocation is essential while greater variety can be introduced into the rest of the phrase and its immediate cotext. However, other idioms seem to rest entirely on an image schema, allowing variety in the lexical slots, but not in the syntactic patterning. Moving a step further into abstraction, it is also possible that the syntax alone can hold an expression together, even when the slots are not occupied by members of the preferred semantic set, as will be seen in §7.5.

Having outlined the full nature of the problem, a solution is called for. The next section sets out the procedures followed to extract the maximum number of creative variants of *the grass is always greener*, using techniques developed in earlier studies (Philip 2000, 2003, 2008).

6.1.2   Finding variants: The case of *the grass is always greener*

It is impossible to predict the creative potential of other language users. It is also impossible, at present, to formulate a single search which can retrieve all relevant variant forms of an idiom, while discarding those which are irrelevant. In the face of these givens, the only reliable way to find variations is to carry out an exhaustive series of queries to extract all possible permutations. Once this has been done, the results have to be collated then sifted manually to eliminate duplicates and identify marginal or irrelevant examples before proceeding with analysis. The sequence of queries which is to be described here is more complex than the ten search strings required to locate all the variant forms of *red rag to a bull / red flag before a bull* described in Philip (2008), but is based on the same principles, which are:

1. Identify the salient words and the syntactic patterning;
2. Formulate a maximally-inclusive search string for each possible combination of the salient words. This should include:
    a. the maximum possible number of intervening words that the software permits, and
    b. inversion of the normal sequence of the collocates.
3. Formulate a maximally-inclusive search string for the syntactic pattern. This should include:
    a. each possible word which can fill the slot (closed classes)
    b. the maximum possible number of intervening words between the identified syntactic elements, and
    c. if possible, the use of POS tags for open classes (e.g. noun, verb).
4. Repeat the above should any new tendencies emerge in the variants retrieved.

A list of the combined searches carried out for *the grass is always greener* is shown in Table 6.1, the number of hits recorded relating to the WordBanks Online collocations sampler. The search queries are listed in order of formulation, to show how series of related queries emerge as different hypotheses are tested.

   As can be seen from the strings listed in Table 6.1, the collocation *green+grass*, was queried first because it was felt to be the salient core collocation of the phrase. The second set of queries tested the verb inflections; the third set, the preferred adverbial; and the fourth set, the colour. After this came a series of more syntactically-oriented searches, testing adjectival forms (so not limiting the search to comparatives alone), the complementary but not necessary locative completion of the idiom (*over there / on the other side*). The final searches, as can be appreciated, search for the structure in the absence of the salient key words. The relevant unique hits are discussed in §7.5.

Table 6.1 Search strings for locating variant forms of *the grass is always greener*

Query	Total hits	Relevant hits
grass # green*	35	15
the grass is	21	7
the grass was	11	1
always greener	6	5
the # is always green*	3	3
is still green*	0	0
is usually green*	0	0
tends to be green*	0	0
is always green*	3	3
is always black*	1	0
is always white*	0	0
is always red*	0	0
is always blue*	0	0
is always yellow*	0	0
is always pink*	0	0
is always grey*	0	0
is always brown*	0	0
is always purple*	0	0
is always orange*	0	0
the # is always more ADJECTIVE	2	0
ADJECTIVE on the other side	20	3
grass # over there	1	1
green* over there	1	1
the # is always * over there	1	1
the # is always # on the other	0	0
Total	105	40
Total unique hits	77	15

Represents 1–3 intervening words permitted.
* Represents any characters (including 0) after specified string.

   Once the variants were found in corpus data, the trends that they were following could be identified with more certainty, making it possible to carry out further searches on the Internet to find more examples of variation and thus enrich the description of *the grass is always greener*. Of the variants found in the corpus search, every one contained the key word *grass*, so in order to find out if there were variants involving this word in Internet texts, the search string *"the * is always *er"* was formulated to find further variations to the description of the

grass within the comparative structure in which *greener* could be excluded, but other comparatives included. This string takes advantage of HTML coding which allows the wildcard (*) to specify a character string occurring not only at the beginning of a word (e.g. *green**), but also at the end (e.g. **er*), thus making it possible to find comparative adjectives with relative ease; it also acts as a dummy (unspecified) word. A second query, *"the grass is always"* – *greener* (= minus *greener*), was formulated.[1] The second search string was designed to find variants in the slot usually occupied by *greener*. The variants found from this search are also be discussed in an extended case study in §7.5.

Deliberately searching for variants in a general reference corpus requires time and patience, but the results make the effort worthwhile. As the numbers in Table 6.1 indicate, there will inevitably be irrelevant examples to discard and duplicates to eliminate; yet while no claim of exhaustiveness can ever be made, it is far preferable have too much data which then has to be sifted through, than not to have enough data, or not to have sufficiently representative data because the queries carried out were too specific and/or too few.

6.2   Variation of key constituents

While Chapter 5 was concerned with unmarked variation, defined as variation which left the idiomatic core intact, this chapter and the one following it look at the marked forms of variation that have been found in the data. Of particular relevance to this study are the types of changes that are made to colour words. This is not an arbitrary choice. As I mentioned at the start of Chapter 1, colour words are particularly useful candidates for studies of figurative meaning because on the one hand they have a very clear "literal" meaning, and, on the other, they have an incredibly wide range of figurative meanings. Metonymy accounts for most of these non-literal meanings: the colour word can be used in place of any entity that is (or typically is) that colour. Additionally, there are symbolic and other conventionalised meanings associated with colours, such as the adoption of a particular colour by a political faction, or the conventionalised use of *red* to mean 'danger' or 'stop', and *green* to mean the opposite (as found in traffic lights, safety flags, and so on).

The kinds of regular changes that were observed, and which will be discussed in this chapter, include substitution of key words with semantically-related words,

---

and the substitution of one colour word for another colour word (lexical set substitution). Idiosyncratic and humorous coinings, including the translation of key words and their substitution with phonetically-similar words, are reserved for Chapter 7, where changes made either to the canonical form itself or to its typical cotextual patterning are seen to trigger aesthetically pleasing effects due to their connotative resonance.

6.2.1   Semantic variation

In Chapter 5, we saw that lexical substitution can reactivate the conceptual schema suggested by the idiom. *Bull* was seen to be substituted by a range of quite unrelated alternatives, but they could all be made to fit somehow into a semantic schema with *bull* and its relevant attributes (especially anger) at the centre. A similar type of substitution is common in the idiom *black sheep of the family*. In this case, it is *family* which undergoes creative substitution by attributively related terms, and it does so quite often – in 32 of the 195 occurrences (including literal uses) of the collocation *black sheep*. A selection of these non-standard "families" is reproduced as Examples (1)–(5).[2]

(1)   As a result **the black sheep** *of the lending sector* is now at the heart of the financial establishment.

(2)   The focus is on the sources of incomes earned by a herd of unusual and anomalous "laborers", the **black sheep** *of economic science*.

(3)   If it refuses to comply, France will become the **black sheep** *of the EU*.

(4)   "We were the **black sheep** *of the community*," adds Melon bassist Brad.

(5)   We're the **black sheep** *of the New Wave*.

The same idiom in Italian, *pecora nera della famiglia*, undergoes exactly the same kind of substitution, and with comparable frequency (in 5 of 35, or 14%, of occurrences): Examples (6)–(10).[3]

(6)   La considerano la **pecora nera** *dei Wagner*.
     [It's considered the black sheep of Wagner's operas]

2nd proofs

(7) In un capriccioso gioco delle parti, l'Italia, fino a due mesi fa **pecora nera** *del gregge europeo*, viene ora promossa a pieni voti a Bruxelles.
     [In a strange twist of fate, Italy, black sheep of the European flock until two months ago, has passed with flying colours in Brussels.]

(8) Fino a qualche giorno fa la Olivetti era la **pecora nera** *dell'industria italiana* e molti giornalisti, commentatori e persino analisti sviluppavano con disinvol- tura giudizi e previsioni catastrofiche.
     [Until a few days ago, Olivetti was the black sheep of Italian industry and many journalists, commentators and even analysts were gaily going around making judgements and predicting catastrophe.]

(9) Secondo me – dice lui – mi tenevano d'occhio, perché ormai mi considerano la **pecora nera** *del quartiere.*
     [I think – he says – they were keeping an eye on me because they already had me down as being the black sheep of the neighbourhood.]

(10) La raccolta di saggi dal titolo *Gli irregolari*, con prefazione di Pierluigi Battista, non è soltanto una galleria di grandi incompresi o **pecore nere** *della Prima Repubblica.*
     [The collection of essays entitled *Gli irregolari*, with a preface by Pierluigi Battista, isn't just a line-up of the misunderstood and black sheep of the First Republic.]

While the transfer of attributes from *family* to these "synonyms" is perhaps not as semantically rich as that between *bull* and its substitutes, precisely the same mechanisms can be seen to be at work. It is also important to notice that the same system is in operation in two languages, which allows the same kinds of modifica- tions to take place, and with similar frequencies. In every one of these variants, as well as those which feature *family* with premodification (e.g. *the Mitchell family*, *the compositae family*, *a philanthropic Baltimore family*), the variant always refers to some kind of community or kinship group, albeit framed metaphorically. So although the collocation *black sheep* may seem to operate independently of the rest of the idiom, it is still obliged to adhere to the semantic preference of the ca- nonical form if it is to convey the idiomatic meaning, "bad or worthless".

6.2.2   Lexical variation

Substitution of a part of an idiom with a lexically-related term achieves quite dif- ferent effects to those noted for semantically-related forms. In this section, the fo- cus closes in on colour words. These belong to a very clearly-delimited lexical set, so the results may be more marked for this lexical set than for some others. This is no bad thing however, as it makes it easier to demonstrate the mechanisms at

work in changing meaning, even if it is true that not all meaning changes caused by substitution and variation are expected to be as marked as these.

Colour terms veer towards the metonymical area of figurative meaning, and although they can be used metaphorically, most of the time their secondary meanings are grounded in experience. Their symbolic meanings can be detached from the original symbols but the connection can be restored though experiential knowledge; and arbitrary links such as those between colours and emotions are usually linked to other symbolic meanings if they are not directly motivated by metonymy.

Colour-word idioms often feature variation to the colour word, and the rea- sons for such modification are no different to those already cited in this chapter and the preceding ones: emphasis, detail, and better contextual fit being the most common functions in operation. There is a difference between the lexical set of colours and many other lexical sets, though, and it is not simply the fact that co- lours are highly salient and imageable. This factor makes substitution of one co- lour word with another more noticeable, certainly, but that greater visibility (both in terms of meaning and of imageability) is not colour words' only forte. Colours are also, fundamentally, gradable.

Since colours are subdivisions of an otherwise uninterrupted continuum, they are available for metaphorical exploitations wherever gradable concepts are to be expressed. The intensity or shade of the colour can be metaphorically mapped onto the intensity or shade of meaning being expressed. This is not the end of the matter. They have a further peculiarity in that their gradability operates on three distinct scales: tone (hue), shade (light/dark), and saturation (intensity). There are, in addition to these three gradability scales, a range of synaesthetic metaphors which are conventionally associated with colours, including mapping of tactile sensations such as touch (*soft/harsh*; *smooth/rough*), temperature (*warm/cool*; *hot/icy*), and also taste (most notably sweetness: *mellow/cloying*). These are also scalar concepts and therefore suitable for mapping in the same way as the non- metaphorical scales are, but as none were found in the data studied, they will not be discussed further.[4]

The gradable scales associated with colours are easily evoked metaphorically, and can be activated in isolation or in combination to create a veritable spectrum of nuances. The data analysed suggest that hue change is the preferred means of expressing intensification of meaning, while shade change is the preferred

---

4. The use of a synaesthetic metaphor in the re-metaphorisation of idioms would require dou- ble processing (deriving the metaphor from another metaphor) rather than a single process of mapping from an "imageable" concept to an abstract one. This probably explains why, although it is theoretically possible, it has not been attested in this data.

means for expressing mitigation. Hue change can also be used in mitigation, but if it is, it works in combination with saturation change to create an impression of bleaching (i.e. whiter, more transparent colours). In the sections below, examples of these metaphorical mappings will be illustrated with the variants found in corpus data.

6.3   Variation of hue: Emphatic nuances in the expression of emotion

Colours are conventionally linked to the expression of emotion (D'Andrade & Egan 1974), and many colour words are used to express emotional states. In this section we will examine envy, anger, and embarrassment as expressed in colour-word idioms.

The intensification (or mitigation) of these emotions is usually expressed by changing the colour word to one which represents a different hue. As far as the words used to refer to these different hues are concerned, there are two possibilities open. One is to emphasise by reinforcing the colour, which is done by substituting the original, basic colour term with a specific colour term, which may in fact refer to the same colour that the basic colour term normally refers to. This is the only kind of gradable colour change found in to occur in the idiom *green with envy*. A total of six colour-term variants were found for this idiom, four of which were intended to reinforce the emotion. Two of these used *emerald*, and the other two *pea-green* (Examples (11)–(12) and (13)–(14) respectively). Both *emerald green* and *pea green* fall within the range of the focal colour, so they are expressing emphasis through their specific reference. The choice of *emerald* in Example (11) is also contextually grounded, as *money*, *art market*, and *art lover* occur in the immediately-preceding context; the choice of *pea-green* in Examples (13) and (14) may well be related to references to high school, because (pea) green is a colour that is sometimes used for school uniforms, especially at girls' private schools, although the contextual information does not make this connection explicit.[5]

(11)   When it comes to serious money, the art market is led from New York these days and just listing the works coming up for sale is enough to make an art lover's mouth water and turn his eyes **emerald with envy.**

(12)   Mars is strongly placed, imbuing you with the kind of energy, courage and sheer sex appeal that make lesser mortals **emerald with envy.**

---

5.   Source: Bank of English 450m version. Ex. 11: Times; Ex. 12: British Magazines; Ex. 13: Sun; Ex. 14: Sun.

(13)   If I still had them all today, I could have worn them all at the same time and looked like a real fashion victim. Those Taras and Clarissas would have been **pea-green with envy**.

(14)   One girl used to receive about half a dozen of those gigantic overblown cards which could comfortably house a family of four. We'd all be **pea green with envy** and it was only years later that I discovered she actually sent them to herself.

An alternative to using a specific colour as a substitute for the basic colour term is to use a different hue altogether. This is favoured in both English and Italian for the expression of anger and rage, as well as for embarrassment. These emotions are themselves gradable, but the colour words used reflect the fact that this emotive gradability is reflected by the colour of the face, which gets redder and redder the more intensely the emotion is felt.

A search was carried out for the idiom *red in the face* to discover which other colours could appear in the place of *red*. This expression is interesting because *in the face* has distinct meanings depending on which colour word precedes it, and all refer quite transparently to the colour of the face caused by emotion, temperature, state of health, and so on. *Blue in the face*, for example, is usually found within the longer string of *talk until you are blue in the face*, *blue* representing the physical manifestation of a lack of oxygen in the bloodstream caused by talking too much and forgetting to breathe (Examples (15)–(16)). *Pink in the face* can refer to exertion from exercise and also to mild embarrassment, and can be viewed as a mitigated form of *red in the face* when the causes are the same (Examples (17)–(18)). When the emotion being described is anger or rage, the scale adopted remains within the broadly-defined area occupied by red; it does not encroach on blue, but can occupy *purple* (Example (20)). The same is true of the related expression *to go red* (with anger/rage/embarrassment, etc.), which partially overlaps with *red in the face*. This expression contrasts with *to go white*, when blood drains away from the face as a result of illness or fear (Example (21)). These expressions could hypothetically merge in the area occupied by *pink*, but rarely do: *pink* is connected to the *red* scale, while *white* tends to occupy an isolated, hue-free area, and is typically mitigated by *grey* (Example (22)). Because of this relative isolation, *white with rage* can be seen to be at least as emphatic, if not more so, than *purple with rage*. To summarise, fear causes us to *go pale* or *go white*; embarrassment starts with *pink* and ends at *(bright) red*; anger starts at *red* and moves through more saturated shades (*bright red*, *very red*) to finish at *purple*; lack of

breath starts at *blue* and, unlike the emotions, only moves down the shade scale, ending up at *black*.[6]

(15)   And so to the core of the puzzle: one can demonstrate that this a safe technology until one is **blue in the face** and people simply will not believe it.

(16)   I counted until I was **black in the face**.

(17)   To her surprise, he went **pink in the face**. "Ah, yes," he said, looking guilty.

(18)   My mother for her part would go **red in the face** and spout bitter recriminations.

(19)   I'd always get embarrassed, go **bright red in the face**.

(20)   Just as all the barristers go **purple in the face** and start screaming if anybody says they are not perfect, so it was exactly with the solicitors.

(21)   **White in the face**, Joe frantically rounded the corner.

(22)   He stood up, **grey in the face**, holding on to the arm of the settee.

Italian follows these tendencies, but also produces a few surprises. The range expressed for colours occurring with *rabbia* 'rage' extends farther, including *nero* 'black' after *viola* 'purple'; and *verde* 'green' is also present. The frequencies of the alternative colour terms are shown in Table 6.2, where they are compared with their English equivalents. While *verde di rabbia* 'green with rage' might sound odd to an English speaker, it should be recalled that both black and greenish yellow are colours of bile, and that bile is related to anger.[7] It can be seen that *verde* and

Table 6.2   The colours of anger in Italian and English

Italian	English
Nero di rabbia (8)	*Black* (0)*
Rosso di rabbia (5)	Red with anger (18) / rage (28)
Verde di rabbia (5)	Green with anger (1) / rage (2)
Bianco di rabbia (2)	White with anger (7) / rage (9)
Blu dalla rabbia (1)	*Blue* (0)
Viola (0)	Purple with anger (1) / rage (21)
Rosa (0)	Pink with anger (3) / rage (2)

* Terms in italics are unattested in the data.

---

6.   Source: Bank of English 450m version. Ex. 15: Independent; Ex. 16: US Books; Ex. 17: British Books; Ex. 18: British Books; Ex. 19: Oznews; Ex. 20: Times; Ex. 21: US Books; Ex. 22: British Books.

7.   These associations are preserved in the Italian expressions, but speakers may not consciously process the meanings.

rosso di rabbia 'green/red with rage' occur with the same frequency, indicating that *verde* 'green' is not innovative, but may in fact be an alternant.

A closer look at the cotext of the variants listed in Table 6.2 suggests that different degrees of anger seem to be expressed with different colours. This lends support to Cacciari et al.'s (2004) findings that people associate very specific colours with emotional states. *Rosso di rabbia* seems to be the canonical form, because the types of anger expressed are varied, and in several occurrences the idiom is expressly linked to the face turning red. *Nero di rabbia* has a semantic preference for physically-expressed anger; and *verde di rabbia* is quite clearly an emphatic variant. The rarer forms *bianco* 'white' and *viola* 'purple' are too infrequent for general comments to be made, but in the instances observed, *bianco* indicates cold, controlled anger and *purple* wild mad anger which involves throwing things and shouting, both uses of colour which conform to other models and to the colours' connotative meanings generally.

6.4   Variation of tone: Ameliorative and pejorative nuances

We saw briefly in the discussion above that tone variation can be used to express ameliorative and pejorative nuances: *blue in the face* becomes *black in the face* not only because it is emphatic, but because one's physical condition has worsened (one is, in fact, close to death). When the intention is ameliorative, a lighter colour is chosen, for example when a *black week* (on the stock exchange) is transformed into a *grey week*, or when one *goes pink* (with embarrassment) rather than *red*. Amelioration by change of shade is not as common as hue variation because it seems to be limited by the availability of basic colour terms which contain *white*, which physically and semantically washes out the power of the saturated colour. In both Italian and English, there are only two such terms: *grey* and *pink*.

As has already been noted, *pink* is always related to *red*, and can be used to represent a toned-down version of any of the meanings associated with that colour. *Grey* is associated with vagueness, because it is neither *black* nor *white*. This is used to advantage in the creative variation of colour words in idioms. The colour itself does not participate in the spectrum, so is excluded from the scales of hue and saturation which we have seen to be very semantically productive. The only scale it belongs to is that of light/dark, but the linguistic evidence suggests that while *grey* can be used creatively to mitigate either *black* or *white*, the reverse is not true: no instances were found of *grey* being substituted (refocused) by either *black* or *white*.

The Italian idiom, *pensieri neri* 'gloomy thoughts' is derived from the association of black with negativity, depression, and desperation. English prefers (dark)

blue for these meanings. When *nero* is substituted with another basic colour term, the associations specific to each of those colours are evoked. The only recurring colour word substitution for this expression is *pensieri grigi*, but unlike the other colours, *grigio* does not express its connotative values of vagueness or boredom, but lightens the blackness of *pensieri neri*. The same happens with *anima nera* 'villain', but the effect is more marked.

Anima nera is predominantly found in political contexts, where it has overtones of sinister dealings and policies (Example (23)), but outside of politics it seems to be read from a moral-religious viewpoint (Example (24)).[8]

(23)   Ricordo male, ma Amato non era il braccio destro, l'**anima nera** di Craxi? Provvedimenti impopolari… dicono di essere stati penalizzati da provvedimenti adottati per il risanamento del Paese.
[I may be wrong, but wasn't Amato Craxi's right hand man, his henchman? Unpopular measures… they say they were penalised by the measures adopted for the Country's recovery]

(24)   Respiro. Sono vivo, sto andando verso il Signore. Sono un'**anima nera,** ma Dio la monderà. Egli è buono con me.
[I can breathe. I'm alive, I'm on my way to the Lord. I am a villain, but God will cleanse me. He is good to me.]

(25)   Un danese dalla pelle nera e un nero dall'**anima bianca.** Entrambi keniani, ma di un Kenya diverso.
[A Dane with black skin and a black with a 'white (lit.: pure) soul'. Both Kenyans, but from different Kenyas.]

(26)   Se vivi vicino a una di queste **anime grigie** non hai una sorte molto lieta.
[If you live near one of these villains you won't have an easy time of it.]

Unsurprisingly for a term that has a polar opposite, the negative meaning(s) of *nero* can be cancelled out by substituting *nero* with *bianco*, thus turning the malicious being into a "heart of gold" (Example (25)). The use of *grigio* in Example (26) is fittingly indistinct: it seems to refer to a mitigated form of *anima nera*, but it may equally refer to shadows and ambiguity, both connoted by the colour grey.

---

8.   Source: CorIS. Ex. 23: Ephemera; Ex. 24: Fiction; Ex. 25: Daily Newspapers; Ex. 26: Non-fiction.

6.5   Variation of saturation: Emphasis and defocusing

Variation of colour saturation is typically used to express emphasis. It takes the colour in the canonical form as its starting point, and either increases or decreases its saturation, depending on whether the meaning is to be intensified or mitigated. In most cases, the colour word in an idiomatic expression represents the focal colour in the language, which may alternatively be known as the prototypical (Rosch 1973) or basic (Berlin & Kay 1969) colour. In both English and Italian, focal colours are highly saturated.

The strongest and purest (primary) colours in the spectrum are both highly saturated and maximally bright, as can be observed by playing around with the settings on any image editing program on the computer. As the focal colours already refer to highly saturated hues, there is not much scope for intensification by increasing saturation, and in fact it is more usual for intensification to follow the tendency visible in the spectrum by complementing saturation with brightness.[9] In so doing, the emphatic form of a basic colour term reflects the colour's full saturation and maximum luminescence. Modification of *red in the face* is reproduced in Figure 6.1.

Considering that the emphatic colour term (fully saturated and bright) refers to something that is very close to the focal colour, it should not come as a surprise that mitigation is more common than emphasis, and that it is expressed by gradually moving away from the fully saturated, bright colour.

As we move away from saturated colours in the spectrum, the decreasing saturation also brings us closer to the black–white continuum. For mid-range and dark colours (*red–purple–blue–green*), the farthest point from the focal colour is

```
tly or not at all? <p> b) Go bright red in the face? <p> <f> 4 If you are anxiou
d always get embarrassed, go bright red in the face. <p> When they found out I'd
Tracey began to cough and go bright red in the face, but Sharon and I thought it
 new man because I know I go bright red in the face, groan and look terrible. Ho
hen upset or angry do you go bright red in the face? 4 If you had to speak to a
away. Frank Little # she was bright red in the face now – `I can't believe you'r
  get drunk, becoming puffier and redder in the face, looking as if he was about
ing hoops and getting redder and redder in the face until he was rescued by an a
the scene, Brecht got redder and redder in the face. Finally, he unscrewed the to
  fficial Mike Knowles became even redder in the face as he pumped a football up d
```

Figure 6.1   Full saturation of colours in *red in the face*

---

9.   Note that Italian seems to prefer intensity to brightness, typically modifying the basic colour term with the adjectives *forte* 'strong' or *intenso* 'intense'/'deep', or with the augmentative suffix *-one* (e.g. *verdone*). However, *forte* is highly polysemous and can also refer to strong (bright) light, suggesting that the use of *forte* to modify a colour term simultaneously increases saturation and brightness.

a shade of grey. But rather than mitigate by *dullness* lexis (e.g. *dreary*, *dull*) which would take us down the greyness scale (shade), we find that mitigation involves the adoption of less focal colour words, realised linguistically in English with the suffix *-ish* (*reddish*, *greenish*), and in Italian by the equivalent, *-astro* (*verdastro*, *bluastro*), or the diminutive suffix *-ino* (*verdino*, *azzurrino*).

This scale is reflected in the variants found for *purple prose*. A total of 68 concordance lines were found relating to the idiom (canonical and non-canonical forms). Reproduced below is the full array of non-standard occurrences, which can be seen to span the use of the superlative form *purplest* (Examples (27)–(29)), the mitigated form *purplish* (Example (35)), modified forms (Examples (31)–(34)), shade change (Example (35)), hue change (Examples (36)–(37)), absence of hue (Examples (38)–(39)), and non-specification of hue (Examples (40)–(41)).[10]

(27)   The owners are driven to **the purplest** of Beverley Nichols **prose** as they rhapsodise over their patches…

(28)   His **prose** is at its **purplest** when he praises the authority imminent in institutions and glorified by tradition.

(29)   One understands the passion of the early commentators as they described in **the purplest** of **prose** what the moving-picture theater meant to the city's working people.

(30)   If you like your **prose purplish** and your plots steamy, try this sample from a book coming out next month and see if you can identify the author.

(31)   If however, your reading revolves around the straight and narrow, the **purple** (*verging on blue*) **prose** misogyny will not amuse you.

(32)   In **sub-purple prose**, they chronicle his passion for another woman – unarguably, unmistakably the unfortunate judge.

(33)   This includes long stretches of imagined dialogue (often sounding risibly like a re-make of Gone with the Wind) and passages of **deep purple prose** describing intense experiences.

(34)   This is a man with a past writ large – in **prose of darkest purple**.

(35)   Question: what do these two scenic descriptions have in common, apart from the **violet prose?**

---

(36)   Who else would describe Labour's 1997 election victory as "a cornucopia of luscious psephological fruit?" But for Lord Jenkins, the report is more than an exercise in **clarety prose**.

(37)   The occasion was a piece I had written dedicated to **scarlet prose,** called "Novel Reactions", which Peter read and I attempted to play out.

(38)   Mr Dobson's, which arrived a day earlier, was a plain photostat of four sides of **grey prose**.

(39)   As for The Levellers, well, there's no validity in their accusations, because all the great journalism of the late-20th century lacks any objectivity and **dull, grey** John Major **prose** the Levellers want, they should go to the crappy dailies.

(40)   With vivid imagery and **colourful prose**, Margulis explores the origins of life, the essence of humanity and the functioning of the Earth.

(41)   Often penned in **colourful prose** by Lord St John himself, these letters don't mince words.

As in other cases, the effects of these colour changes may be determined by contextually-relevant features: Examples (36) and (39) are undeniably contextually determined, several of the others may well be. The change in colour is dictated partly by the hue, and also by the connotative values assigned to it in the culture. *Purple* in *purple prose* is intended to recall the richness and sumptuousness of the regal colour purple (see Appendix), although if the trend for delexicalisation observed elsewhere is any guide, this association is probably not activated in the use of the canonical form. The connotation of wealth is recalled in both *claret* (colour of the wine of the same name) and *scarlet* (once an extremely costly dye).[11] In the context of this idiom the new colours are intended to provide extra information – information that goes beyond the gradable concepts that we have seen in other examples. This happens because *purple prose* is not gradable: *prose* either is, or is not, *purple*. So although the extent to which a text is mannered and flowery can be emphasised (Examples (27)–(30)) or toned down (Examples (32) and (35)), other changes, even if related along the colour spectrum, are not seen to be relevant because they do not map onto an equivalent scalar concept. Thus *claret* refers to the

---

wine, the culture of good eating and drinking, and is a close enough colour match to purple for the idiom to be recovered, *blue* co-occurs with *purple* to emphasize the sexually explicit content of the prose. *Grey* can be interpreted equally as relating to *colourful prose* (a semi-compositional collocation) or the idiom *purple prose*, but in either case, it expresses the conventional meanings "dull", "tedious". Thus, in the variants presented here, it can be seen that variation is not only able to enhance the meaning expressed by the canonical form, but may in fact take over from it by causing the meaning to be detached from the wording in a typical context. In fact, the meanings expressed by different hues (Examples (36)–(41)) are all supported by elements in the cotext. The delicate balance between phrase and cotext will be further investigated in Chapter 7.

6.6   Culturally constructed colour scales

In addition to the natural scales (hue, shade, and saturation), colours are incorporated into culturally determined "artificial" sequences. These do not follow the spectrum, so require shared knowledge if they are to be used in successful communication. As a general rule, the colours cannot merge into one another to express nuances, because the mapping of the natural scales impedes this. The colours of traffic lights illustrate the concept clearly. *Red* means "stop" and *green* "go"; but any intermediate movement ("get ready to go", "get ready to stop") can only be mapped onto the intermediary colour in the established sequence – *amber* – and not onto any of the colours which actually come between *red* and *green* in the spectrum.[12]

A similar case can be found in the adoption of colours by political parties. *Red*, as the colour of revolution and blood, has long been the colour of the political Left. *Blue*, symbolising constancy (and by extension, reliability, but also resistance to change), justly belongs to conservativeness and hence to the political Right. These colours merge to form *purple* in the chromatic spectrum, but not in the political spectrum. In the United Kingdom, where parties are commonly referred to obliquely by their associated colour, the sequence for mainstream politics is (from left to right) *red–yellow–blue*, exploiting the distinctiveness and separateness of the primary colours to indicate their distinct and separate

---

12. As red and green stand at opposite ends of the colour spectrum, a "natural" colour scale would permit any intervening hue to represent the intermediary stage. Amber occurs along the red-yellow axis, but if the sequence were not culturally fixed, blue (the opposite of amber/orange) could just as readily serve the same function.

political ideals.[13] *Purple* is used as a colour in politics, but is far less well established than either *blue* or *red*.[14] *Green* represents environmentalists all over the world, but in Italy it has been hijacked by the right-wing separatist *Lega Nord* 'Northern League', while *black* is irretrievably associated with Fascism.

Given that these kinds of sequences do not relate to the ordering of refracted colours in the chromatic spectrum, they are left only with the shade and saturation scales for their creative modification. Thus New Labour, whose politics are distinctly less left-wing than its predecessor, (old) Labour, is assigned the colour *pink*, rather than the traditional *red* used to denote the political left.

That said, it is uncommon for gradation to be used in politics. As we saw in the variant forms of *purple prose*, the different colours embody different sets of concepts and do not mix. Because the colours of political parties are institutionalised, they participate in creative variation of a similar kind to that seen above whereby the new colour replaces the canonical one not only in hue but also in meaning. This kind of substitution re-metaphorises the canonical form, as it draws attention to the full meaning of that form indirectly while at the same time contributing new meanings, thanks to the variant colour word. These changes can be ad hoc, but sometimes they show signs of becoming themselves institutionalised. This happens when a creative variation starts to gain currency in the language by being picked up and re-used. Two examples of this are presented here – *clear blue water* and *camicia nera* 'black shirt' – illustrating aspects of British and Italian politics respectively.

Clear blue water started life within politics itself. The Conservative government of the 1990s was the first to use the phrase, which expresses the desire to establish a dividing line between one set of political views and another. The presence of *blue* is significant both because it is the colour of water and because it represents Conservative politics; the phrase recalls the insularity of the conservative Briton, who sees the British Isles as separate from the rest of Europe not only because of the English Channel but also through mindset.

Because politics and journalism have a symbiotic relationship, it comes as no surprise that variant forms of the phrase emerged early on in its life. The cotextual environment of the phrase is well-delineated, and its pragmatic meaning is clear: it expresses not just lack of cooperation between the government

---

13. The Green party (environmentalists) are not included in this representation because their choice of colour is motivated by the metonymical relationship with plant life and the environment.

14. Purple tends to be used by fringe parties, e.g. the (rightwing) UK Independence Party (UKIP), which is represented by imperial purple, or where the colour is culturally significant (for the Scottish National Party, it represents heather).

2nd proofs

and opposition, but that the government (at the time) had no intention of tak-ing opposition views into consideration. *Put* (8) is the most common verb, but may be substituted with another with approximately the same meaning in the context, such as *open* (6), *establish* (2), or *create* (2); and in more than 50% of the occurrences the phrase is followed by *between* and the names of the two oppos-ing political parties. The full phrase can be schematised as [to put] + [clear blue water] + [between] + [party X and party Y] (Example (42)),[15] and this separates the political meaning from the semi-compositional reading of *clear blue water* describing water in its natural state.[16]

(42)   In his blatant attempt to establish his patch of "**clear blue water**" between the Tories and new-style Labour, Mr Portillo also attacked European Union countries.

(43)   I have recently noticed the use of the phrase "**clear red water**" in relation to the Labour Party.

(44)   One aim is to put "**clear red water**" between Labour and the Conservatives in the public's mind.

(45)   One of the few remaining expanses of **clear pink water** between the two par-ties is New Labour's old commitment to introduce a minimum wage.

(46)   Leading Article: **Clear yellow water** – Mr Ashdown's pitch is distinctive, but not exclusive.

(47)   We are tickled by the Liberal Democrats' latest foray into the world of comedy. "Paddy Ashdown," reported the Times yesterday, "sought to put '**clear yellow water**' between himself and the government at the start of the European elec-tion campaign". **Clear yellow water** – geddit? Marvellous.

Examples (43)–(47) show the range of variation found for this idiom. As antici-pated, *red* for Labour (Examples (43)–(44)) and *pink* for New Labour (Example (45)), plus *yellow* for the LibDems (Examples (46)–(47)). Interestingly, the three occurrences of *yellow* all refer to a single incident – the then leader, picking up on what was evidently a term in vogue at the time, coined the creative variant, much to the scorn of the Guardian journalist who penned Example (46). The impres-sion in this example is that the variant was forced, and indeed this is true, as the canonical use of *clear blue water* completely ignores the LibDems (traditionally

viewed as a minority party), juxtaposing only the government and the opposition. The variants in Examples (43)–(45) are less marked because they simply reverse the direction of the expression (i.e. Labour scorning the Conservatives instead of the Conservatives scorning Labour).

A similar phenomenon can be observed in the Italian data for *camicia nera*. The recent political history of Italy is marked by the wearing of shirts – red or black depending on one's political allegiance. But it is *camicia nera* (131 occurrences including plurals) with its relation to the Fascist past, which seems to have left more of a mark on the national psyche. It occurs nearly four times as frequently as *camicia rossa* (35 occurrences, including plurals) in the CorIS corpus. Of course, being potentially compositional, both *camicia nera* and *camicia rossa* can refer to the shirts themselves, but the currency of the idiomatic usage seems to block such a compositional reading: it is more common for the expressions to be used as met-onyms for the shirt wearers. This is also indicated by the frequency with which both variants are found in the plural (just over one third of the total occurrences).[17]

(48)   Per mettere in ombra la vicenda giudiziaria degli operai anarchici emigrati Nicola Sacco e Bartolomeo Vanzetti (accusati, senza prove, di omicidio) fu messo in atto un piano per far finire carcere altri due anarchici innocenti, accusati per l'assassinio di due **camicie nere.**
[To detract from the case of the emigrant anarchists Nicola Sacco and Bartolomeo Vanzetti (accused, without evidence, of murder) a plan was put into action to have another two innocent anarchists jailed for the murder of two blackshirts.]

(49)   Zio Matteo era tornato a parlare dei Mille, delle **Camicie Rosse**, ma il piccolo non capiva.
[Uncle Matteo had gone back to talking about Garibaldi's thousand men, the Red Shirts, but the little boy didn't understand.]

Political shirt-wearers are certainly not a thing of the past, however, even if *cami-cia nera* and *camicia rossa* mainly occur in texts discussing historical events. A new development has emerged in the past few years, which maps the political fervour of the shirt-wearers onto a new colour, green. Green is the colour of the far-right *Lega Nord*, whose homeland "Padania" in the north of Italy is character-ised by its green pasture land. There are 84 occurrences of *camicia verde* (singular and plural) in CorIS – more than twice the number of occurrences of the well-established *camicia rossa*.[18]

---

15. Source: Bank of English 450m version. Ex. 42: Today; Ex. 43: Guardian; Ex. 44: Times; Ex. 45: Independent; Ex. 46: Guardian; Ex. 47: Guardian.

16. The phrase was already formulaic before it was given a new political spin, and although semantically transparent, it is not fully compositional.

---

17. Source: CorIS. Ex. 48: Daily News; Ex. 49: Fiction.

18. Source: CorIS. Ex. 50: Daily News; Ex. 51: Daily News.

(50)   Giura che non è lì che finirà tutta la Storia della Padania, delle **camicie verdi,** del Roma ladrona, il Nord non perdona.
[Swear that it won't be there that the whole thing will end up – the History of Padania, the green shirts, 'the North won't condone/thieving Rome'.]

(51)   Forza Italia copia i leghisti ACQUI TERME. Dopo le camicie verdi, le **camicie azzurre;** dopo la milizia padana, quella italoforzuta.
[Forza Italia imitates the Northern League ACQUI TERME. First green shirts, now blue shirts; first pro-Padania militancy, now that of Forza Italia]

The *blackshirts* were Fascists, while the *redshirts* were Garibaldi's men, those who strove for the unification of Italy. It is the connection with nationhood that shirt-wearing has acquired in Italy which allows it to move into this new domain, coupled to living memories of Fascism associated with the *blackshirts*. The extremist politics of the *Lega Nord* allow it to slot into this context, and the supporters of its politics also actually wear green shirts, thus fuelling the image schema. Two creative variants on this theme are found in the data: both *camicie azzurre* 'blue shirts'. The blue shirt-wearers belong to the extremist wing of *Forza Italia* and are here referred to as *italoforzuti*. (Example (50)). This is evidently an attempt to link what was at the time *Forza Italia*, the political party that Silvio Berlusconi formed when he first entered politics in the 1990s, with the *Lega Nord*, *Forza Italia*'s volatile coalition partner.

6.7   Pandora's box

This chapter began by outlining the problems that can be encountered in finding variants to study, and offered suggestions on how to overcome the limitations of corpus query software. Taking the variants extracted by this general method, it looked at how idiomatic meaning can be enhanced through metaphorical mapping. The kind of metaphorical processes described here are more complex than the ones mentioned in Chapter 5. Class inclusion involves the temporary, ad hoc creation of a semantic class which allows meaning to be transmitted effectively. In this chapter, however, we have seen how entire conceptual schemata, derived from real-world knowledge of what colours are and how they relate to one another, are metaphorically mapped onto meaning.

When conceptual schemata interact, the result is richer and more meaningful. In fact, we have seen that the three main schemata identified for colours – hue, shade and saturation – are *not* exploited on an ad hoc basis, but rather are conventionalised intensifiers, albeit ones that can only be exploited in contexts

where a colour word is expected: gradation of the meaning expressed by the canonical form can be introduced by referring to the gradations present in the colour spectrum, whether they be the hues (tones) themselves, their lightness or darkness (shade), or their intensity (saturation). Additionally, culturally-familiar colour scales, even if they do not reflect "natural" scales, can be used for the same purpose. However, in a break from the tendency observed for the "natural" scales, the change of meaning in these cases is more clearly defined as there is no gradual shifting from one colour to another: a clean break is the only option available.

While these variations are creative and change the meaning to some extent, they stay faithful to the pragmatic meaning of the canonical form. Typically, the cotextual environment mirrors that of the canonical form, maintaining the overall unit of meaning, but with the addition of expressive nuance, and the enhancement of the image schema. Not all variants behave in this way. In the next chapter we will see how highly marked variants differ from the metaphorising variants discussed here and we will see how the remaining types of connotative meaning can be activated in text.

Punning, word play and other linguistic special effects

The kinds of variations that have been examined in the previous chapters have demonstrated that creativity, in order to be communicative, must be solidly grounded in knowledge that is shared between the creator of the variation and his or her intended audience. Yet no single facet of knowledge is sufficient in itself. The phraseological combinations that words enter into are essential for conventional meanings to be expressed, but these interact with knowledge of a word's other combination possibilities, and hence, its other meanings.

Lexical patterns, image schemata, context of use and salience all interact in word play, sometimes with spectacular results. Building on the findings of the preceding chapters, the most creative forms of variation are finally discussed here, and they are pegged to the outstanding categories of associative connotation: paronymy, homophony and symbolism. This chapter introduces the concept of "optimal innovation" (Giora 2003), incorporating it into the analysis of linguistic special effects including humour, punning, contextualising, and even the creative variation of established exploitations.

7.1 Optimal innovation

Some instances of word play work better than others. We are all familiar with the "groan factor" which accompanies "lame" jokes, just as we revel in the novel associations that a good comedian or satirist comes up with. The perception of effective and ineffective word play is closely tied up with aesthetic appreciation, and this relationship has been studied, most notably by Rachel Giora and co-workers (Giora 2003; Giora et al. 2004; Giora & Shuval 2005). Through their investigations, it has come to their notice that there is an important connection between familiarity and creativity. The combination has been labelled *optimal innovation*.

A stimulus would be optimally innovative if it involves
(a) a novel – less or nonsalient – response to a given stimulus, which differs not only quantitatively but primarily qualitatively from the salient response(s) associated with this stimulus and

(b) at the same time, allows for the automatic recoverability of a salient response related to that stimulus so that both responses make sense (e.g., the similarity and difference between them can be assessable. (Giora et al. 2004: 116)

Optimal innovations not only enhance an image schema, but add a new plane of interpretation so that the familiar is recognised but cast in a new light. The re-metaphorisation associated with *red tape* (§5.5) is one example of this, as the tape – benign in itself – is given a life of its own, strangling, binding and otherwise ensnaring its victims. This kind of re-metaphorisation is separated from fully optimal innovation simply by a matter of degree. If variation tendencies can be identified, as they can be for the idiom *red tape*, then even one-off variations adhering to a tendency are not truly innovative and inventive, although they will be perceived as such when an individual comes across them for the first time. Optimal innovations, on the other hand, involve a juxtaposition of familiar and unfamiliar elements in a meaningful and aesthetically pleasing way. Consider all the potential meanings suggested in Example (1).[1]

(1) **Pinko** commie comic Ben Elton will be left **red faced** after being seen **in the pink** in a **red** hot **blue movie** soon to be released.

Here we have an incontestably innovative piece of language, but it is one that is put together from familiar elements: a *pinko commie* is a left winger or moderate communist (from *red* being washed out to *pink*); *to be red-faced* is a conventional means of referring to embarrassment, but also to physical exertion, anger and rage; *to be in the pink* typically means "to be healthy" but also "in the nude"; *red hot*, as well as meaning "extremely hot", can mean "dangerous", "exciting", or "controversial"; a *blue movie* is a "pornographic movie". The juxtaposition of colour word expressions creates surface novelty, and pleasure is also obtained by piecing together the underlying meanings of the items, most of which are polysemous, to reconstruct the writer's message. Successfully recreating the meaning of an elaborate piece of word play like Example (1), which combines surface and underlying features, invokes in the reader what has been aptly labelled the "smugness effect" (Partington 1998: 140): a sense of self-satisfaction in having been clever enough to appreciate all that is being communicated.

What optimal innovation means in linguistic terms is that two meanings – the normal/conventional one and an innovative one – are activated simultaneously. In actual fact, there may well be more than two meanings involved. In the specific case of idiom variants, the meaning of the canonical form alone is actually composed of three separate meanings which function simultaneously – the

---

1. Source: OUP Reading Programme database.

non-compositional "long-word" meaning, the pragmatic implications (semantic prosody) associated with that long-word in its usual cotext, plus the pictorial or conceptual representation of the meaning. In delexicalised canonical forms it is likely that the richer semantic elements such as the image schema are probably latent, but we have already seen that they can be reactivated as a result of relexicalisation, which in turn is triggered by changes to the cotext (Chapter 5) or indeed to its "long-word" core (Chapter 6). On top of all these meanings, a variant brings a new interpretation and also reactivates the latent image schema and the delexicalised compositional meanings, so really what we are dealing with in creativity is the convergence of as many as five meanings – perhaps more if aspects such as phonology and syntax are also considered.

This chapter deals with a range of optimal and near-optimal innovations, explaining the linguistic features that set them apart from the variants that were discussed in Chapters 5 and 6. §7.2 deals with changes to the canonical form which come about either by translation or by substitution with a phonetically-related word, and some special cases where expansion of the colour word has a contextualising and focusing effect. Changes to the cotextual environment which prompt a context-specific reinterpretation of the meaning of the canonical form are dealt with in §7.3 and §7.4. By way of summary and conclusion, §7.5 will show where the limits to optimal innovation lie in a final case study which features the whole spectrum of variation, from unmarked to optimally innovative and beyond.

7.2 Changes to the canonical form

One of the most obvious kinds of word play is substitution of one or more of an idiom's constituents. Chapters 5 and 6 looked in detail at the mechanisms that operate here, but did not touch on two particular forms of variation: substitution of a key word with a translated form, and substitution with a rhyming or near-homophonous word. These kinds of variation are the most marked and also the rarest observed in the data, with only a handful of instantiations found in all the idioms studied. The important point is that in spite of their rarity, they can be found both in corpus data and on the Internet by carrying out multiple queries of the kind outlined in §6.2.

7.2.1 Translation

Two types of translation were found in the data, neither of which are common, but for different reasons. The first kind to be discussed is the back-translation or

nativisation of foreign words. In the English data, this was only encountered in the concordances for *éminence grise* (even though *bête noire* and *carte blanche* are also borrowings) and in the Italian data, no translated variants were found at all. Sixteen of the 103 concordances (15%) extracted for *éminence grise* appear in the Anglicised form *grey eminence*. However, apart from the change of language, there are no discernable differences between the French form and the English: the cotextual features of *éminence grise* remain unaltered in spite of the translation, leading us to suppose that the reasons underlying the translation are concerned with the accessibility of the phrase in a world that is increasingly dominated by the English language.

The inverse process, translating an expression wholly or partially into a foreign language, produces much more interesting semantic effects, although it is equally uncommon, if not more so. Two examples of substitution by translation were found for *the pot calling the kettle black* (Examples (2)–(3)),[2] which undergoes a variety of other variation types including ellipsis, expansion, reversal of components, substitution, and personification (see Philip 2008).

(2)   "It is downright hypocritical for the *French* to ban *British* beef produced under some of the world's most stringent safety regulations." Mr Gill, president of the National Farmers' Union, added: "It is time **the pot** stopped **calling the kettle 'noir'**".

(3)   Last year, a top London food critic awarded *null punkte* for every *German* restaurant visited. <p> Surely a case of **the pot calling the kettle schwarz.**

In both cases, the schema of *the pot calling the kettle black* is adhered to: there are two participants (*pot* and *kettle*), both are at fault (*black*), but one has the gall to criticise the other for the fault which they share. Hypocrisy is expressed, but in a roundabout manner (it is certainly more polite say that somebody's actions are *like the pot calling the kettle black* than to accuse them outright of being a *hypocrite*).

We can see that the language used for the translation of *black* corresponds to the nationality being referred to in the preceding cotext: *French* in Example (2), and *German* in Example (3). The effect is to introduce a cultural connotation by lexical analogy (Kerbrat-Orecchioni 1977: 117; cf. §3.6.2), specifically by the use of a paronym which activates the schema of "Frenchness" or "Germanness" through the presence of the foreign word.

Typically, paronymy involves 'identity of root, but difference of syntactic category' (Cruse 1986: 55), such as *black:blacken*. The term has wider application, however, including that used here (after Kerbrat-Orecchioni 1977), in which the

base of the paronym is not a lexical item, but the entire language that it forms part of. Thus a foreign word can conjure up the whole conceptual schema associated with the foreign nation(s) in which the language is used.

> Le signifiant "Panzani" rappelle paronymiquement, non un mot particulier, mais tout un ensemble de mots italiens … il s'agit de valoriser le produit alimentaire en indiquant qu'il possède le pouvoir quasi magique de transporter à peu de frais son consommateur dans le pays de ses rêves.   (Kerbrat-Orecchioni 1977: 117)
> [The signifier "Panzani" paronymically conjures up not a particular word, but a whole range of Italian words … it gives added value to the food product, suggesting that it has the almost magical power to transport its consumer, at minimal cost, to the country of his dreams.]

Examples (2) and (3) both exploit another facet of *the pot calling the kettle black*, which is personification. In fact, it can be seen that it is the nationality adjective, not the nation, which appears in the cotext, and through a process of metonymy, the nationality stands for the people of that nationality, even if not explicitly stated (as in Example (4)). Personification is the dominant variation tendency for *the pot calling the kettle black* (Philip 2008), suggested by the image schema and fuelled by the verbal process it describes. Example (4), although a paraphrase of the idiom, is easily reconnected to the canonical form because of the juxtaposition of *pot* and *kettle* (the singular form of *pot* in this collocation triggering the idiomatic meaning alone) combined with the verbal process in a preceding cotext where hypocrisy is clearly signalled.[3]

(4)   I've been known to accost people in a public place with a cigarette hanging out of my mouth, saying: "Do you realise this is a no-smoking area?" **Hello pot, my name's kettle**.

If the *pot* (criticiser) is French, then it would call the English kettle *noir*, not *black*, surely? The justification for *schwartz* in Example (3), however, operates at a greater semantic distance, because it is the English speaker abroad (*a top London food critic*) who is using the native language, not just *schwartz* but also by awarding *null punkte* 'no/zero points', which sets off a range of cultural connotations of its own.[4] The image schema is also forced out of kilter because the participants in the

2nd proofs

context of situation are usually of equal status, but here we have a critic (in a superior role) criticising chefs (subordinates). Later in this chapter we will see that changes to the context of situation can destabilise delexical, idiomatic meaning, although for now we are still on relatively safe ground.

In the previous chapter, we saw how changes to colour words in colour-word expressions made use of the variety of meanings that are culturally attributed to colours (as entities, not words), and how metaphorical conceptualisations could be used expressively. In the case of translation as illustrated in this section, however, the colour does not change, nor does the meaning of the idiom, but a particular kind of cultural connotation is activated. The translated variants in Examples (2) and (3) activate a very rich form of connotation, and this interacts with the idiom's full meaning, enhancing the image schema and metaphorical entailments. This area of variation appears an exciting one to explore, but there were very few instances found in the data (five instantiations in total), making it impossible and inadvisable to make generalisations about their communicative purpose and its effectiveness. However, that cultural connotations are being activated, and deliberately so, is irrefutable.

7.2.2   Variation exploiting rhyme and assonance

Connotations that are activated through translation are partly influenced by the different phonetic traits of the translated words, which evoke the language of translation and its culture as a whole. Another way in which the sounds of substituted parts of idioms can contribute to meaning is when there is rhyme or assonance present. This is probably a more familiar kind of pun, and is a type which is often cited in the literature: for instance, Giora et al. (2004: 118) list a range of shop names which are based on puns, including *body and sole* (a shoe shop), *curl up and dye* (a hair salon), and *sofa so good* (a furniture store). There is no shortage of such language play to be found, and it is certainly not limited to English-speaking contexts. In Italy, I have noticed an ice-cream parlour called *Peter Panna* (*panna* is 'cream'), and, to my great personal delight, *Scarpe Diem*, a shoe shop whose name fuses *scarpe* 'shoes' with the well-known Latin proverb *carpe diem* ('seize the day').

However much these examples please us, they are decontextualised linguistically, so they can tell us little about how punning can contribute to textual meanings; and while they are prominent, they are not particularly frequent. Variation which is based on similar-sounding words defies probabilistic computation even more effectively than that which is based on attributive semantic sets does, but can be found in a corpus through repeated searches in which variable slots are

Table 7.1   Variation exploiting rhyme or assonance

Variant	Canonical form	Additional information
White Lie-kea	white lie	IKEA furniture store
greasy eminence	eminence grise	unctuous character
Post calling the Kestle black	the pot calling the kettle black	Post Office; proper name (Kestle)
bestia nerazzurra	bestia nera [villain]	blue and black (colours of Atalanta FC)
La signora in giallorosso	La signora in giallo [the lady in yellow]*	red and yellow (colours of Roma FC)
ira rossonera	ira nera [black rage]	red and black (colours of Milan FC)
periodo nerazzurra	periodo nero [bad time]	blue and black (colours of Atalanta FC)

* The TV serial known in the English-speaking world as 'Murder She Wrote'.

left open. Table 7.1 lists the few examples found in the English and Italian data examined together with the canonical form from which they are derived and the additional information referenced in the pun.

Substitution by rhyme or assonance forms an important category of contextualising variation, even if examples are hard to come by. Due to the difficulty in finding variant forms in general, many existing corpus-based studies (including Partington 1998; Herrera Soler & White 2007) have resorted to manual analysis, and have typically focused on newspaper headlines which are easy to locate (they are coded separately) and are known to represent a rich source of creative variants. While it is true that marked variations are often found in newspaper titles and headlines, many are not, and it can be interesting to observe how the creativity displayed in headlines differs from that used in other parts of a text, as well as to compare marked variation with the unmarked forms that were discussed in Chapter 5.

It can be appreciated from a look at Table 7.1 that substitution by near-homophones is an extremely marked, and rare, form of variation. Like translated elements of an idiom, a phonetically-induced pun brings outside information into play. If exploiting true homophony, it is relatively easy to coin and comprehend (e.g. *Subtitled **Greene with Envy**, this episode sees the debut of Alan Alda as Dr Gabriel **Greene***, Table 3.5 above §3.6.4). The substitution of a similar word for a similar meaning-referent requires a little more effort to produce (unless of course it is institutionalised, e.g. *red **rag** to a bull / red **flag** before a bull*), but the phonetic

resemblance is tied to semantic resemblance. When the examples have a phonetic resemblance with the canonical constituent, but no semantic resemblance, the meaning of the underlying idiom is understood thanks to previous use of the expression in context. This can be classed as a form of intertextual connotation (Kerbrat-Orecchioni 1977: cf. §3.6.2), while true homonyms connote by lexical analogy (ibid.), and similar forms by a signifier-signified relationship (ibid.).

Some of the examples shown in Table 7.1 maintain a semantic relation to the canonical constituent they substitute while others do not. In the Italian data, it can be seen that the institutionalised fusion of colour words to refer metonymi-cally to sports teams (Roma FC is *giallorosso*; Juventus is *bianconero*) allows puns to be created with ease. All the Italian examples occur in the football reports of national broadsheets, and manage to convey evaluative information about the team by incorporating its colours into an idiom which normally features only one of those colours. These Italian puns are really examples of lexical expansion (see §7.3, below) rather than substitution per se: the substituted colour word in these cases specifies the referent, but does so in a less intrusive way than in puns where fully-fledged substitution takes place.

The English data is more problematic, because there is no semantic relation-ship to accompany the phonetic one in the examples found. The result is deliber-ate word play through the activation of double-meanings, fusing an idiom with the person(s) or thing(s) that it refers to. This is a sophisticated technique which allows information to be presented in a particularly concentrated way. Unsurpris-ingly, then, puns like these are found in newspaper headlines, where condensa-tion of information is paramount as a space-saving device. Puns in a headline also serve an important pragmatic function in that word play attracts a reader's attention thus increasing the chances of an article being read. They operate as cataphoric devices, with the relevance of the pun, i.e. the reason for particular substitutions, unravelled in the ensuing text: *White Lie-kea* and *Post calling the Kestle black* both occur as headlines, as do many more of the idiom variants that will be discussed later on in this chapter. The remaining pun, *greasy eminence*, does indeed refer to an *éminence grise*, but with the addition of an evaluative slant. The political figure is, however, played by an actor, rather than being the real thing (Example 5).[5]

(5)   Yet Trevor Eve played him sexy and slimy at the same time, and that's not easy. <p> Ian Bannen was great as the party's **greasy eminence**; Anton Lessor made you understand why the crypto-gay aide hated his boss.

---

5.   Source: Bank of English 450m version. Today.

Greasy eminence is an example of a false paronym (Cruse 1986: 131) as *greasy* connects phonically with an unrelated base (*grise*). However the relative rarity of the underlying expression (*eminence* occurs only 112 times, 87 of which col-locate with *grise*), especially if compared to *greasy*, the dual meaning may not be appreciated by all readers.

7.3   Expansion, addition and combination

The Italian sports puns in Table 7.1 demonstrate that additional referential mean-ing can be introduced when the colour word is expanded. In §6.3 too we saw that *pea-green* can take the place of *green* in the idiom *green with envy*, causing a change of emphasis, but not a change of referent. The effect that expansion of an idiom constituent can have on the complete meaning conveyed by the vari-ant differs according to what kind of expressions are included in the expansion, and how the expansion fits into the surrounding cotext. In this section we will look at expansion (*green with envy* → **pea** *green with envy*), addition (*see red* → *see red*, **white and blue**), and combination (*caught red-handed*; *paint the town red* → *caught painting the town red-**handed***), which were the types identified in the data examined.

Colour word expansion is here defined as the extension of the colour word by pre- or post-modification in order to specify the precise hue, shade or satura-tion of the colour indicated. The examples found in the data have been discussed already in Chapter 6, so only a brief summary of the findings is required here.

As far as the lexical set of colour words is concerned, expansion is used for intensification and mitigation. When the precise hue of the colour is specified, e.g. **pea** *green* / **emerald** *green with envy* (§6.3), it makes reference to the focal point of the colour category (Berlin & Kay 1969), i.e. the prototypical or "best example" of that colour in the language concerned. Expanding the colour lends emphatic weight to the meaning already expressed by the canonical form, increasing the salience of the colour word in the delexical string. Specification of the intensity (saturation) of the colour has much the same effect, although it can also be used effectively as a downtoning device: *His prose is at its **purplest**… / If you like your prose **purplish**…* (§6.5).

The addition found in the data involved conventional colour sequencing (e.g. *red, white and blue*), each time alluding to the colours of a flag. The colours of a flag have symbolic meaning, and for this reason the variants found in this category activate associative connotations as a result of their referential relation-ship (Kerbrat-Orecchioni 1977: cf. Table 3.5, §3.6.2). In particular, the referential relationship is a form of metonymy whereby the flag represents nationhood, and

the effect of a variant following this pattern is to evoke this cultural connotation in addition to the meaning conveyed normally by the canonical form. Possibly significant, though difficult to gauge from the few examples found, is that the colour found in the canonical form must be present, and that it must appear either as the first or as the last in the string, presumably so as not to sever the phraseological connection. Examples (6) and (7) illustrate the creative addition of colour words.[6]

(6)   Stunning *Miss Ireland* Emir Holohan-Doyle wraps our national flag around her – hoping to make her Miss World rivals **green, white and orange with envy**!

(7)   *REPUBLICANS* **saw red, white and blue** at the Spring Hill Fair yesterday, when they spied *monarchists* and *Union Jack wavers* running three stalls.

In Example (6), Miss Ireland wraps herself in her national flag, whose colours are green, white and orange; and as the winner of the beauty contest, she leaves the other contestants *green with envy*. Example (7), from the Australian press, finds *Republicans* contrasted with *monarchists* and *Union Jack wavers*, who represent opposing political factions (pro- and anti-Commonwealth respectively). The Republicans' anger is caused by the presence of the Union Jack flag, whose colours are *red, white and blue*, and supporters of the (British) monarchy at an event which, if not actually anti-Commonwealth, is certainly intended to celebrate independent nationhood.

The combination of idioms can involve juxtaposition (Example (8)) or fusion (Example (9)), the choice probably depending on the desire to avoid ambiguity: *light green fingers* (cf. Example (8)) suggests neither of the idiomatic meanings rather than both together.[7] Irrespectively of whether they are juxtaposed or fused, each idiom expresses its full meaning. The distinct idiomatic meanings interact, each enriching the other by highlighting its contextually-relevant features through a process of class inclusion. Thus the overall meaning is enhanced and refocused onto its context of use, with less relevant features of the individual idioms being played down.

(8)   <h> **Light fingered green fingers** </h> <p> Report by Tony Henderson <p> Light fingered – as well as green fingered – visitors are making off with *valuable cuttings and plants* from *gardens* which open to the public.

---

6.   Source: Bank of English 329m version. Ex. 6: Sun; Ex. 7: Oznews.

7.   Source: Bank of English 450m version. Ex. 8: British Magazines; Ex. 9: Oznews.

(9)   **Caught painting town red-handed** By Peter Hansen. *HIDDEN cameras* planted on Brisbane suburban train stations have been silently *capturing on videotape* some of Brisbane's worst *vandals*, as well as a variety of *thieves*, *graffiti artists and other criminals*.

Example (8), a newspaper headline, combines two idioms which have the word *finger* in common, though different meanings of *finger* are meant. To be *light-fingered* is to have the propensity to steal: *finger* refers metonymically to *hand*; *hand* refers metaphorically to *taking* (MacArthur 2005). *Green fingers* (in the US, *a green thumb*) is the ability to make plants grow. In juxtaposition, one would assume that theft is involved, and that the thieves are gardeners. The first line of the article has been included to provide the context which resolves what is in effect an innovative sequence. It is necessary because the idioms do not normally collocate with each other (idioms on the whole do not collocate with other idioms), and idioms featuring salient words of any type, such as colours (but also body parts, e.g. *fingers*) also normally avoid collocating with members of the same lexical sets, for reasons which will become apparent in §7.5.1. The article does indeed speak of a theft, but whether or not it is committed by gardeners can only be inferred, as the thieves have not yet been apprehended. All that is known is that the stolen goods are plants and cuttings, not from garden centres, but from public gardens. If one takes a cutting of a plant, one usually also knows what to do with it to make it grow, so the thieves probably are *green-fingered*. The pun is successful because it can be decoded, and the decoding is confirmed in the following text without being made too explicit.

Example (9) illustrates a fused form which combines *paint the town red* with *caught red-handed*. The idioms share the colour word *red*, but its meaning is different in both idioms, and the idioms also have very different contexts of situation and image schemata. The reason why the paint in *paint the town red*, should be red in particular is a mystery, whereas *red* in *caught red-handed* is obvious (but misleading, as we saw in §4.3). *Paint the town red* is an emphatic form of *go out on the town*, meaning to go out and have fun (usually drinking, socialising, and dancing), whose pragmatic inferences contrast starkly with being *caught red-handed*. As a headline, the fusion creates a semantic anomaly which has to be resolved. Were revellers arrested? If so, why? What crime (*caught red-handed*) were they committing? The first sentence of the article explains the reasoning behind the juxtaposition. The people *caught red-handed* were indeed criminals (*vandals, … thieves, graffiti artists and other criminals*), and they were indeed *painting the town red*, both literally and metaphorically, as they were literally (spray-)painting graffiti all over the town, but presumably doing so as a form of entertainment. In this example too, there is a third, contextualised meaning

which is activated, and this is lexically and semantically related to elements in the two idioms which have been fused.

The feature that all the idiom variants discussed both in this chapter and in the preceding ones is that although the core of the unit of meaning can withstand variation, it does so within a relatively stable cotextual environment. The collocates may undergo a degree of re-metaphorisation, as was seen to occur with *red tape* (§5.5), but on the whole, the variant expresses the same cotextual preferences as the canonical form, adheres to the same image schema (which is enhanced, but not actually modified), and expresses the same pragmatic meaning. The novel element has to be interpreted and related to the cotext and to the underlying canonical form, and meaning can be reprocessed, relexicalised, and re-metaphorised, but the communicative function does not change. Even in the fused idioms (Examples (8)–(9)), the full weight of the pragmatic meaning is realised while the extra meanings are created. But this does not always hold true. We will now look at what happens to the meaning of canonical forms when their "canonical cotext" undergoes variation: §7.4 introduces the main aspects of explicit and implicit meaning in context, while §7.5 investigates them in a case study of the idiom *see red*.

7.4 Canonical forms in non-canonical cotext

Idioms are defined as such because their non-compositional meaning does not match their compositional meaning. Given this, it is evident that polysemy is always lurking just below the surface. Up to this point, we have focused on changes to the composition of the idiom, to its internal configuration. But it is not necessary for the idiom to change its internal composition for it to change its meaning. Other factors which can destabilise the delexical idiomatic meaning include the addition or change of words in the cotextual environment. These can trigger off salient meanings within the idiomatic string simply by their presence, specifically by their juxtaposition, as we have just seen (Examples (6)–(9)).

New meanings can be signalled explicitly in text, but they need not be, and in the vast majority of the examples of word play encountered in the data, they are not. In these cases, it is up to the reader to make the connection between the contextually appropriate meaning and the canonical expression. When additional meanings are not signalled, there is always the risk that some of the finer points that a writer wishes to convey will fail to reach their target. In some cases, this simply means that the "cherry-on-the-cake" aspect of word play is lost while the basic, underlying meaning is still conveyed. But it can also happen that the

meaning conveyed is diminished or even absent if the implicit meaning cannot be accessed. This is what Gerrig & Gibbs (1988) call the "dark side" of creativity, because 'the intimacy it creates can serve as an agent of exclusion' (ibid.: 8). Yet sometimes the reverse happens, and the signalling is so strong that the fun and "smugness effect" (Partington 1998: 140) of word play is spoiled. These three types of contextually-induced word play are now illustrated in turn (Examples (10)–(12)).[8]

(10) A view clearly shared by Lennon, whose crisp *strike* from all of 20 yards in the 71st minute came like a **bolt from the blue**. It was too much for Chalmers in the visitors' *goal* to handle and the *ball* landed at the feet of Crawford who swept into the *net*.

(11) Now is the time for *Celtic* to put matters right and stop **the air turning**, dare I say it, **blue** at *Parkhead*! It's time the *Bhoys* made the other side of *Glasgow* **green with envy**!

(12) A fanatical *Star Trek*, *Alien* and *Star Wars* fan, Tim has a collection of *space age* memorabilia which would make most *extra-terrestrials* **green with envy** – if they weren't that colour already.

Bolt from the blue (Example (10)) is read idiomatically ("surprise", cf. §2.3.1), but additional knowledge that in football, a powerful or forceful strike is also known as a *bolt* allows the savvy reader to appreciate the double meaning. Absence of this knowledge does not diminish the information being conveyed, however, because the contextual cues fill in any gap in knowledge: it is clear that the striker is shooting at the goal, and the conventional reading of *bolt from the blue* tells us how the action was perceived. Some might even argue that a double meaning is not intended.

Full appreciation of Example (11) depends on the reader being in possession of a considerable amount of cultural knowledge, because if not, the evident jokiness (indicated by exclamation marks and "dare I say it") is perceived, but what it actually refers to is not. The reader has to know what relationship *Celtic*, *Parkhead*, *the Bhoys* and *Glasgow* have to one another. Only from this premise is it possible to see the significance of *blue* and *green* in the idioms *turn the air blue* and *green with envy* respectively. The connection is fairly straightforward for anyone familiar with the Scottish football scene, making this a very successful pun. But if the reader is not familiar with the sociocultural background, the whole piece comes over as nonsense. Readers unfamiliar with the context can compare

---

8. Source: Bank of English 450m version. Ex. 10: Times; Ex. 11: Sun; Ex. 12: Oznews.

the "before" and "after" effect by consulting the footnote provided;[9] those in possession of the knowledge will already be feeling the "smugness effect". It will be appreciated that it is the cultural knowledge that creates the pun, and when that knowledge is missing the communication is ineffective. If the reader fails to pick up on all the intended meanings, *turn the air blue* and *green with envy* can only be understood with their canonical meanings, whose relevance to the context is at best tenuous.

The opposite effect can be found when the basis for a pun is so explicitly signalled that no play is left in the word play. Example (12) is certainly not an example of optimal innovation, as it leaves the reader rather deflated. The reader is bombarded with contextual information that regards space – *Star Trek, Alien* and *Star Wars*, *space age memorabilia* and *extra-terrestrials*. *Green with envy* would be acceptable in this context, because the reader could make the connection between the conventional representation of aliens – Martians in particular – as *green*, and the same word used delexically within the idiom. But the motivation for the pun is then spelled out (*if they weren't that colour already*), thus spoiling the fun of finding the connection and the satisfaction that accompanies the appreciation of a double meaning.

7.5   More on *seeing red/blue/green…*

Despite sharing the same etymology as *red rag to a bull*, *see red* is not really used to report wild rage, as the associated image schema might suggest. But another interesting aspect of this idiom is that its most common use is within double-meanings, with only about a quarter of its occurrences referring to the state of anger

---

9.   The necessary cultural background is as follows: Celtic FC is one of Glasgow's two largest football clubs, and plays at Parkhead stadium. Being originally founded as the city's catholic team (representing the Irish immigrant population) its colours are green and white. The team is colloquially referred to *the Bhoys* /bʰoiz/ to reflect the Gaelic-influenced, aspirated pronunciation. Rangers, Glasgow's other large team, was founded as the protestant club, and its stadium (Ibrox) is located, literally, on "the other side of the city". Rangers players wear a blue strip. The catholic/protestant divisions were rigidly in place until as recently as 20 years ago, with Celtic fielding only catholic players and Rangers only protestants. This has changed, but the fan base is still largely divided; the religious division carries with it further implications regarding social class and nationality/ethnicity, and so the rivalry exists on several inter-related planes. All these elements combine to support the pun. Celtic has to win the "Old Firm" match against Rangers to save face. If they fail to do so, the atmosphere at their stadium will be filled with swearing (*turn the air blue*), *blue* also being the colour of their arch-rivals. What they should do instead is play so well that Rangers fans would be envious – so envious, in fact, that they would turn *green* – Celtic colour – *with envy*.

---

alone. This case study will focus on the different kinds of cotextual changes that occur, and what effect they have on the meaning of the idiomatic core. In order to do so, the core needs to be outlined first.

See red is different from other expressions of anger in that it centres on the contrast between a mild-tempered person and his or her unusually vehement response to a given situation. Examples (13) and (14) illustrate typical uses of canonical *see red* in the corpus data, complete with causative *make…* (Example (13)), the downtoner *just* (Example (14)), comment on the emotion felt (Example (13)), and indication of the reason for anger, *because …* (Example (14)), all of which are features which help to construct the semantic prosody of "surprise caused by an uncharacteristically strong angry reaction to a particular event or situation".[10]

(13)   Mike: 'I was seeing this girl before, and once I was at a bar and I saw her giving out her number to another guy. I **saw red** thinking if she had sex with him, I'd kill them in bed. I grabbed her arm and said, "Don't you ever do that again in front of me." She hit me in the head and ran out. *My own blow-up bothers me. I never did that before. Could it happen again*?'

(14)   I think John just **saw red** *because the clamping has been going on for some time.*

7.5.1   Juxtaposition, delexicalisation and salience

One of the interesting features to be found in word play associated with *see red* is the appearance of colour words in the immediate cotext, where they are not normally found. It is, in fact, quite safe to say that this kind of juxtaposition is normally avoided in the language as a whole because it generates ambiguity. Hoey (2005) discusses this general principle as being related to the priming of lexical patternings and their conventionally-attributed meanings.

In the first place, it is now a given, thanks largely to corpus lexicography, that 'co-hyponyms and synonyms differ with respect to their collocations, semantic associations and colligations' (ibid.: 13). It is also true of members of a lexical set such as the one we are discussing here: colour words too are co-hyponyms: although they fit less neatly into the hierarchical description "X is a kind of Y". With colours, the relationship is one of greater specificity, not hierarchy (red *is* a colour, not *a kind of* colour). Added to this is that 'when a word is polysemous, the collocations, semantic associations and colligations of one sense of the words

---

10.   Source: Bank of English 329m version. Ex. 13: US Books; Ex. 14: Today.

differ from those of its other senses' (ibid.). This lies behind the avoidance, in normal text, of juxtaposing two colour words (or indeed any other polysemous terms belonging to the same lexical set). The phraseological environment of a colour word when used in an idiom is different from the phraseological environment its literal, referential meaning attracts. When two words belonging to the same lexical set are found in close proximity, the probability that they will be interpreted salient-meaning first is greatly increased, *irrespective of phraseology*. Tension ensues between the salient meaning and the phraseological one. This is undesirable in unmarked discourse, but ideal in word play. The result is the creation of a contextually-motivated pun such as those illustrated in Examples (15)–(17).[11]

(15)   **Green sees red** over card GLADSTONE'S Green candidate Cedric Williams said yesterday he was cheated out of votes by a "questionable" how-to-vote card infuriated Gladstone's National Party.

(16)   Colleagues **see red** when medics turns out to be **black**.

(17)   IBM no longer **sees red** when called "**Big Blue**."

The puns arise because of the forced reprocessing of the idiom which unavoidably happens when the normal process of delexicalisation is interfered with. In Example (15), the *Green* in question is a member of the Green Party. There is already a considerable amount of metonymical activity going on in this context, even before the pun kicks in, with the following chained metonymy: *green* stands for *environment* stands for *environmentally-oriented individuals* stands for *environmentally-oriented political party* stands for *parliamentary representative of that party* stands for *individual* (who is therefore a member of the party and, by extension, concerned about the safeguarding of the environment). Now, to report the anger of a politician or of an environmentalist there are alternative expressions available – "MP sees red", "Angry Green", and so on – but a deliberate choice has been made to exploit the tension that the juxtaposition the two colours generates. The salient meanings of *green* and *red* (the colours) are activated together with the metonymy unpacked above and the idiomatic expression of anger. Yet in spite of the complexity of the figurative language involved here, the pun is actually very easy to process, with no need for the activation of new image schemata or conceptualisations. The same is basically true of Example (16), where the racial sense of *black* is contrasted with the metaphorical *red* in the idiom and the salient meanings of both colours, and of Example (17) where again alternative wordings are available to express the sense of anger and refer to the business in question,

---

11.   Source: Bank of English 329m version. Ex. 15: Oznews; Ex. 16: British Magazines; Ex. 17: US News.

but these have been rejected in favour the pun. While these three examples are neither particularly sophisticated nor innovative, they all achieve their communicative purpose, which is to express meaning and to contextualise that meaning simultaneously.

Juxtaposition of colour words does not always result in punning, but if one of those colour words is normally delexicalised and occurs within an idiomatic string, it does. If the cotextual colour word is also used in a figurative sense, it is certain that reprocessing will occur, and in such cases, it will result in a more satisfying effect than if the colour were simply used referentially. If, on the other hand, the colour terms are both used in their compositional, referential sense, there is no need for any reprocessing to occur, and hence there is no word play.

7.5.2   Implicit meaning

Juxtaposition of colour words does not make for the most satisfactory of puns, mainly because the source of the double meaning is too obvious. More intriguing, because more intellectually demanding, are those puns which essentially refer to situations of the same nature while leaving the connections implicit. This is extremely common in *see red*, in particular in sports reporting where players on the pitch have been sent off (signalled by being shown a *red card*). Here, context is all. The inability to connect *see red* to the context of being shown a *red card* in a team game means that its implications cannot be appreciated, especially since it tends to arouse anger in the player being sent off (as well as sometimes being the penalty paid for an angry outburst in the first place).

Although no claim can be made that the reader always "gets" implicit word play, in the case of this particular idiom it can be taken for granted. Many occurrences of punning *see red* in the corpus are found in the sports sections of newspapers, both broadsheet and tabloid. People who have no interest in sport are unlikely to read this section, and it is also inconceivable that any British-newspaper-reading individual has no inkling of a *red card*'s value in sport, however detached his or her interest in sport is. So if a reader comes across *see red* in the sports section, the chances are that it is in a football match report when a player was sent off, or *red-carded*. The player is shown a *red card*, so, in literal sense, *sees red*. An interesting twist on what could be a rather banal and tired use, however, is that is seems to be used only when the player concerned refuses to bow out gracefully, deciding to continue to vent his anger on the referee or on other players (Example (18)). In fact, those players who have been booked (shown a *yellow card*) but then argue with the referee about the decision risk getting a *red card* – which is then reported as *seeing red* (Example (19)). So despite the context-dependency

of the expression, it remains true to the general sense of anger that is associated with the canonical form.[12]

(18) Tosh was booked for arguing the decision, then **saw red** when he continued his protests in the tunnel at the break.

(19) Booked earlier for dissent, he then delivered a volley of abuse at the referee after being awarded a foul and **saw red**.

Even when we move outside of sports contexts and into journalism at large (the vast majority of these puns are found in journalism) we find again that it is the context which signals that a double-meaning is intended, and provides the necessary background to understand what the additional significance of "red" might be. When a reader comes across a pun in written discourse, he or she already has some contextual information at hand – the genre, the subject matter, the political slant, perhaps even a photograph. So although in theory any meaning can be triggered, in practice the principle of least effort generally applies, for no better reason than because the writer intends to communicate something to the readership, so is unlikely to make the meaning entirely opaque.

See red, meaning simultaneously *red card* and *anger*, appears exclusively in football commentaries, but other double meanings are triggered in relation to different subject matter. For instance, the idiom also occurs in economics and business texts, a context in which *red* is inextricably linked to debt. Here, the presence of *see red*, alongside *red ink* and *into the red* (or other, more direct expressions of debt) should not be surprising. The anger sense of *see red* fits nicely with the link between *red* and *debt*, though for anger to be induced, the debt should presumably be considerable (Example (20)) and, in keeping with the semantic prosody, be unexpected for some reason (Examples (20) and (21)).[13] Explicit reference to surprise appears not to be necessary when the amount of the debt itself is unbelievable (Example (20)); in other circumstances, such as that reported in Example (21), the surprise has to be made explicit so that *see red*'s appearance is properly justified.

(20) A CUSTOMER **saw red** when her bank said she was *overdrawn* – by *£40 million*.

(21) *INTEREST rate* **blues** had the stock market **seeing red** yesterday as traders pondered Monday's *surprise move*.

---

12. Source: Bank of English 329m version. Ex. 18: Sun; Ex. 19: Independent.

13. Source: Bank of English 329m version. Ex. 20: Today; Ex. 21: Today.

Incidentally, it is thanks to there being two figurative meanings of *see red* at work in Example (21), combined with the figurative meaning of *blue* and the salient meanings of both colours, that it is more complex, and more satisfying, than the colour word juxtapositions discussed in §7.4.1 (Examples (15)–(17)).

Another contextualised meaning of *see red* connects anger with the red of Communism. The relationship between *red* and Communism is a symbolic one which extends through chained metonymy just like *green* was seen to do (§7.4.1), i.e. from *political party* to *political regime* to *country under that regime* to *nationals of that country*. This provides evident scope for idiom exploitation, not just with *see red* but with any idiom containing the word *red* e.g. **Labour are socialists red in tooth and claw**. Examples (22)–(24) demonstrate that Communist connotation, brought about through the referential relationship within the symbolic meaning (Kerbrat-Orecchioni 1977: cf. Table 3.5, §3.6.2).[14]

(22) <h> European Business: *Russians* **see red** over conference calls – </h> *Eastern firms* are flocking to London seeking partners to help them grow. But, as Yevgenia Borisova reports, it can lead to disappointment and mistrust By YEVGENIA BORISOVA <p> *RUSSIAN businessmen* are paying between £6,000 and £10,900 each to attend business conferences in London or other west European cities, in the belief that they will establish direct contacts with the most prominent and reliable *western* banks and other financial, industrial and legal organisations. They expect that the conference, organised by London-based companies, will help them find western partners for their business. Instead, the Russians claim, they simply get entertainment, excursions and luxurious dinners.

(23) LONDON Monarchs were **seeing red** yesterday when they signed *ex-USSR* rugby international Oleg Sapega, the 24-year-old defensive lineman who played in the World League last year for the Raleigh-Durham Skyhawks.

(24) <h> **Seeing red** </h> Six years after *Hungary* embraced *democracy*, *communist* symbols and leaders have returned as kitsch, from decor for restaurants to tourist souvenirs. <p>

The context surrounding Examples (22)–(24) features lexis which triggers Communist associations and thus recontextualises *see red*. While in all cases this derives from the chained metonymy which connects *red* with a national of a Communist state, Example (23) adds to this the colours of two American football

---

14. Source: Bank of English 450m version. Ex. 22: Guardian; Ex. 23: Today; Ex. 24: Oznews.

teams, both of whose team colours included red.[15] It is one of only two instances (out of 14) of "Communist" *see red* puns not to occur in a political context, suggesting that the relationship generally requires more solid contextual support in the discourse topic as well as in the local cotext.

As well as recontextualising *see red*, the association of Communism with the colour *red* signals the presence of a second meaning in the text, to be read in conjunction with – or indeed *instead of* – the idiomatic "anger" sense. In fact, unlike the sports and economics *see red* puns, the idiomatic meaning here does not always seem to be intended or even relevant to the contexts of use, as can be appreciated in Example (24). The apparently irrelevant presence of *see red* marks the limit between what is word play and what is an empty pun, defined as a device which attracts notice but fails to offer a meaningful contribution to the text in which it occurs.

The other context-specific, cotextually-modified instances of *see red* were not absolutely faithful to the full unit of meaning associated with its canonical form. This is most notable in the absence of causatives and downtoners, but the sense that usually mild people find themselves angered also fades out when double meanings are set up. However, at least the generic meaning, anger, was seen to be essential to justify the presence of the idiom. This is not the case in Examples (23)–(24), where the connection between a meaningful idiom and the meaning to be conveyed has been severed. In other words, the idiomatic core remains in its canonical form but the extended unit of meaning has undergone variation. This is important to note, as in the variations presented up to this point, the meaning expressed by a non-canonical form was sustained by the "canonical cotext". We now see the canonical phrase acting as a focal point, counterbalancing the destabilising effect of a "non-canonical cotext".

7.5.3   Ringing the changes: Variation inside and out

Finally, we come to a variation in which both the canonical form and its normal cotext undergo substitution of some kind. This is extraordinarily delicate because the linchpins which normally hold idiomatic meaning in place are being attacked from both sides, and there is a particularly high risk of meaning collapsing in on itself like a house of cards. Once again, this form of variation can be exemplified in the occurrences of *see red*, but it is not common and it must be noted that it

---

15. Red was one of the four colours used by both American Football teams: red, metallic gold, blue and white (London Monarchs: http://en.wikipedia.org/wiki/London_Monarchs) and red, Kelly green, black and white (Raleigh-Durham Skyhawks: http://en.wikipedia.org/wiki/Raleigh-Durham_Skyhawks). Neither of these teams is still in existence.

may succeed here only because of the particular salience of colour words. That is to say, other lexical sets may not permit this kind of variation if their members have more than one salient meaning.

When *see red* undergoes substitution of *red* with another colour word, it sometimes also occurs in a "non-canonical context"; and yet it can still be recognised as meaning *see red*. How can this be possible? The reasons lie in the interaction between item and environment, phrase and cotext. Unlike longer idioms, *see red* contains only two words, and so there is not much scope for variation to take place within the phrase. When variation does occur, it has to be well motivated, otherwise the connection between the canonical form and the innovative variant may not be made and the idiomatic meaning simply not activated. The substituted colours are therefore absolutely unambiguous in the context of use and they also appear in close proximity to the necessary disambiguating lexis in the cotext, as can be appreciated from Examples (25)–(28).[16]

(25)   *Jeans* ads dropped after protesters see **blue**.

(26)   *UNITED* FANS SAW **BLUE**

(27)   Hunched figures slurp yaki soba, using the disposable wooden chopsticks that make *environmentalists* see **green**.

(28)   This was a derby: *niggling* meanness, haste and waste, *frustration* and *fury*. Fowler saw **red**, Cusani, Baldwin and Erskine **yellow**, the crowd **bleakness** on a **golden** afternoon.

Although *see red* is correctly an unambiguous idiom, in that it has no literal counterpart, it has already been seen that one of the effects of relexicalisation is to force a compositional reading. This happens when the idiom undergoes modification, but equally when it is unchanged, as in Examples (18)–(19), above. However, there is one important difference between changes to the cotext alone, and changes to both the cotext and the canonical form: the conventional meaning of *see red*, with all its pragmatic inferences in place, must be unambiguous when multiple variation is going on. If the idiomatic meaning is not relevant to the context there can be no justification for connecting e.g. *see green* (Example (27)) to "anger", so the meaningfulness would be lost. In the examples shown, we find that the idiomatic meaning is always strong, whether the colour word is substituted entirely (Examples (25)–(27)) or is a hybrid of substitution and addition (Example (28)).

Apparently *see red* can also be expressed as *see blue*, if the source of the anger is blue in colour (denim in Example (25), a football strip in Example (26)). So

---

16. Source: Bank of English 329m version. Ex. 25: Guardian; Ex. 26: British Books; Ex. 27: Economics; Ex. 28: Guardian.

just as red is traditionally said to be the source of a bull's anger, here we have *blue* responsible for the equivalent reaction in the context. Example (27) suggests that angry environmentalists might well *see green* rather than *see red* when they get irate, but it can only do so because the cotext forces that meaning out: the anger is caused by wasteful actions – the use of *disposable wooden chopsticks*, rather than eco-friendly re-usable ones – and while this would probably anger environmentalists, it might not worry other groups in the slightest. This meaning does admittedly seem somewhat forced: it is after all the only example of *green* being used to refer to anger in the English data, and although there are three other examples of *see green* with ecological overtones, these are literal and hence do not express "anger", "envy", "youth" or any other of *green*'s extended meanings.

The variants examined here underline that the context of situation plays an integral part in the creation and interpretation of intentional word play. The reader is invited to make the connection and therefore appreciate the writer's craftsmanship in creating the expression, as well as feeling self-satisfaction (Partington's "smugness effect", 1998: 140), for having done so successfully. But the role of context cannot be underestimated. Every single instance of word play found in the corpus was located in what the psycholinguists refer to as a "richly supportive" context – one which provides all the necessary information for deciphering the meanings(s) being conveyed. Word play can only exist where ambiguity is possible, but the ambiguity has to be constrained and resolvable. There are simply too many possible meanings, so an innovative variant has to be contextually grounded in order to be successful. And the writer is unlikely to draw a great deal of satisfaction from creating superficial, empty word play which is unlikely to be appreciated.

7.6   *The grass is always greener…*

The creative variants that we have studied in this chapter and in those preceding it have illustrated that variation is usually a surface phenomenon, because the underlying semantic and pragmatic meaning of the canonical form is preserved. The image schema may be reactivated and enhanced; component words may be relexicalised; expressive nuances may be added; connotative meanings may be evoked. But the pragmatic meaning expressed by the idiom – its semantic prosody and true raison d'être – is always present.

There are occasions when this general rule is violated. Sometimes creativity can be so extreme that the canonical form is no longer recognisable, and the communicative intent of the variant can only be inferred from contextual cues. Other times, the relationship to the canonical form is quite clear, but the innovation

serves as a mere embellishment which does not communicate meaning. Having lost the connection between form, content, and function, 'the string of words just "means" – it is not put to use in a viable communication' (Sinclair 1996b: 34). This case study will illustrate the progression from optimal innovation to meaninglessness using creative variants of the idiom *the grass is always greener*, found in the WordBanks Online collocations sampler and on the Internet using the search strings reproduced in the previous chapter (cf. Table 6.1, §6.2).

All the unique forms found in the corpus data are reproduced in Figure 7.1. At the top are the canonical forms (lines 1–2), followed by unmarked variants (lines 3–7), truncated forms (lines 8–9), paraphrases (lines 10–13), and apparently literal renderings (lines 14–15). Of the paraphrases, the extended context of lines 10–11 demonstrates clearly that the idiom is intended to be understood both compositionally and non-compositionally; lines 12–13 are also interpretable both ways, but the double meaning is not signalled explicitly.

While the low number of canonical forms may seem striking, it is simply a demonstration that the longer the phrase is, the more variation is it likely to undergo as it is slotted into the surrounding text. Headlines and titles (lines 8–9) appear to favour the truncated form, which we surmise constitutes the core, or key part, of the idiom – that element which allows the idiomatic meaning to be accessed. Of interest is the way in which the insertion of an adverbial into the two literal expressions (lines 14–15) differs from the other adverbials used in the idiom, contributing to an exclusively compositional reading. While adverbials are also inserted in lines 10–13, it can be seen that these are all intensifiers: these fit into the pattern set by *always* in the canonical form (lines 1–2) and in fact intensify the entire phrase. In contrast, the adverbs in lines 14–15 exert a more local action, modifying only *greener*, and it is this local action which strengthens the literal-only reading.

```
 1  mprove the consumer flow, the grass is always greener over there, and now that
 2  in the yearbook said [f] The Grass Is Always Greener. [f] Mack's first reacti
 3  a deeper question: why is the grass always greener on the other side [p] It ha
 4  re said: `Remember Pally, the grass is not always greener on the other side [p
 5  h life and a feeling that the grass may be greener somewhere over the hill. [p
 6  se-knit Washington world. The grass may not be greener on the other side of th
 7  n to fresh fields. But is the grass greener? [p] To find out more call 0839106
 8  all 0303 226602.  [p] [h] The grass is greener - It's Far And Away the fairest
 9  still survives.  [p] [h] The grass is greener; Beautiful places; Weekend Toda
10  perficial thrills enough. The grass does indeed sprout greener on t'other side
11  ry and the discovery that the grass really was greener outside the Wimbledon g
12  ] Dreadful things. Erm so the grass was never quite as green [ZF1] as [ZF0] as
13  ch only goes to show that the grass is always greener.) [p] But this year
14  ne elegance. South, where the grass is perceptibly greener, the River Derwent
15  he park of Purslem Manor, the grass already greener for yesterday's rain, the
```

Figure 7.1 *The grass is always greener*

The concordance lines retrieved, while limited in number, suffice to reveal the meaning of the idiom.

> If you say **the grass is greener** somewhere else, you mean that other people's situations always seem better or more attractive than your own, but may not really be so. (COBUILD)

In addition to the semantic meaning, the idiom will have a semantic prosody, because it has been seen up to this point that the semantic prosody of an idiom is the meaningful element which holds it together in spite of variation. There is not really enough information in the corpus examples in Figure 7.1, there being too few of them for the more abstract elements of the unit of meaning to be identified, so reference was made to canonical forms found on the Internet (in particular, to those found in running text rather than in isolated positions such as in titles and headlines). Dissatisfaction with one's lot and yearning for pastures new are already included within the semantic meaning of the idiom, but what is less evident is that the idiom is used as a criticism towards individuals who undervalue their (privileged) position and are resentful toward others (who may be no more privileged than they are themselves). It is this semantic prosody of perceived unfairness and perceived slight (which combine to form an unjustified sense of grievance) that must be borne in mind when discussing variant forms of the idiom. As we will see shortly, the point at which it disappears is the point at which innovation is no longer "optimal" and disintegrates into meaningless empty punning.

The non-compositional idiomatic meaning and the semantic prosody are completed by an image schema. As has already been mentioned, the conceptual schema of an idiom is likely to be latent in the canonical form, but it is reactivated when variation occurs. The image, like all prototypes, will vary from person to person on the basis of their personal experience. As far as *the grass is always greener* is concerned, the canonical idiom refers to the stereotype whereby English houses have gardens to the front (sometimes) and to the back (almost always). The only essential characteristic of this commonplace as far as the image schema is concerned is that there must be a stretch of grass in one's own garden and a neighbour living next door who has an equivalent stretch of grass. Grass that is well tended and watered grows a lush green; grass that is parched, beaten, worn down or burned will not be so green. If one's neighbour's grass looks in better condition than one's own, this will be a source of grievance, resentfulness and petty rivalry.

The variations found on the Internet for *the grass is always greener* display a range of variation types, both conforming to trends we have already examined in this and in the preceding two chapters, and others which still require

elucidation.[17] Variation of the colour word is by far the most common type: there are only four instances where *grass* is changed (Examples (29)–(32)), and it can be appreciated that the relation of these variants to the idiom hangs on a very slender thread.[18]

(29) **The arm is always golden-er** on the other kid.

(30) **The fence is always straddley-er** on the greener side.

(31) Sure it does, the **planet is always blue-er** on the other side! And its grass definitely looks greener! But it has people.

(32) **The hair is always blue-er** on the … [ellipsis in original]

As can be seen from Examples (29)–(32), the comparative seems to be a crucial element. Indeed, in all four examples, comparative forms have been coined on an ad-hoc basis. The same is true in Examples (33) and (34). However, if *green* remains, it can appear either in the base form (7 times) or the canonical comparative (18 times), or an either/or combination (Example (35)). In other words, if the canonical lexical item is present, it can withstand grammatical variation, but if it is changed, the substituted term must appear in the canonical comparative form.[19]

(33) Tom Green. **The grass is always Green er** on the internet side.

(34) **The Grass is Always Greeny-er** ~ Spotlight on David Greenwalt.

(35) Know that **the grass is not always green(er)** on the other side but we all have a roll to play in this world and god made us all different for a reason with many levels so no matter who u are or where u come from.

Unsurprisingly, *green* is most frequently substituted by other colour words. The other substitutions relexicalise *grass*, including one example where the *grass* referred to is *marijuana*, rather than the grass that covers lawns, but there are a handful of variations which make a clean break with the image schema. This has

---

17. In the examples to follow, variations occurring in body text appear in sentence context, those used as headings appear together with the relevant block of following text, and those which appear in isolation (as page headers, captions, etc.) are reproduced as stand-alone items. No editorial correction has been carried out.

18. Sources: Ex. 29: <bamareport.blogspot.com/2005_11_27_archive.html>; Ex. 30: <www.poetserv.org/SRR16/goodell.html>; Ex. 31: <fencestraddling.blogspot.com/2007/02/cupid-needs-glasses.html>; Ex. 32: <http://www.43things.com/entries/view/450352>.

19. Sources: Ex. 33: <it.video.yahoo.com/watch/439876/2469841>; Ex. 34: <http://www.cityofangel.com/council/crew.html>; Ex. 35: <http://profile.myspace.com/index.cfm?fuseaction=user.viewprofile&friendid=357503696>.

also been seen in Examples (29)–(32), but none of these mentioned with word *grass*, Examples (36)–(37) demonstrate this kind of variation in completion of the string *the grass is …*[20]

(36) **The grass is always more frequent** on the other side. How often are you waiting for an F train (at Carroll or Bergen) and 2 G's go by? You are like where the F is the F?

(37) Oscar hype: **The grass is always screener**. Studios and indies are putting a lot of time, energy and money into something that they're not even sure they believe in: screeners.

(38) **THE GRASS IS ALWAYS MEANER** The smell of summer is without a doubt that of freshly cut grass. God it's good. But did you know it's deadly? Lawn mowers cause 80,000 injuries a year in the US. That's a decent sized town full of lawn mower victims. Think about it.

Example (36), like Examples (29)–(32) before it, does not bear enough similarity to the canonical form to succeed effectively. If all the lexis is changed, but the structure is preserved, it is possible to activate a new schema which can be mapped onto the original one, and in this way it is separate from, but related to, the canonical form. In Example (36), only some of the lexis has been changed so the schema is spliced: it corresponds neither to the canonical form nor to the innovation. The relationship to the canonical form and the meaning conveyed by it is therefore rather weak. Examples (37) and (38) overcome the problem of the image schema by adopting a rhyming substitute. The phonetic resemblance between *greener* and *screenier* (Example (37)) and *meaner* (Example (38)) reinforce the connection to the canonical form and, while as odd-sounding as Example (36), they succeed communicatively in a way that it does not.

In most of the variants, however, the image schema is preserved. The features of *grass* are expressed in terms of our conceptual knowledge of grass – not only its colour, but the fact that it is a plant, and that it is a desirable plant to have in a garden. "Grassiness" is emphasised by substituting *green* with a normal collocate of *grass*, as in Examples (39) and (40).[21]

(39) **The grass is always** *sweeter* on the other side of the bridge.

20. Sources: Ex. 36: <http://bergencarroll.com/2007/06/25/the-grass-is-always-more-frequent-on-the-other-side/>; Ex. 37: <http://findarticles.com/p/articles/mi_hb1437/is_200101/ai_n5952893>; Ex. 38: <www.grist.org/news/daily/2006/08/16/4/index.html>.

21. Sources: Ex. 39: <www.lexington1.net/technology/instruct/ppts/LAppts/Folklore%20Jeopardy.ppt >; Ex. 40: <danismyfriend.blogspot.com/2008/06/grass-is-always-weedier.html>.

(40) **The grass is always** *weedier*… My neighbor must use fertilizer. My yard is just growing the bermuda grass. This guy's got all sorts of stuff taking root.

But by far and away the preferred type of variation is by substitution of the colour word, and there are two very strong tendencies present. The first is to reverse the directionality of the image schema, making the neighbour's grass the less attractive. This occurs six times, and the substituted word is always *browner* (never *yellower*, which would also be possible but is less emphatic). Two extended examples are provided below (Examples (41)–(42)), illustrating that the meaning has indeed been reversed as a result of the schema reversal.[22]

(41) **The Grass Is Always Browner**, Los Almos Labs Pending Lay off. Though I don't like writing about where I work on my blog, I will mention that I work at Intel in Rio Rancho. I am part of that group (my department specifically) that could be laid off. It's bad times right now and the worse is yet to come.

(42) Sounds like a case of "**the grass is always browner**." The person whose partner had a one-night stand might look over the fence at someone whose partner got emotionally entangled and think, "Lucky him/her. At least his/her partner didn't act on impulse. At least he/she can trust his/her partner when he/she says that he/she won't do it again. My partner just leapt on some stranger, so how can I ever trust him/her again when he/she goes off alone/unaccompanied to some other city/state?"

Of all the colours that are used in place of *green*, the most common one comes as a surprise. It does not fit into the image schema, as lawn grass is never naturally this colour, and, as a result, the idea that grass should be *bluer* is incongruous. This variant, *the grass is always bluer*, occurs ten times (out of a total of 49 variants).

Away from the influence of context, *blue* would probably be interpreted to mean 'depressed', which is its dominant connotative meaning. Yet this interpretation contrasts with the meaning of the canonical form, which expresses disgruntlement with one's own state in comparison with another's, not the reverse (as is expressed in *the grass is always browner*). In context, it is seen that this interpretation is not intended since every one of these variants occurs in connection with bluegrass music. But although the connection with grass and the unusual colour can be traced, the question of meaning is not resolved. *The grass is always greener* conveys a sense of disgruntlement, so why is it present – albeit in variant form – if disgruntlement is evidently not being expressed? Examples (43)–(45) seem to

22. Sources: Ex. 41: <http://www.greginthedesert.net/files/820a5cc563b3d78738f767aa45be2da9-1107.html>; Ex. 42: <http://chemistry.typepad.com/the_great_mate_debate/dan_savage/page/2/>.

2nd proofs

convey sentiments of quite a different kind, picking up on the comparison made in the canonical form (*the grass is always greener*) as a positive attribute, while disregarding its pragmatic implications.[23]

(43) Where **the grass is always bluer**. Three local **bluegrass** musicians give life to a dying art form at Zoey's Café.

(44) Lexington's **Bluegrass** Region – Come play where **the grass is always bluer**: Our Backyard!

(45) **The grass is always bluer**. **Bluegrass** music has always been present in Athens, but over the past several years it's seen a resurgence in popularity.

While *the grass is always bluer*, as a variant form of *the grass is always greener*, is contextually relevant on the lexical plane, the meaning that it is expressing is not relevant to the meaning expressed by the canonical form; the image schema is violated and the context of situation fractured. The result of this is to move into the territory of non-optimal innovation. The variant seems to have become institutionalised to a certain extent, but is clichéd – because while it is a clever play on words, it is not a *meaningful* play on words. The remaining instances are listed in Examples (46)–(50), with both the variant itself and the reference to bluegrass music highlighted. This reference can come at quite a distance from the pun (Example (48)), but there is a further indication that the form has become institutionalised within the bluegrass community. Example (46) in fact refers not to bluegrass music but to an evolution of it, *nu-grass*, and a further exploitation of the variant is also found in Example (51), where the pun is used for yet another musical genre.[24]

(46) **The Grass Is Always Bluer** – Various Artists (Artist) This fourteen-track collection features the brightest Americana and "**nu-grass**" artists of today.

(47) **The grass is always bluer** on the other side of rock n' roll.

(48) **The Grass is Always Bluer** on the Other Side – by Sara Gorelick – For busy, energetic professionals, retirement can be a double-edged sword, an enjoyable prospect accompanied by a dilemma – "Now what?" some ask. For Ron

---

23. Sources: Ex. 43: <http://www.vcreporter.com/cms/story/detail/?IssueNum=114&id=4406>; Ex. 44: <www.vacationfun.com/programs.asp?id=8>; Ex. 45: <www.redandblack.com/news/2001/08/10/Variety/Athens.Is.A.Melting.Pot.For.Music.Junkies-2587631.shtml>.

24. Sources: Ex. 46: <http://www.amazon.com/Grass-Always-Bluer-Various-Artists/dp/B000 22XE5O>; Ex. 47: <hardcore-troubadours.com/index.php?s=hillbillies>; Ex. 48: <http://www.makingmusicmag.com/features/07mar03.html>; Ex. 49: <www.exclaim.ca/articles/multiarticlesub.aspx?csid1=70&csid2=779&fid1=3901>; Ex. 50: <www.bluegrassusa.net/NewsRumors.htm>; Ex. 51: <www.yourdailymedia.com/post/1120469059>.

Whitlow, of Brentwood, Tennessee, that dilemma never arose. "There was no question of what to do with my time," says former UPS regional manager, Ron Whitlow. "I knew I wanted music to occupy a central part of my life." Specifically, that music had to be **bluegrass**.

(49) **The Grass Is Always Bluer** Canadian **Bluegrass** Pushes Traditional Boundaries By Jason Schneider Sunday afternoon outside the Cameron House.

(50) **The grass is always blue** on his side of the fence. By JEANNIE WILEY WOLF – Staff Writer. It's fair to say that Joe Steiner has **bluegrass** in his blood.

(51) **The grass is always emo-er** on the other side!

This time, however, there is the additional connotative meaning of *blue* (depressed) which connects with the emo phenomenon in youth culture.[25] *The grass is always emo-er* cannot be legitimately defined as an exploitation of *the grass is always greener*, as it does not fit with any other the variation tendencies that have been observed, except with *the grass is always bluer* which, as we have seen, is a full step away from the others. It is a superficial, empty pun, devoid of pragmatic meaning. *The grass is always emo-er* can therefore be defined as a *secondary exploitation*, i.e. a variation of a variation, and its pleasure-rating is dependent on the connection being made from bluegrass music to emo music. This apparently involves a fairly straightforward variation by semantic class, but it is not a kind of variation that can be derived directly from the canonical form. It may also make use of an unstated intermediary term, *blues*, which it is connected to bluegrass by lexical analogy, has a referential relationship to emo music (depression), and is connected to both by semantic affinity (musical genre). In this secondary exploitation, the pragmatic meaning of the canonical form is absent, because it is absent in the variant form from which it is itself derived.

7.7   The effects of variation

The detailed analysis of variants of *the grass is always greener* offers an insight into how a single canonical idiom can be exploited in many different ways, pulling together into a single extended example all the types of variation that have been discussed in Chapters 4–7. It also offers a convincing argument in favour of judging highly marked variants against variation tendencies, because only in this way is it possible to measure and evaluate each step taken away from the canonical form.

---

25. Emo music is the preferred choice of the "emo[tional]", sensitive teen. It is drearily focused on suicidal thoughts and world-weariness in general.

That some highly marked forms are in fact exploitations of variations, rather than exploitations of canonical forms, is one of the most important findings to come out of such an analysis, because they are identified through their divergence from norms and, crucially, from "norms of variation".

Normal variation preserves the idiom's image schema and is faithful to the semantic prosody, and it enhances the meaning conventionally expressed by the canonical form. Marked variation does so to a lesser extent, but if it is to fall within the area of "optimal innovation", the sense of the original must remain and combine with the new sense(s) brought by the variation. This provides enjoyment and satisfaction to the reader who appreciates that multiple meanings are present and can try to identify the ways in which they are connected (by identifying the relevant shared attributes, in the way that happens when metaphors are decoded). What I have spoken of here as superficial or empty punning goes beyond the boundaries of optimal innovation, because the fundamental rule of optimal innovation is that two (or more) meanings are simultaneously activated. If the connection to the canonical form – its wording, context of situation and image schema – is lost, and the variant form has an incompatible or indiscernible pragmatic function, it no longer expresses two meanings but simply makes a surface connection between lexical items in the novel form and those in the context. 'Pure innovations do not […] involve the recovery of any salient response on top of the novel one' (Giora & Shuval 2005: 244). Such a variant does not really "mean" anything; it is a mere figure of speech; an embellishment or ornament with minimal communicative value.

Words and meanings

8.1   A few words about meaning

The aim of this book has been to investigate the relationship that words have with meanings, and how that relationship shifts when language is used in unconventional ways. It is hardly the first study to have attempted to do this, but what sets this research apart from existing work into linguistic creativity is that it has drawn together findings from a range of fields within the discipline of linguistics – corpus linguistics, psycholinguistics, cognitive linguistics, phraseology, lexicography, and semantics – in order to account for meaning from several different but related standpoints. It has deliberately sought not to dismiss any of the viewpoints expressed within these fields, because although it is evident that in conventional language, meanings have words, once creativity kicks in, words equally evidently have meanings. It is pointless to try to prove that only one of these views of meaning is the "right" one: they are both right, but they apply to different kinds of language, in different contexts. "Words have meanings" stands at one end of the linguistic continuum, while "meanings have words" stands at the other, and between those extremes words and meanings interact in sophisticated and occasionally stunning ways.

The underlying principles that have guided this study ultimately derive from the work of J. R. Firth, some in its "pure" form, some filtered through neo-Firthian (Sinclairian) corpus linguistics. Meaning by collocation (§3.3) comes directly from Firth's work, as does the 'context of situation for linguistic enquiry' (Firth 1957: 182; §3.7.1). The extended unit of meaning (Sinclair 1996b), while not explicitly deriving from Firth's work, can be seen as a schematisation of how the context of situation is realised linguistically. Collocates indicate who the relevant participants and objects are; colligation sets them in relation to one another; semantic preference provides insights into the broader identity of the participants and objects; and the semantic prosody is the expression of the effect of the verbal action. What is missing here, however, is the associative aspect of language.

Semantic associations – evaluation, connotation, and metaphorical entailments – are also important parts of meaning, but are often hidden beneath the surface of a text. In corpus linguistics, which is by definition an empirical,

performance-based branch of the discipline, meaning in the mind is an unwelcome topic. This book has "outsourced" it, relying on the wealth of knowledge that has been accrued in the psycholinguistics literature, but this cannot be seen as a permanent solution, and it will have to be addressed sooner or later. The origins of connotative and associative meanings of a textual nature lie at the basis of literary studies into intertextuality; within corpus linguistics, the closest equivalent is likely to be the emerging area of collocational resonance, applied to the language as a whole (Williams 2008), to domains of discourse (Williams & Millon 2009) and other forms of restricted language (Philip 2010a). But this awaits us in the future.

8.2   Meaning what you say vs. saying what you mean

In the course of this book, we have looked at the effects of variation in fixed expressions. This has been done comprehensively, and with a purpose. It is not enough merely to bring forward interesting examples of word play and let them speak for themselves; it is equally inadequate to list types of variation without addressing the effects that these have on the transmission of meaning in a text. This book sets out the complete range of variation observed in the colour word expressions chosen for study, and applies a double analysis. On the one hand there is the textual study which takes the unit of meaning (Sinclair 1996b) as the frame of reference. The other facet of the analysis is to identify types of connotation, which followed Kerbrat-Orecchioni's (1977) classification of *connotations associatives* combined with an unravelling of the metaphorical and metonymical processes at work.

Connotation is all too often treated as if it were context-free. The data examined in the preceding chapters has hopefully demonstrated that this is not true, and that if the relationship between word meaning (literal or metaphorical) and connotative or associative meanings is to be properly understood, it is necessary to look at how those meanings are created in texts. In purely quantitative terms, some types of text-based connotation are more common than others. The most common ones connect effectively to the parameters set out in the unit of meaning and in the corresponding image schema, and involve the semantic affinity which holds between lexical relations, including synonymy, (co-)hyponymy and hyperonymy, as well as similar word forms with similar meanings. It should be mentioned that many generic descriptions of connotation would not include these, but the very notion of semantic relatedness is viewed in a different light when dealing with language in use. Traditional kinds of lexical and semantic relations are supplanted by metaphorically-derived shared attributes, often of an ad-hoc and temporary nature. They certainly operate at a secondary level of meaning,

connecting with image schemata, but whether one chooses to call them metaphor or connotation most likely depends on ones' theoretical leanings rather than any other single factor.

The less common kinds of connotation draw in information which is external to the normal context of situation, defined either linguistically or conceptually. These include connotations which have a referential-symbolic basis, or which are derived from cultural paronyms, or which make intertextual reference to citations and catchphrases. These kinds of connotation are "connotations proper", being the kinds of referential meaning that are most often defined as connotation. They are also the most marked and the rarest. In fact, the data indicates that it is very unusual for these connotations to be evoked in text, and when they are, there must be clear cues present in the context to ensure that the secondary meaning is activated. Precisely which meanings are activated by any given individual, however, can never be ascertained from corpus data alone: corpora can provide the norms and the external evidence of the meaning-making process, but it cannot account for the connections that individuals make in their heads on the basis of what is salient for them culturally, socially, or personally; or on connections that arise because of what that individual is thinking about as s/he encounters text. These personal aspects of meaning are certainly important, but they are private and ephemeral, part of *parole*, while a corpus provides evidence of *langue* (Tognini Bonelli 2002: 98–99).

8.3   Meanings in words

At the beginning of this book, in §1.2, six real-world examples of word play were put forward as being the type of language which this study sought to explain. Now is the time to return to them, applying the knowledge of idiomaticity, phraseology, salience, metaphor, connotation, and pragmatic meanings accumulated through the past chapters. Rather than simply indicate which idiom(s) they are derived from, it is possible to reconstruct the complete unit of meaning, which in turn leads to the setting out of its context of situation. It is also possible to investigate the metaphorical and connotative activity which is likely to be activated given the topic and source of the text, and even decide if it is optimally innovative or not.

The first example, ***Light relief from the winter blues***, is a headline to an article about the use of light boxes in the treatment of Seasonal Affective Disorder (SAD). The focal point for investigation is the collocation *light relief*, whose unit of meaning is summarised in Table 8.1a and mapped onto the relevant context

Table 8.1a Unit of meaning: *light relief*

Collocates	*PROVIDE* (7), *a little* (11), *a bit of* (8), *some* (5)
Colligates	Quantifiers: *a little* (11), *a bit of* (8), *some* (5)
Semantic preference	(i) provision: *PROVIDE* (7), *afford*, *offer*, *bring*, *give* (2), etc.
	(ii) tedious/difficult situations: *the sombre business of teaching science*; *the boredom or demands of the shift*; *after a winter of grim industrial strife*, etc.
Semantic associations	Light-heartedness here positively evaluated for its alleviating function (though might otherwise be negative, i.e. "frivolous")
Semantic prosody	Expresses appreciation of respite from a disagreeable situation

Table 8.1b Context of situation: *light relief*

Persons/personalities	People involved in difficult and/or boring situations, seeking some temporary reprieve
	Providers of *light relief* (e.g. friends, actors, comedians)
Verbal action	(Implicit) providers of light relief speak in a way that provokes laughter/enjoyment
Non-verbal action	(Implicit) providers of light relief behave in a way that provokes laughter/enjoyment
Relevant objects	Jobs and work
Effect of verbal action	Indicates appreciation of respite from a disagreeable state

of situation in Table 8.1b.[1] *Light relief* usually is something that is provided, and although it occurs in small quantities, these are deemed sufficient; it is offered by way of reprieve from tedious and unenjoyable activities and situations, and most often provided through humour. It is felt to be both desirable and necessary, and its temporary nature is acknowledged and accepted. The headline essentially fits with these schemata, suggesting that the text will discuss something trivial which might help beat depression. Only in the subsequent text does it transpire that it is an inanimate object (the light box) and not a person who provides *light relief*, and that no humour is involved. The double meaning can only be inferred during reading of the article. *Light* turns out to be polysemous in this context, exploiting the structure to allow the normal adjective+noun collocate to be read as a compound noun (*relief* involving *light*). Triviality is the salient meaning in the collocation, so even in a text whose topic is "brightness" this meaning is also activated. The language play is therefore caused by lexical analogy (polysemy) of the type discussed in §7.4.

---

1.   Source: British National Corpus. 55 occurrences of *light relief*.

Table 8.2a Unit of meaning: *grasp the nettle*

Collocates	*GRASP* (17)
Colligates	Partitive: *the nettle of NN* (6)
	Clause-final position
	Deontic modals: *(have) got to*, *need to*, *fail to*
Semantic preference	Politicians and government departments
Semantic associations	Delay and prevarication, unpleasantness
Semantic prosody	Expresses reluctance to do what needs to be done

Table 8.2b Context of situation: *grasp the nettle*

Persons/personalities	Politicians, government ministers, civil servants
Verbal action	Recommendations to take action, criticism that action has not yet been taken
Non-verbal action	To accept and implement change
Relevant objects	Policies and procedures
Effect of verbal action	Indicates reluctance to do what needs to be done

The second example operates along similar lines, although the pun is evident from the outset. *Greens need to grasp the nettle* is the first half of a headline which is completed with *aren't there just too many people?* The article discusses the controversial topic of population management, arguing that environmental problems such as pollution, shortages of food and resources, and deforestation, are ultimately caused by there being too many humans on the planet. Population management, therefore, is essential in forestalling environmental disaster. The pun centres on the juxtaposition of *greens* and *nettle*. In §7.4.1 the chained metonymy linking *green* to *person* was uncoiled, so the connection between *green* ("plant life") and *nettle* should be transparent.

To *grasp the nettle* is to avoid delay in dealing with a difficult situation or issue which is used aptly in this headline, as can be appreciated from the summary in Tables 8.2a and 8.2b.[2] Less apt, or less normal, as can be seen from the same Tables, is the presence of *Greens*, because it seems that the text refers to *Greens* as "environmentally-oriented individuals" rather than the political party, although this is not made clear. Overall, this example exploits the metonymically-motivated polysemy of *green* to enhance the meaning of the idiom and refocus its scope onto the environmental topic of the article it introduces.

The third headline to be discussed is more complex and more interesting than the previous two. *Old Lady takes a hammer to her furniture* contains multiple ambiguities and an example of a "fracture" in the context of situation (Louw 2007).

---

2.   Source: British National Corpus. 18 occurrences of *grasp/grab the nettle*.

Table 8.3 Conceptual schema: *hammer*

Persons/personalities	*(Strong) men* use hammers
Action	(i) Hammers *tap/drive/hit/whack* nails into wood/walls
	(ii) People *brandish*, *bring down on sth*, *grip*, *hold*, *swing*, *wave*, *wield, use* a hammer
	(iii) People *attack*, *batter to death*, *crush*, *hit*, *kill*, *smash*, *split sb's head open*, *strike*, *threaten* other people with a hammer
Relevant objects	Hammers are: *large*, *small*, *heavy*
	Man-made objects; these are undamaged prior to hammer being used;
	Nails, chisels, screwdrivers
Effect of action	(i) construction/ modification of man-made object (tool)
	(ii) damage or destruction (weapon)

Before explaining the context of use and giving everything away, the meaning of *take a hammer to sth* has to be explained. This expression is one possible realisation of the lexicogrammatical frame *take a [tool or instrument] to [an object]*. The meaning is idiomatic: "to destroy something with said tool or instrument", but the expression is too infrequent to schematise as a unit of meaning.[3] Instead, what is essential to the humour in the headline is found in the conceptual schema surrounding *hammer* when used as a tool or instrument, as set out in Table 8.3.[4]

The incongruity of the headline comes from the *Old Lady* being the wielder of the hammer. *Old ladies* do not share the relevant attributes of strong men; they are typically portrayed as being forgetful, small in stature, vulnerable. Of course, there is the occasional "Supergran" who stands out as an exception, but she is atypical. The end result is humorous because the context of situation associated with hammer use is "fractured" (Louw 2007) by the presence of *old lady* instead of (strong) *man*.

This is not the only aspect, however. In fact, the word play works on several levels, making this a particularly optimal innovation. Before reading the article, it is the above interpretation that is the most likely one, but what it refers to, and why it should be reported in the Financial Times rather than a tabloid, only emerges in the text of the article itself.

> Fans of *second-hand office furniture* will have the chance to *buy* a little piece of British economic history this month with the *auction* of, among other things, the hat stands on which generations of *Bank of England* workers hung

---

3. Only one occurrence of idiomatic *take a hammer to sth* in the British National Corpus.

4. Source: British National Corpus: 322 occurrences of *hammer* (noun; literal "tool" sense only).

their bowlers. The *Old Lady of Threadneedle Street* is offloading old-fashioned chairs, desks, cupboards and – bizarrely – a backgammon table which are surplus to requirements now the Bank has adopted a modern, open-plan look.[5]

The Bank of England has been known as *The Old Lady of Threadneedle Street* since the appearance of a satirical cartoon in 1797; this moniker can be expected to be familiar to readers of the *Financial Times*, even if it is not the most salient meaning of *old lady*. After reading the first paragraph of the article, the meaning of *hammer* in *take a hammer to sth* is reprocessed: the phrasing is consistent with the tool meaning, but the contextually relevant meaning of *hammer* is not the workman's tool but the auctioneer's hammer. Sale at auction is more conventionally referred to as *putting sth under the hammer*. *Hammer* is rightly interpreted as being a key word in the headline, but the phraseology used deliberately misleads the reader, and four types of associative connotation are in operation at once: *Old Lady* vs. *old lady* exploits polysemy, but *Old Lady* also has a symbolic connotation (the historical and cultural role of the Bank of England, and possibly the cartoon itself, which still has currency and can be found on the Internet). *Take a hammer to sth* vs. *put sth under the hammer* exploits combinatorial affinity by deliberately using the wrong phrasal configuration for the intended meaning; and the whole image schema is interpreted in the light of previous use of *hammer* and *old lady*, both linguistically and in terms of extra-linguistic experience and knowledge.

The remaining three examples come from fiction rather than journalism. From the fifth book in the *Harry Potter* series comes "*... well, it's no good **crying over spilt potion**, I suppose ... but **the cat's among the pixies now**.*" In fiction, word play can be expected to be more firmly embedded in the preceding text, and this is a fine example of that expectation. There are two idiom variants present: *cry over spilt milk* and *cat among the pigeons*. Both introduce features of the magical world – *potion* and *pixies* respectively – to root the word play within that world.

The first idiom involves a simple modification at the level of semantic preference (cf. §5.5); *milk* is one of the key words in the idiom, so the change to *potion* is highly marked. *Potion* fits into the semantic class of "drinkable liquids" to which *milk* also belongs, although it can be appreciated that this membership is ad hoc because only particular relevant attributes are involved: even in the magical world, potions are not all-purpose, everyday beverages, but are only drunk for magical purposes. The idiom variant is used in accordance with its conventional

---

5. *Financial Times* Monday July 12, 2004. Text originally retrieved from <http://search.ft.com/search/article.html?id=040712000930&query=old+lady&vsc_appId=totalSearch&state=Form> (accessed 12/7/2004). The article was initially headlined *Old Lady prepares to sell off her antiques*, which is also nicely ambiguous but nowhere near so incongruous as the form examined here, which appeared on the same website later that day.

Table 8.4a Unit of meaning: *crying over spilt milk*

Collocates	*no* (7); *no use* (4), *no good* (2)
Colligates	Negative polarity
	Direct speech and citation forms
	Clause/sentence final position
Semantic preference	Pointlessness: *no use* (4), *no good* (2), *no point*, *a waste*
Semantic associations	Impatience and criticism
Semantic prosody	? Expresses impatience, possibly veiled criticism; nothing can be changed, so dwelling on the topic is futile

Table 8.4b Context of situation: *crying over spilt milk*

Persons/personalities	Older/wiser/more experienced person (speaker)
	Younger/less experienced person (listener)
Verbal action	Older/wiser/more experienced person concludes a topic of conversation with advice in form of proverb (cliché)
Non-verbal action	Younger/less experienced person seeks consolation or advice regarding a regrettable situation or event
Relevant objects	Regrettable situation or event (in recent past)
Effect of verbal action	To end discussion of topic instead of offering sympathy

phraseology (see Table 8.4a) and context of situation (Table 8.4b).[6] Harry, the younger participant, has just been party to a frightening and dangerous attack, but in defending himself has violated the "Decree for the Reasonable Restriction of Underage Sorcery". Mrs Figg, the older participant, is bringing discussion of the topic to a close. The context and wording are thus appropriate, but *potion* instead of *milk* serves no purpose other than to create a superficial link to the magical world.

The second idiom fares better, because even if the same strategy has been adopted (i.e. substitution of an idiomatic key word), the substituted term, *pixies*, has already built up a rich associative history for *Harry Potter* readers, having featured in an episode in *Harry Potter and the Chamber of Secrets* (Rowling 1998), which serves as back-story for the interpretation of this idiom variant (Figure 8.1). In this context, which differs from many accounts in traditional folklore, *pixies* can be seen to bear a series of resemblances to *pigeons*. They are not domesticated, so must be caught and caged in order to be controlled; they are of a similar size; both can fly. *Pixies*, being magical creatures, also wilfully wreak havoc, and this is the source of the semantic associations that the word substitution contributes

---

6. Source: British National Corpus. 11 occurrences of *spilt/spilled milk*, which is really too few to reliably posit a semantic prosody: a provisional one is suggested in Table 8.4a.

```
 dramatically. "Freshly caught Cornish pixies. " Seamus Finnigan couldn´t control
cky little blighters they can be!" The pixies were electric blue and about eight
ened the cage. It was pandemonium. The pixies shot in every direction like rocket
d them up, round them up, they´re only pixies ..." Lockhart shouted. He rolled up
t had absolutely no effect; one of the pixies seized his wand and threw it out of
m?" roared Ron as one of the remaining pixies bit him painfully on the ear. "He j
ence," said Hermione, immobilizing two pixies at once with a clever Freezing Char
 "said Harry, who was trying to grab a pixie dancing out of reach with its tongue
p] Since the disastrous episode of the pixies, Professor Lockhart had not brought
 on´t fit. And Millicent Bulstrode´s no pixie." "Good thinking," said Ron, unlocki
ed anything from him except not to set pixies loose." Neville Longbottom had been
```

Figure 8.1 *Pixies* in *Chamber of Secrets*

Table 8.5a Unit of meaning: *cat among the pigeons*

Collocates	*PUT* (16) *the* (cat) (13)
Colligates	Past tense (14);
	Topic is introduced cataphorically (and at some distance)
Semantic preference	Controversial actions/statements
Semantic associations	Moral and social acceptability judgements
Semantic prosody	Expresses acknowledgement that some individuals will react unfavourably towards the described action/situation on the basis that it is morally/socially unacceptable, but *does not share that judgement*

Table 8.5b Context of situation: *cat among the pigeons*

Persons/personalities	Actor/speaker + more conservative-minded listener/audience
Verbal action	Controversial or socially unacceptable comments and/or reaction to the same
Non-verbal action	Controversial or socially unacceptable behaviour and/or reaction to the same
Relevant objects	[None identified]
Effect of verbal action	Functions as cataphoric reference reporting a reaction of disapproval by others to the topic discussed

to the existing phrase. This richness of associations combines with the meaning expressed by the canonical form of the idiom, which is outlined in Tables 8.5a and 8.5b.[7]

Of the idioms studied in this book, several have been seen to be clause-final, indicating that they have a summative function in text. *Cat among the pigeons* does the opposite: it sets up an expectation for information, as it describes a judgement of a reaction before describing the reaction itself, and often before even describing its cause. *The cat's among the pixies* conforms to this patterning, setting in

---

7. Source: British National Corpus. 19 occurrences of *cat among the pigeons*.

2nd proofs

motion not only the most immediate reaction (one page later) but also the events which unfold over the course of the rest of the book. Harry's actions are seen to anger his adoptive parents and are cause for potential expulsion from school; but the cause of Harry's actions, the release of Dementors – dark magical creatures which he was defending himself against – is the true cause of the pandemonium which ensues and which is not properly resolved until the end of the *seventh* book (around five hundred pages later). *The cat's among the pixies* therefore combines cultural paronyms (*pixies* in this fictional context), lexical analogy (homophony in first syllable), a signifier-signified relation (similar word – similar referent), in-tertextuality (or rather *intra*textuality, if the series is viewed as an episodic whole, cf. Philip 2010a) in the associations attached to *pixie*, plus the cataphoric refer-ence to the chaos and disorder yet to come.

Moving away from fictional worlds, and back to fictional representations of the real world, the fifth example is an ironically-worded, hypothetical Lonely Hearts insert worded by Angelique, a young female police officer whose mother is hoping that she will soon find herself a boyfriend (Brookmyre 2003: 44), *Embit-tered, disillusioned, increasingly shrewish and expertly violent female workaholic seeks **horizon-dwelling male** to make nothing else matter*. The irony in the citation above stems from the relevant personality in the context of situation. In a typical Lonely Hearts insert, the writer advertises his or her most attractive features, in-cluding physical attractiveness, wealth, trustworthiness, pleasantness, and so on. The brevity of the text ensures that no space can be wasted on negative charac-teristics that might detract a potential partner. Thus *embittered*, *disillusioned* and *shrewish* are immediately perceived as being out of place – they have "fractured" the context of situation (Louw 2007) – and this prepares the ground for an ironic interpretation of the rest of the insert.

However, being found in a work of fiction, the ground is prepared earlier, and the remaining phrases, *on the horizon* and *to make nothing else matter*, are delib-erately used to echo Angelique's mother's words which are reported on the page preceding the citation:

> The job would never love her, she could agree with her mum about that now. Where they continued to differ was over the theory that *nothing else would matter* if she could find someone who did. (ibid.: 43)

> "is there still no-one *on the horizon*?" was Mum's preferred oh-so-reluctant-but-I-have-to-ask-because-I-care phrase for dredging up the subject… In fact, An-gelique considered her mother's choice of words unwittingly apposite. If there was a right guy for her in this world, then the *horizon* was exactly where she'd expect to see him, it being a place you can never actually reach. (ibid.: 43–44)

Table 8.6a Unit of meaning: *on the horizon*

Collocates	(i) *bright* (5), *hope* (5), *light* (4)
	(ii) *cloud* (9), *looming* (6) *dark* (3)
Colligates	Preposition (*of, for*) to specify horizon
	Clause-final position (119) or cleft sentence (7)
Semantic preference	(i) hope and change for the better
	(ii) ominous events, bad weather, lack of (job) opportunities
Semantic associations	Positive or negative, depending on semantic preference
Semantic prosody	(i) Relief that improvement to present situation is possible [with positive collocates]
	(ii) Fear of future as an unknown quantity [with negative collocates]

Table 8.6b Context of situation: [significant other] *on the horizon*

Persons/personalities	Young single female
	Older female (usually mother) who broaches the subject
Verbal action	To inquire about daughter's love life
Non-verbal action	–
Relevant objects	The sought-for male partner (absent)
Effect of verbal action	Tension arises within the mother-daughter relationship because the comment is perceived as intrusive

On the horizon, meaning "foreseeable", can refer to positive and negative events (see Table 8.6a),[8] but it can be appreciated that the meaning has been restricted in the preceding text to refer significant others. Because of this preparation, *horizon-dwelling male* is understood both inter- and intratextually as being a paraphrase of the cliché, and the telegraphic form mimics the style of genuine Lonely Hearts inserts (exploiting intertextuality of a different order). As is true of many idiom variants, *horizon-dwelling male* relexicalises the formulaic string *no-one on the horizon* and refocuses the meaning by specifying *dwelling* (his home is on the ho-rizon) rather than *living* (he lives his life there) or *being* (he is there). If his home is located in a place which we have learned "you can never actually reach", any hope of meeting him are slim indeed, which is consistent with the context of situation (Table 8.6b) which shows that the sought-for male is an absent "relevant object".

The final example is a pun, occurring in direct speech, at the culmination of a context-specific joke. The novel in question (O'Brien 1979: 48–49) is set during the Napoleonic wars, and revolves around the interlocking stories of the ship's captain Jack Aubrey and his surgeon (and spy) Stephen Maturin. Sitting at mess, Maturin has observed two weevils emerging out of some ship's biscuit. He sets

---

8. Source: British National Corpus. 130 occurrences of *on the horizon*.

Table 8.7a Unit of meaning: *lesser of two evils*

Collocates	*CHOOSE* (2), *SEEM* (2), *was* (4)
Colligates	Downtoners: *perhaps* (2), *at least*, *somehow*
Semantic preference	Options in (i) military strategy and (ii) medical treatment; Choices: *choose, consider, decide, prefer*
Semantic associations	Difficult situations
Semantic prosody	Reluctance to make choice; avoidance of blame for potentially making the wrong choice

Table 8.7b Context of situation: *lesser of two evils*

Persons/personalities	Person(s) faced with making the difficult decision
Verbal action	Decision taken
Non-verbal action	Reluctance to take the decision
Relevant objects	Potential consequences of the course of action chosen
Effect of verbal action	Avoidance of blame for potentially making the wrong choice

them in the same place and asks his mess-mates to decide which one will reach the bread the fastest. The vote is clearly in favour of the larger, fatter, stronger of the two. Yet Maturin disagrees, stating *Don't you know that in the Service one must always choose* **the lesser of two weevils?**

The pun is based on *the lesser of two evils* ('the less terrible of two bad options'), and exploits the intrusive /w/ which occurs between *two* and *evils* to permit vowel elision: *two evils* and *two weevils* are therefore true homophones. The unit of meaning for *lesser of two evils* is schematised in Table 8.7a, its context of situation in Table 8.7b.[9] The link between the context of use in the novel and the semantic preference for decisions relating to military strategy is important, as it ensures that the pun is relevant in general terms. However, the pragmatic meaning is reversed because there is no blame to be apportioned in choosing one or other *weevil*, unlike choosing between *evils*. Furthermore, the relevant objects in the context of situation have changed dramatically: now there are concrete objects, *weevils*, to be chosen for amusement, while in the canonical form, the objects are abstract in nature and are likely to have important consequences. Once again the context of situation is fractured so a special effect is created in text. The pun is facetious, deliberately mocking the maxim observed in the Navy, and since it does not respect the expected semantic prosody, it is perceived as functionally meaningless (cf. §7.6) and if only for this reason, non-optimal.

---

9. Source: British National Corpus. 21 occurrences of *of two evils*.

8.4   Famous last words

Idiomatic language is itself complex, but creative modifications of it can be seen to increase that complexity many times over. This book has sought to span the range of that complexity in an attempt to explain how words and meanings can relate to one another. What is has not done – what it cannot do – is resolve the "words have meanings" and "meanings have words" dispute. By now it should be apparent that both views are right, but are dependent on the co(n)text of use.

Idioms in their canonical forms are not so dissimilar to single words in that their usage patterns are regular and predictable. Taking the core of an idiom, which in this book has nearly always been a two-word collocation, it is possible to draw up a detailed summary of its unit of meaning. Doing so demonstrates that even if they can be used as stand-alone items, idioms attract collocates and colligates, usually have a preferred syntactic position, are used within well-delimited semantic fields, and have a distinct pragmatic function in text. This is the "real" or "actual" meaning of an idiom which belies its surface realisation, and this meaning must be relevant in idiom variants if they are to be communicatively successful.

Semantic prosody – the pragmatic force underlying phraseological expressions – seems to be the "glue" that holds phraseological meanings together. As long as it is present, conventional phrases can be altered and yet still be traced back to the canonical form and its full meaning potential. It works in tandem with the conceptual representation of the phrase, but while that representation may be cast in various lights, activating different metaphoremes (Cameron & Deignan 2006), the prosody stands firm. Were it not to do so, idioms – whether in their canonical form or a variant, located in their canonical context, in an unusual context, or standing alone – would no longer be able to mean.

One of the major contributions this book has made to corpus linguistics is in its treatment of associative meanings. Connotations are all too often taken as given, and connotation itself is rarely defined except in the most simple of terms. Distinguishing between the different degrees of connotation, many of which operate simultaneously, means that different levels of secondary meaning activity can be identified and viewed in relation to canonical forms and contexts. This is simply impossible if connotation is taken to mean 'any and all secondary meanings', especially because connotation is distinct from another facet of secondary meaning, metaphor. Metaphor and metonymy contribute to the conceptual schema that accompanies idiomatic language, enhancing and refocusing meaning without altering its fundamental message. In doing so, they fill out the details of the context of situation – identified by means of the lexical unit of meaning – by

enriching and refining information regarding the relevant features of the persons, personalities and objects.

The context of situation, reconstructed lexically in the unit of meaning and conceptually in the metaphoreme, tells us about normal patterns of use. Less normal uses stretch the context of situation, but can be admitted. Highly marked uses can be seen to arise from a fracture in the context of situation: the participants or objects deviate too far from the norm, or their relevant attributes do, or their context of utterance does. Successful marked uses exploit this to their advantage, creating many-layered interpretations rich in associative potential. But sometimes the communicative intents of word play and canonical form part company. Puns are primarily aesthetic effects with no pragmatic function other than to entertain. Empty puns are empty not because they lack semantic content, but because they do not transmit any purposeful message. Although they can be traced back to their canonical form, their behaviour can be compared to that of ungrateful offspring disowning their parents. On the surface, the resemblance between the two forms is apparent, but the underlying intents and purposes are quite at odds with one another. To paraphrase Carroll (1865: 69), canonical forms say what they mean; empty puns mean (only) what they say.

References

Ayto, J. & Crofton, I. 2006. *Brewer's dictionary of modern phrase and fable.* Weidenfeld and Nicholson: London.

Baranov, A. N. & Dobrovol'skij, D. O. 1996. Cognitive modelling of actual meaning in the field of phraseology. *Journal of Pragmatics* 25: 409–429.

Barnbrook, G. 1996. *Language and Computers: A Practical Introduction to the Computer Analysis of Language.* Edinburgh: Edinburgh University Press.

Bartsch, S. 2004. *Structural and Functional Properties of Collocations in English.* Tübingen: Gunter Narr.

Bednarek, M. 2008. Semantic preference and semantic prosody re-examined. *Corpus Linguistics and Linguistic Theory* 4(2): 119–139.

Benson, M., Benson, E. & Ilson, R. 1986a. *The BBI Combinatory Dictionary of English.* Amsterdam: John Benjamins.

Benson, M., Benson, E. & Ilson, R. 1986b. *Lexicographic Description of English.* Amsterdam: John Benjamins.

Berlin, B. & Kay, P. 1969. *Basic Color Terms: Their Universality and Evolution.* Berkeley CA: University of California Press.

Black, M. 1962. Metaphor. In *Models and Metaphors*, 25–47. Ithaca NY: Cornell University Press.

Black, M. 1993. More about Metaphor. In *Metaphor and Thought,* 2nd edn, A. Ortony (ed.), 19–41. Cambridge: CUP.

Blasko, G. D. & Connine, C. 1993. Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 657–668.

Boers, F. & Demecheleer, M. 2001. Measuring the impact of cross-cultural differences in learners' comprehension of imageable idioms. *ELT Journal* 55(3): 255–262.

Brookmyre, C. 2003. *The Sacred Art of Stealing.* London: Abacus.

Budanitsky, A. & Hirst, G. 2004. Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics* 32(1): 13–47.

Bussmann, H. 1996. Connotation. In *Routledge Dictionary of Language and Linguistics*, 96. London: Routledge.

Cacciari, C. & Tabossi, P. 1988. The comprehension of idioms. *Journal of Memory and Language* 27: 668–683.

Cacciari, C., Massironi, M. & Corradini, P. 2004. When color names are used metaphorically: the role of linguistic and chromatic information. *Metaphor and Symbol* 19(3): 169–190.

Cameron, L. 2003. *Metaphor in Educational Discourse.* London: Continuum.

Cameron, L. & Deignan, A. 2006. The emergence of metaphor in discourse. *Applied Linguistics* 27: 671–690.

Carroll, L. 1872. *Jabberwocky*. In *Through the Looking Glass*. London: Macmillan.

Carroll, L. 1865. *Alice's Adventures in Wonderland*. London: Macmillan.

Čermák, F. 2001. Substance of idioms: Perennial problems, lack of data or theory? *International Journal of Lexicography* 14(1): 1–20.

Chafe, W. 1968. Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language* 4: 100–127.

Cheng, W., Greaves, C. & Warren, M. 2006. From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics* 11(4): 411–433.

Chomsky, N. 1964. *Aspects of the Theory of Syntax*. Cambridge MA: The MIT press.

Church, K. & Hanks, P. 1990. Word association norms: Mutual information and lexicography. *Computational linguistics* 16(1): 22–29.

Clear, J. 1993. From Firth principles: Computational tools for the study of collocation. In *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis & E. Tognini Bonelli (eds), 271–292. Amsterdam: John Benjamins.

Considine, J. 1996. The *meanings*, deduced logically from the etymology. In *EURALEX '96 Proceedings*, M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström & C. R. Papmehl (eds), 365–371. Göteborg: Göteborg University Press.

Coulson, S. 2000. *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge: CUP.

Cowie, A. 1998a. Introduction. In *Phraseology: Theory, Analysis, and Applications*, A. P. Cowie (ed.), 3–20. Oxford: OUP.

Cowie, A. 1998b. Phraseological dictionaries – some east-west comparisons. In *Phraseology: Theory, Analysis, and Applications*, A. P. Cowie (ed.), 209–228. Oxford: OUP.

Cresswell, J. 2007. *The Cat's Pyjamas*. London: Penguin.

Cruse, D. A. 1986. *Lexical Semantics*. Cambridge: CUP.

D'Andrade, R. & Egan, M. 1974. The colors of emotions. *American Ethnologist* 1: 49–63.

Davies, I. & Corbett, G. 1994. The basic color terms of Russian. *Linguistics* 32: 65–89.

De Knop, S. 1985. Linguistic and extralinguistic aids for reconstruction and interpretation of metaphors in headlines. In *The Ubiquity of Metaphor: Metaphor in Language and Thought* [Current Issues in Linguistic Theory 29], W. Paprotté & R. Dirven (eds), 243–262. Amsterdam: John Benjamins.

Deignan, A. 1999. Corpus-based research into metaphor. In *Researching and Applying Metaphor*, L. Cameron & G. Low (eds), 177–199. Cambridge: CUP.

Deignan, A. 2010. The evaluative properties of metaphors. In *Researching and Applying Metaphor in the Real World* [Human Cognitive Processing 26], G. Low, Z. Todd, A. Deignan & L. Cameron (eds), 357–374. Amsterdam: John Benjamins.

Dilts, P. & Newman, J. 2006. A note on quantifying 'good' and 'bad' prosodies. *Corpus Linguistics and Linguistic Theory* 2: 233–242.

Edmonds, P. & Hirst, G. 2002. Near-synonymy and lexical choice. *Computational Linguistics* 28(2): 105–144.

Élouard, P. 1929. La terre est bleue. *L'amour la poésie*. Paris: Gallimard.

Fauconnier, G. & Turner, M. 1998. Conceptual integration networks. *Cognitive Science* 22: 133–187.

Fernando, C. 1996. *Idioms and Idiomaticity*. Oxford: OUP.

Fernando, C. & Flavell, R. 1981. *On Idiom: Critical Views and Perspectives* [Exeter Linguistic Series]. Exeter: University of Exeter.

Finke, R. A. 1989. *Principles of Mental Imagery*. Cambridge MA: The MIT Press.

Firth, J. R. 1935 [1957]. The technique of semantics. In *Papers in Linguistics 1934-1951*, J. R. Firth (ed.), 7–33. London: OUP.

Firth, J. R. 1935 [1957]. The use and distribution of certain English sounds. In *Papers in Linguistics 1934–1951*, J. R. Firth (ed.), 34–46. London: OUP.

Firth, J. R. 1950 [1957]. Personality and language in society. In *Papers in Linguistics 1934–1951*, J. R. Firth (ed.), 177–189. London: OUP.

Firth, J. R. 1951 [1957]. Modes of meaning. In *Papers in Linguistics 1934–1951*, J. R. Firth (ed.), 190–215. London: OUP.

Firth, J. R. 1952. Linguistic analysis as a study of meaning. In *Selected Papers of J. R. Firth 1952–1957*, F. R. Palmer (ed.), 12–26. London: Longmans.

Firth, J. R. 1956. Descriptive linguistics and the study of English. In *Selected Papers of J. R. Firth 1952–1957*, F. R. Palmer (ed.), 96–113. London: Longmans.

Firth, J. R. 1957. A Synopsis of linguistic theory 1930-55. In *Selected Papers of J. R. Firth 1952–1957*, F. R. Palmer (ed.), 168–205. London: Longmans.

Francis, G. 1993. A corpus-driven approach to grammar – principles, methods and examples. In *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis & E. Tognini Bonelli (eds), 137–156. Amsterdam: John Benjamins.

Frazier, L. & Rayner, K. 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language* 33: 39–68.

Games, A. 2006. *Balderdash and Piffle*. London: BBC books.

Garza-Cuarón, B. 1991. *Connotation and Meaning*. Berlin: Mouton de Gruyter.

Gerrig, R. J. & Gibbs, R. W. 1988. Beyond the lexicon: Creativity in language production. *Metaphor and Symbolic Activity* 3(1): 1–19.

Gibbs, R. W. 1994. *The Poetics of Mind. Figurative Thought, Language, and Understanding*. Cambridge: CUP.

Gibbs, R. W. 2001. Evaluating contemporary models of figurative understanding. *Metaphor and Symbol* 16: 317–333.

Gibbs, R. W. 2002. A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics* 34: 457–486.

Gibbs, R. W., Buchalter, D., Moise, J. & Farrar, W. 1993. Literal meaning and figurative language. *Discourse Processes* 16: 387–403.

Gibbs, R. W. & Matlock, T. 1999. Psycholinguistics and mental representations. *Cognitive Linguistics* 10: 263–269.

Gibbs R. W. & Nyak, N. 1991. Why idioms mean what they do. *Journal of Experimental Psychology: General* 120: 93–95.

Gibbs, R. W., Nyak, N. & Cutting, C. 1989. How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28: 576–593.

Gibbs, R. W. & O'Brien, J. E. 1990. Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition* 36: 35–68.

Giora, R. 1997. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics* 7: 183–206.

Giora, R. 1999. On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics* 31: 919–929.

Giora, R. 2002. Literal vs. figurative language: Different or equal? *Journal of Pragmatics* 34: 487–506.

Giora, R. 2003. *On our Mind. Salience, Context, and Figurative Language*. Oxford: OUP.

Giora, R. & Fein, O. 1999. On understanding familiar and less-familiar figurative language. *Journal of Pragmatics* 31: 1601–1618.

2nd proofs

Giora, R., Fein, O., Kronrod, A., Elnatan, I., Shuval, N. & Zur, A. 2004. Weapons of mass distraction: Optimal innovation and pleasure ratings. *Metaphor and Symbol* 19(2): 115–141.

Giora, R. & Shuval, N. 2005. Beyond Figurativeness: optimal innovation and pleasure. In *Literal / Nonliteral Distinction*, S. Coulson & B. Lewandowska-Tomaszczyk (eds), 239–253. Berlin: Peter Lang.

Gläser, R. 1988. The grading of idiomaticity as a presupposition for a taxonomy of idioms. In *Understanding the Lexicon*, W. Hüllen & R. Schulze (eds), 246–79. Tübingen: Niemeyer.

Gläser, R. 1998. The stylistic potential of phraseological units. In *Phraseology. Theory, Analysis, and Applications*, A. P. Cowie (ed.), 125–143. Oxford: OUP.

Glucksberg, S. 1995. Commentary on nonliteral language: Processing and use. *Metaphor and Symbolic Activity* 10: 47–57.

Glucksberg, S. 2001. *Understanding Figurative Language: From Metaphors to Idioms*. Oxford: OUP.

Glucksberg, S. 2003. The psycholinguistics of metaphor. *TRENDS in Cognitive Science* 7(2): 92–96.

Glucksberg, S. & Keysar, B. 1993. How metaphors work. In *Metaphor and Thought,* 2nd edn., A. Ortony (ed.), 401–424. Cambridge: CUP.

Glucksberg, S. & McGlone, M. S. 1999. When love is not a journey: What metaphors mean. *Journal of Pragmatics* 31: 1541–1558.

Goatly, A. 1997. *The Language of Metaphors*. London: Routledge.

Greaves, C. 2009. *ConcGram 1.0: A Phraseological Search Engine*. Amsterdam: John Benjamins.

Greaves, C. & Warren, M. 2007. Concgramming: A computer driven approach to learning the phraseology of English. *ReCALL* 19(3): 287–306.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Gries, S. T. 2008. Phraseology and linguistic theory. In *Phraseology: An Interdisciplinary Perspective*, S. Granger & F. Meunier (eds), 3–25. Amsterdam: John Benjamins.

Gustawsson, E. 2006. Idioms Unlimited. A Study of non-canonical forms of English verbal idioms in the British National Corpus. PhD dissertation, Göteborg University.

Guthrie, D., Allison, B., Liu, W., Guthrie, L. & Wilks, Y. 2006. A closer look at Skip-Gram modelling. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, 1222–1225. Genoa, Italy.

Halliday, M. A. K. 1992. Language theory and translation practice. *Rivista internazionale di tecnica della traduzione* 0 (pilot issue): 15–25.

Hanks, P. 1986. Typicality and meaning potentials. In *ZüriLEX '86 Proceedings*, M. Snell-Hornby (ed.), 37–47. Zürich: Francke.

Hanks, P. 1987. Definitions and explanations. In *Looking Up*, J. M. Sinclair (ed.), 116–136. London: Collins.

Hanks, P. 1998. Enthusiasm and condescension. In *Euralex '98 Proceedings*, T. Fontenelle, P. Hiligsmann, A. Michiels, A. Moulin & S. Theissen (eds), 151–166. Liège: University of Liège.

Hanks, P. 2004. The syntagmatics of metaphor and idiom. *International Journal of Lexicography* 17: 245–274.

Hanks, P. 2005. Resonance and the phraseology of metaphors. In *Phraseology/phraséologie 2005: The Many Faces of Phraseology*, C. Cosme, C. Gouverneur, F. Meunier & M. Paquot (eds), 181–184. Louvain-la-neuve: Université catholique de Louvain.

Hanks, P. 2006. Metaphoricity is gradable. In *Corpus-Based Approaches to Metaphor and Metonymy*, A. Stefanowitsch & S. T. Gries (eds), 17–35. Berlin: Mouton de Gruyter.

Hanks, P. & Ježek, E. 2008. Shimmering lexical sets. In *Proceedings of the XIII EURALEX International Congress*, E. Bernal & J. DeCesaris (eds), 391–402. Barcelona: Universitat Pompeu Fabra.

Hausmann, F. J. 1984. Wortschatzlernen ist Kollokationslernen. *Praxis des Neusprachlichen Unterrichts* 31: 395–406.

Herrera Soler, H. & White, M. 2007. A contrastive view of British and Spanish business press headlines. *Rassegna Italiana di Linguistica Applicata* 39: 295–316.

Hidalgo Tenorio, E. 2009. The metaphorical construction of Ireland. In *Politics, Gender and Conceptual Metaphor*s, K. Ahrens (ed.), 112–136. Basingstoke: Palgrave Macmillan.

Hillert, D. & Swinney, D. 2001. The processing of fixed expressions during sentence comprehension. In *Conceptual and Discourse Factors in Linguistic Structure*, A. Cienki, B. J. Luca & M. B. Smith (eds), 107–121. Stanford CA: CSLI.

Hockett, C. F. 1958. *A Course in Modern Linguistics*. New York NY: Macmillan.

Hoey, M. 2005. *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hunston, S. 2007. Semantic prosody revisited. *International Journal of Corpus Linguistics* 12: 247–268.

Jack, A. 2004. *Red Herrings and White Elephants*. London: Metro.

Jones, M. (ed.). 2000. *The New Cambridge Medieval History,* Vol VI: *c. 1300 – c. 1415.* Cambridge: CUP.

Kerbrat-Orecchioni, C. 1977. *La Connotation*. Lyons: Presses Universitaires de Lyon.

Keysar, B. & Bly, B. M. 1999. Swimming against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics* 31: 1559–1578.

Kilgarriff, A. 2005. Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory* 1(2): 263–276.

Kövesces, Z. & Szabó, P. 1996. Idioms: A view from cognitive semantics. *Applied Linguistics* 17(3): 326–355.

Lakoff, G. 1986. A figure of thought. *Metaphor and Symbolic Activity* 1(3): 215–225.

Lakoff, G. 1987a. Image metaphors. *Metaphor and Symbolic Activity* 2(3): 219–222.

Lakoff, G. 1987b. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago IL: University of Chicago Press.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live By*. Chicago IL: Chicago University Press.

Leech, G. 1974. *Semantics*. Harmondsworth: Penguin.

Lin, D. & Pantel, P. 2001. Induction of semantic classes from natural language text. In *Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining* 2001, 317–322. <http://www.cs.ualberta.ca/~lindek/papers.htm>

Louw, B. 1993. Irony in the text or insincerity in the writer?: The diagnostic potential of semantic prosodies. In *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis & E. Tognini Bonelli (eds), 157–176. Amsterdam: John Benjamins.

Louw, B. 1997. The role of corpora in critical literary appreciation. In *Teaching and Language Corpora*, A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (eds), 240–251. London: Longman.

Louw, B. 2000a. Contextual prosodic theory: Bringing semantic prosodies to life. In *Words in Context: A Tribute to John Sinclair on his Retirement*, C. Heffer & H. Sauntson (eds), 48–94. Birmingham: University of Birmingham. (Reprinted in *Texto!* 13(1). <http://www.revue-texto.net/docannexe/file/124/louw_prosodie.pdf>).

Louw, B. [W. E.] 2000b. Some implications of progressive delexicalisation and semantic prosodies for Hallidayan metaphorical modes of expression and Lakoffian 'Metaphors we Live

2nd proofs

By'. Privately-distributed version of 'Progressive delexicalization and semantic prosodies as early empirical indicators of the death of metaphors'. Paper read at the 11th Euro-International Systemic Functional Workshop: Metaphor in systemic functional perspectives, University of Gent (Belgium), 14–17 July 1999.

Louw, B. 2005. Collocation as the determinant of verbal art. In *Verbal art*, D. Miller & M. Turci (eds), 149–180. London: Equinox.

Louw, B. 2007. Truth, literary worlds and devices as collocation. In *Corpora in the Foreign Language Classroom*, E. Hidalgo, L. Quereda & J. Santana (eds), 329–362. Amsterdam: Rodopi.

MacArthur, F. 2005. Whose hand and what is it holding? HAND FOR CONTROL revisited. In *Proceedings of the Third Interdisciplinary Workshop on Corpus-Based Approaches to Figurative Language*, A. Wallington, J. Barnden, S. Glasbey, M. Lee & L. Zhang (eds), 35–41. Birmingham: University of Birmingham.

Makkai, A. 1972. *Idiom Structure in English*. The Hague: Mouton.

Martin, J. 2000. Beyond exchange: Appraisal systems in English. In *Evaluation in Text*, S. Hunston & G. Thompson (eds), 142–175. Oxford: OUP.

McGlone, M. S., Glucksberg, S. & Cacciari, C. 1994. Semantic productivity and idiom comprehension. *Discourse Processes* 17: 67–190.

Mel'čuk, I. 1995. Phrasemes in language and phraseology in linguistics. In *Idioms: Structural and Psychological Perspectives*, M. Evereart, E. van der Linden, A. Schenk & R. Schreuder (eds), 167–232. Hillsdale NJ: Lawrence Erlbaum Associates.

Mel'čuk, I. 1998. Collocations and lexical functions. In *Phraseology. Theory, Analysis, and Applications*, A. Cowie (ed.), 23–53. Oxford: OUP.

Moon, R. 1992. There is reason in the roasting of eggs: a consideration of fixed expressions in native-speaker dictionaries. In *Euralex '92 Proceedings*, H. Tommola, K. Varantola, T. Salmi-Tolonen & J. Schopp (eds), 493–502. Tampere: University of Tampere.

Moon, R. 1996. Data, description and idioms in corpus lexicography. In *EURALEX '96 Proceedings*, M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström & C. R. Papmehl (eds), 245–256. Göteborg: Göteborg University Press.

Moon, R. 1998. *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon.

Morley, J. & Partington, A. 2009. A few *frequently asked questions* about semantic – or *evaluative* – prosody. *International Journal of Corpus Linguistics* 14: 139–158.

Musolff, A. 2006. Metaphor scenarios in public discourse. *Metaphor and Symbol* 21: 23–38.

Niemeier, S. 1998. Colourless green ideas metonymise furiously. *Rockstocker Beträge zur Sprachwissenschaft* 5: 119–146.

O'Brien, P. 1979. *The Fortune of War*. London: Collins.

Ortony, A., Schallert, D., Reynolds, R. & Antos, S. 1978. Interpreting metaphors and idioms: some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior* 17: 465–478.

Özgen, E. & Davies, I. 1998. Turkish color terms: Tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics* 35–36: 919-956.

Partington, A. 1993. Corpus evidence of language change: The case of the intensifier. In *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis & E. Tognini Bonelli (eds), 177–192. Amsterdam: John Benjamins.

Partington, A. S. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching* [Studies in Corpus Linguistics 2]. Amsterdam: John Benjamins.

Partington, A. S. 2004. 'Utterly content in each other's company': Semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9: 131–156.

Peleg, O., Giora, R. & Fein, O. 2001. Salience and context effects: Two are better than one. *Metaphor and Symbol* 16: 173–192.

Philip, G. 2000. An idiomatic theme and variations. In *Words in Context: A Tribute to John Sinclair on his Retirement* [ELR Monograph 18], C. Heffer & H. Sauntson (eds), 221–233. Birmingham: University of Birmingham.

Philip, G. 2003. Connotation and Collocation: A Corpus-based Investigation of Colour Words in English and Italian. PhD dissertation, University of Birmingham. <http://amsacta.cib.unibo.it/archive/00002266/01/Thesis.pdf>

Philip, G. 2006. Connotative meaning in English and Italian colour-word metaphors. *Metaphorik* 10: 59–93.

Philip, G. 2008. Reassessing the Canon: 'Fixed' phrases in general reference corpora. In *Phraseology: An Interdisciplinary Perspective*, S. Granger & F. Meunier (eds), 95–108. Amsterdam: John Benjamins.

Philip, G. 2009. Why prosodies aren't always present: Insights into the idiom principle. In *Proceedings of the Corpus Linguistics Conference CL2009*, M. Mahlberg, V. González-Díaz & C. Smith (eds.). Liverpool: University of Liverpool. <http://ucrel.lancs.ac.uk/publications/CL2009/317_FullPaper.rtf>

Philip, G. 2010a. 'You have your mother's eyes': Collocational resonance and symbolism in *Harry Potter*. Paper presented at RaAM 8 – 8th Conference on Researching and Applying Metaphor. Amsterdam 30th June – 3rd July 2010.

Philip, G. 2010b. Metaphorical keyness in specialised corpora. In *Keyness in Texts* [Studies in Corpus Linguistics 41], M. Bondi & M. Scott (eds), 185–203. Amsterdam: John Benjamins.

Piao, S., Rayson, P., Archer, D. & McEnery, T. 2005. Comparing and combining a semantic tagger and a statistical tool for MWE extraction. *Computer Speech and Language* 19: 378–397.

Piirainen, E. 2008. Figurative phraseology and culture. In *Phraseology: An Interdisciplinary Perspective,* S. Granger & F. Meunier (eds), 207–228. Amsterdam: John Benjamins.

Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol* 22(1): 1–39.

Rayson, P. 2007. Wmatrix: A web-based corpus processing environment. Computing Department, Lancaster University. <http://www.comp.lancs.as.uk/ucrel/wmatrix>

Reddy, M. J. 1979 [1993]. The conduit metaphor: A case of frame conflict in our language about language. In *Metaphor and Thought,* 2nd edn., A. Ortony (ed.), 164–201. Cambridge: CUP.

Renouf, A. 2007. Corpus development 25 years on: From super-corpus to cyber-corpus. In *Corpus linguistics 25 Years On*, R. Facchinetti (ed.), 27–49. Amsterdam: Rodopi.

Renouf, A. & Sinclair, J. M. 1991. Collocational frameworks in English. In *English Corpus Linguistics. Studies in Honour of Jan Svartvik*, K. Aijmer & B. Altenberg (eds), 128–143. London: Longman.

Richards, I. A. 1936. *The Philosophy of Rhetoric*. London: OUP.

Ritchie, L. D. 2006. *Context and Connection in Metaphor*. Basingstoke: Palgrave Macmillan.

Ritchie, L. D. & Dyhouse, V. 2008. Hair of the frog and other empty metaphors: The play element in figurative language. *Metaphor and Symbol* 23(2): 85–107.

2nd proofs

Rosch, E. 1973. On the internal structure of perceptual and semantic categories. In *Cognitive Development and the Acquisition of Language*, T. E. Moore (ed.), 111–144. New York: Academic Press.

Rossini Favretti, R. 2000. Progettazione e costruzione di un corpus di italiano scritto: CORIS/CODIS. In *Linguistica e informatica: Corpora, multimedialità e percorsi di apprendimento*, R. Rossini Favretti (ed.), 39–56. Rome: Bulzoni.

Rowling, J. K. 1998. *Harry Potter and the Chamber of Secrets*. London: Bloomsbury.

Rowling, J. K. 2003. *Harry Potter and the Order of the Phoenix*. London: Bloomsbury.

Schmid, H. J. 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Mouton de Gruyter.

Schweigert, W. A., Cintron, J., Sullivan, K., Ilic, E., Ellis, S., Dobrowits, C. & Roberts, C. 2003. Novel Figurative phrases and idioms: Phrase characteristics over multiple presentations. *Journal of Psycholinguistic Research* 32(4): 455–475.

Scott, M. 1998. *WordSmith Tools* 4.0. Oxford: OUP.

Siepmann, D. 2005. Collocation, colligation and encoding dictionaries. Part I: Lexicological aspects. *International Journal of Lexicography* 18(4): 409–443.

Sinclair, J. M. (ed.). 1987. *Looking up*. London: Collins.

Sinclair, J. M. 1991. *Corpus, Concordance, Collocation*. Oxford: OUP.

Sinclair, J. M. 1992. Trust the text: The implications are daunting. In *Recent Advances in Systemic Linguistics*, M. Davies & L. Ravelli (eds), 5–19. London: Pinter.

Sinclair, J. M. 1996a. The empty lexicon. *International Journal of Corpus Linguistics* 1(1): 99–119. (Reprinted in J. M. Sinclair. 2004. *Trust the Text: Language, Corpus and Discourse,* 149–163. London: Routledge.)

Sinclair, J. M. 1996b. The search for units of meaning. *Textus* 9: 75–106. (Reprinted in J. M. Sinclair. 2004. *Trust the Text: Language, Corpus and Discourse,* 24–48. London: Routledge.)

Sinclair, J. M. 1998. The lexical item. In *Contrastive Lexical Semantics*, E. Weigand (ed.), 1–24. Amsterdam: John Benjamins.

Sinclair, J. M. 2003. *Reading Concordances*. London: Longman.

Sinclair, J. M., Jones, S. & Daley, R. 2004 [1970]. *English Collocation Studies: The OSTI Report*, R. Krishnamurthy (ed.). London: Continuum.

Sinclair, J. McH. & Mauranen, A. 2006. *Linear Unit Grammar: Integrating Speech and Writing*. Amsterdam: John Benjamins.

Sinclair, J. M. & Teubert, W. 2004. Interview with John Sinclair conducted by Wolfgang Teubert. In *English Collocation Studies: The OSTI Report*, R. Krishnamurthy (ed.), xvii–xxix. London: Continuum.

Steen, G. 1994. *Understanding Metaphor in Literature*. London: Longman.

Steen, G. 2002. Towards a procedure for metaphor identification. *Language and Literature* 11(1): 17–33.

Stefanowitsch, A. 2006. Words and their metaphors: A corpus-based approach. In *Corpus-based Approaches to Metaphor and Metonymy*, A. Stefanowitsch & S. T. Gries (eds), 63–62. Berlin: Mouton de Gruyter.

Stewart, D. 2009. *Semantic Prosody: A Critical Evaluation*. London: Routledge.

Stubbs, M. 1995. Corpus evidence for norms of lexical collocation. In *Principle and Practice in Applied Linguistics*, G. Cook & B. Seidlhofer (eds), 245–56. Oxford: OUP.

Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Svensson, M. H. 2008. A very complex criterion of fixedness. In *Phraseology: An Interdisciplinary Perspective*, S. Granger & F. Meunier (eds), 81–93. Amsterdam: John Benjamins.

Tabossi, P. & Zardon, F. 1993. The activation of idiomatic meaning in spoken language comprehension. In *Idioms: Processing, Structure, and Interpretation*, C. Cacciari & P. Tabossi (eds), 145–163. Hillsdale NJ: Lawrence Erlbaum Associates.

Telija, V. N. & Doroshenko, A. V. 1992. The motivational basis in the semantics of idioms and ways of its presentation in the computer database. In *Euralex '92 Proceedings*, H. Tommola, K. Varantola, T. Salmi-Tolonen & J. Schopp (eds), 433–439. Tampere: University of Tampere.

Telija, V. N., Bragina, N., Oparina, E. & Sandomirskaya, I. 1998. Phraseology as a language of culture: Its role in the representation of a collective memory. In *Phraseology: Theory, Analysis, and Applications*, A. Cowie (ed.), 55–75. Oxford: OUP.

Temple, J. & Honeck, R. 1999. Proverb comprehension: The primacy of literal meaning. *Journal of Psycholinguistic Research* 28: 41–70.

Tennyson, A. 1850. *In Memoriam A.H.H.* London: Moxon.

Teubert, W. 1999. Corpus linguistics – a partisan view. *International Journal of Corpus Linguistics* 4(1): 1–16.

Teubert, W. 2007. Corpus linguistics and lexicography. In *Text Corpora and Multilingual Lexicography*, W. Teubert (ed.), 109–133. Amsterdam: John Benjamins.

Titone, D. A. & Connine, C. M. 1999. On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* 31: 1655–1674.

Tognini Bonelli, E. 2001. *Corpus Linguistics at Work* [Studies in Corpus Linguistics 6]. Amsterdam: John Benjamins.

Tognini Bonelli, E. 2002. Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach. In *Lexis in Contrast* [Studies in Corpus Linguistics 7], B. Altenberg & S. Granger (eds), 73–95. Amsterdam: John Benjamins.

Uusküla, M. 2006. Colour term distribution in Ostwald's colour space in Estonian, Finnish, Hungarian, Russian and English. *TRAMES* 10(2): 152–168.

Uusküla, M. 2007. The basic colour terms of Finnish. *SKY Journal of Linguistics* 20: 367–397.

Whitsitt, S. 2005. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics* 10(3): 283–305.

Williams, G. 2008. The good lord and his works: A corpus-driven study of collocational resonance. In *Phraseology: An Interdisciplinary Perspective*, S. Granger & F. Meunier (eds), 159–173. Amsterdam: John Benjamins.

Williams, G. & Millon, C. 2009. The general and the specific: Collocational resonance of scientific language. In *Proceedings of the Corpus Linguistics Conference CL2009*, M. Mahlberg, V. González-Díaz & C. Smith (eds). Liverpool: University of Liverpool. <http://ucrel.lancs.ac.uk/publications/CL2009/129_FullPaper.doc>

Wilton, D. 2004. *Word Myths: Debunking Linguistic Urban Legends*. Oxford: OUP.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Wray, A. 2002. *Formulaic Language and the Lexicon*. Oxford: Oxford University Press.

Wray, A. 2008. *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.

Zipf, G. K. 1935. *The Psychobiology of Language*. Boston MA: Houghton Mifflin.

2nd proofs

Reference works consulted

Collins COBUILD English Dictionary for Advanced Learners, 3rd edn. on CD-ROM. Sinclair, J. M. (ed.). 2001. Glasgow: HarperCollins.

Dictionary of Idioms and their Origins. Flavell, L. & Flavell, R. 1992. London: Kyle Cathie.

Il Ragazzini: Dizionario inglese italiano – italiano inglese, 3rd edn. Ragazzini, G. (ed.). 1995. Bologna: Zanichelli.

Lo Zingarelli 2002. Zingarelli, N. (ed.). 2001. *Vocabolario della lingua italiana.* Bologna: Zanichelli.

OED *The Oxford English Dictionary,* 2nd edn. Simpson, J. A. & Weiner, E. S. C. (eds). 1989. Oxford: Clarendon.

Oxford-Paravia Italian Dictionary, 2nd edn. 2006. Oxford: OUP.

Paravia Dizionario della lingua italiana. De Mauro, T. (ed.). 2002. Milan: Paravia.

The COBUILD Dictionary of Idioms. Moon, R. (ed.). 1995. London: HarperCollins.

The Wordsworth Dictionary of Idioms. Kirkpatrick, E. M. & Schwarz, C. M. 1993. Ware: Wordsworth.

Thesaurus of Traditional English Metaphors. Wilkinson, P. R. 1993. London: Routledge.

Corpora consulted

Bank of English™ <http://www.collinslanguage.com/wordbanks/Default.aspx>
   The data comes from the Bank of English corpus jointly owned by HarperCollins Publishers and the University of Birmingham. The dimensions of the corpus (expressed in running words) refer to the versions used at the time the data was gathered (1997–2003). The author expresses her gratitude for access to the data during those years, granted through the PhD programme of the (then) School of English, University of Birmingham.

WordBanks Online collocations sampler <www.collins.co.uk/Corpus/CorpusSearch.aspx>
   WordBanks online contains 57 million words of written and spoken English, from both American and British sources, from the Bank of English™. It is available online for teachers and students.

Corpus di Italiano Scritto (CorIS) <http://corpora.dslo.unibo.it/CorIS_ita.html>

The British National Corpus, version 2 (BNC World). 2001. <http://www.natcorp.ox.ac.uk/>
   Distributed by Oxford University Computing Services on behalf of the BNC Consortium. All rights in the texts cited are reserved.

Appendix

The tables here list the colour-word idioms and metaphorical collocations, attested in the Bank of English (English data) and CorIS (Italian data), which were included in the study, plus (in square brackets) those whose frequency was attested at < 5. Lack of correspondence between the two languages does not mean that there is no equivalent expression, only that the equivalent does not feature the equivalent colour word. The expressions are grouped by connotative meaning, as suggested by their etymology or 'literal' reading, and arranged in sequence from most to least basic colour, after Berlin & Kay (1969).

Table A1. Black

	English examples	Italian examples
Colour	Black and blue	–
	–	Camicia nera
Illicit affairs	Black market	Mercato nero
	–	Lavoro nero
Malice	–	Anima nera
	Bête noire	Bestia nera
	[Black heart]	Cuore nero
	Black sheep of the family	Pecora nera della famiglia
	Black humour	Umorismo nero
	Not as black as you are painted	–
Misfortune	Black day	Giorno/periodo nero
	–	Cronaca nera
	–	Miseria nera
	–	Crisi nera
Melancholy	[Black thoughts]	Pensiero nero
	Black mood	Umor nero
	–	Vedere tutto nero
	[Black picture]	Fare un quadro nero della situazione
Dirt	Pot calling the kettle black	–
Ink	In black and white	Nero su bianco
	Black mark	–

Table A2. White

	English examples	Italian examples
Colour	White elephant	–
	–	Settimana bianca
Honesty	White lie	–
	Whiter than white	Più bianco non si può
	White as (the driven) snow/whiter than snow	Bianco/candido come la neve
Purity	White wedding	–
	–	Matrimonio bianco
	–	Passare una notte in bianco
Pallor	Bleed someone white	–
White = blank	Carte blanche	Dare carta bianca
	–	Di punto in bianco

Table A3. Red

	English examples	Italian examples
Colour	Go red	Diventare rosso in viso
	Red herring	–
Warning	Red flag	Bandiera rossa
	[Red sky at night shepherd's delight]	[Rosso di sera bel tempo si spera]
	Red card	–
Importance	Red letter day	–
	(Roll out) the red carpet	(Stendere) il tappeto rosso
	Red tape	–
Finance	Bleed red ink	–
	In(to)/out of the red	Avere il conto in rosso
	(not one) red cent	–
Passion	Paint the town red	–
	Red light	Luce rossa
Political passion	Red flag	Bandiera rossa
	–	Camicia rossa
Blood	Catch red-handed	–
	Red in tooth and claw	–
Anger	See red	Vedere tutto rosso
	Red rag to a bull/ Red flag before a bull	–
	Red with rage	Rosso di rabbia
Fire/heat	–	Linea rossa
	Red hot	–

Table A4. Green

	English examples	Italian examples
Colour	Get the green light	–
	–	Essere al verde
Plant life	Green fingers/thumb	Avere il pollice verde
	The grass is always greener	L'erba del vicino è sempre più verde
	The greens	I verdi
Youth	Be green	–
Envy	Green with envy	–
	Green-eyed monster	–
Anger (bile)	–	Verde di rabbia/collera
Good luck	[Rub of the green]	–

Table A5. Yellow

	English examples	Italian examples
Colour	Yellow press	–
Cowardice	Yellow bellied	–
	Yellow streak	–
Envy	–	Giallo di invidia
Anger (bile)	–	Giallo di rabbia
Crime	–	Libro/film/romanzo giallo

Table A6. Blue

	English examples	Italian examples
Colour	Once in a blue moon	–
Sky	Out of the blue/ like a bolt from the blue	–
	Into the wide/wild blue yonder	–
Water	Clear blue water	–
	Between the devil and the deep blue sea	–
Colour of skin	Black and blue	–
	Talk a blue streak	–
	Talk until you're blue in the face	–
Depression	Feel blue	–
	Have the blues	–
Fear	Blue funk	Fifa blu
	Scream blue/bloody murder	–
Obscenity	Blue movie	Film blu
Aristocracy	Blue blood	Sangue blu
	–	Principe azzurro

Table A7. Grey

	English examples	Italian examples
Colour	Grey matter	Materia grigia
Vagueness	Grey area	Zona grigia
Monotony	–	Esistenza grigia
Age	[Grey power]	–
Power	Grey eminence/éminence grise	Eminenza grigia

Table A8. Pink

	English examples	Italian examples
Colour	Go pink (in the face)	–
Optimism	Tickled pink	Vedere tutto rosa
Health	In the pink	–
Romance	–	Letteratura/romanzo rosa
	–	Cronaca rosa
	–	Scandalo rosa
Femininity	Pink collar	–
Homosexuality	Pink economy	–

Table A9. Purple

	English examples	Italian examples
Colour	–	–
Extravagance	Purple patch	–
	Purple prose/passage	–
Homosexuality	–	[Telefono viola]

Table A10. Brown

	English examples	Italian examples
Colour	–	–
Obsequiousness	Brown-nose	–

Table A11. Orange

	English examples	Italian examples
Colour	–	–

2nd proofs