

Shedding light on SHED, CAST and THROW as nodes of extended lexical units

Maria Verde

School of English, University of Birmingham

1. Introduction

1.1 The nature of the lexical item

Considering linguistic form and meaning as two facets of the same phenomenon has been the fundamental assumption behind corpus work on phraseology. This systematic interconnection was originally proposed by Firth and then adopted by Sinclair to become the basis of the theoretical framework he developed. Such a view of language has important implications with regard to analyzing and describing meaning, in that first 'given certain formal parameters in the context of a word, it is possible to arrive at a reliable meaning by formalizing the evidence of language usage' and second 'a variation in the formal profile of a word or an expression will always lead us to a change of meaning' (Tognini-Bonelli 2002:76).

In the light of corpus evidence, Sinclair challenged in the light of corpus evidence the traditional association of the unit of meaning with the word and postulated a redefinition of it (1991, 1996, 1998), based on collocation, colligation, semantic preference and semantic prosody. This new kind of unit is called 'extended', because although it is described by means of using a word (or phrase) as a starting point, it is ultimately defined by the observed interrelation of the node and its context. Tognini-Bonelli (2001) defines meaning as 'function in context' and stresses the fact that it is essential for the extended unit of meaning to have reached its pragmatic function in order to be considered as 'functionally complete' (id.:19), which in turn draws attention to semantics and pragmatics as forming a continuum.

1.2 Delexical verb and noun combinations

The structural tendency of English towards analytic constructions is a current productive pattern constituting a vital part of its lexico-grammar and deeply rooted in the historical development of the language (Strang, 1970, Dixon, 1991). In this paper, we take an interest in periphrastic verbal expressions consisting of a verb and a noun complement, in which the overall meaning of the combination rests primarily on the noun (functioning as direct object in the logical structure), while the meaning of the verb appears to be 'reduced' or 'weakened' (when compared to its basic or prototypical meaning). Different names have been used by scholars with reference to these phraseological units, such as 'verbo-nominal phrases' (Renský, 1964), 'complex verbal structures' (Nickel, 1978), 'composite predicates' (Cattell, 1984), 'V-NP' constructions' (Stein, 1991) or 'expanded predicates' (Algeo, 1995). The most frequent of the verbs employed in these expressions and the most commonly accepted as such in the literature are *'have'*, *'take'*, *'give'*, *'do'*, *'make'*, but there appear to be several others less frequently occurring in this kind of construction, such as *'put'*, *'get'*, *'set'*, *'provide'* etc. They too have been given various names, including 'purely or mainly connective verbs', 'verbs of vague or weakened meaning' (Poutsma, 1926), 'copulas' (Curme, 1935), 'light or insignificant verbs' (Jespersen, 1933), 'function verbs' (Nickel, 1968), 'lexically empty verbs' (Live, 1973), 'general or semantically fuzzy verbs' (Leech & Svartvik, 1975) and 'empty verbs' (Allerton, 1982). We will be referring to the verbs of this type as 'delexical' and the structures they enter 'delexical structures', which is the name commonly employed by corpus linguists.

Sinclair (1991:113) introduced the term *progressive delexicalization* to refer to 'a broad general tendency of frequent words or of frequent senses of words, to have less of a clear, and independent meaning than the less frequent words or senses'. While delexicalization is a growing area of interest, there has not been much corpus research on delexical verbs as yet and their treatment in the existing literature is far from comprehensive. The aim of this paper is the examination of the behaviour of SHED, CAST and THROW¹ as verbs with a delexical tendency in naturally-occurring data taken from the ukbooks subcorpus of the Bank of English. These particular verbs are chosen as they are illustrative of how the phenomenon of delexicalization works, while at the same time they are not as frequent and/or productive as other more

¹ Capital letters represent lemmas, italics are used for forms of a lemma as well as for examples taken from the concordances, and quotes are used for semantic labelling.

'prototypical' delexical verbs. In effect, in this paper we take the view that far more verbs than the few mentioned occasionally -as well as inconsistently- in the relevant literature could be claimed to have a delexical function. Nevertheless, they are more 'peripheral' or of 'minor status' with respect to either the number of discrete units of meaning they form and/or their frequency and distribution patterns.

It was decided that 500 hundred concordances retrieved at random from the corpus would suffice for identifying typical lexicogrammatical patterning. Once yielded the concordance lines were examined manually in order to discount uses not relevant to our study. Our criteria of inclusion are much less restrictive than those postulated by the majority of linguists who have discussed this type of structures in the past; namely, we do not consider factors such as the form of the noun (i.e. morphology) or the existence of a semantically corresponding simple (i.e. non-periphrastic) verb expression as relevant; our focus is primarily on the existence or not of a unitary verbal meaning. The examples retained as instances of delexical uses were then subjected to a thorough analysis in terms of grammatical features of both the noun and verb group² as well as lexical profiles. They provide evidence that various categories of nouns forming a number of lexical sets co-occur with the delexical verbs in question. Indeterminacy of membership is occasionally observed, which we see as relating to the question of the overall function of the predicate frame. Based on the analysed data, claims about typicality of usage are made.

2 Lexicogrammatical patterning and meaning identification

2.1 SHED as node

Out of the 500 concordances that were retrieved at random from the corpus, the verb was observed to have a delexical function in 176 instances, which gives a frequency of 35%. The verb group³ is finite 53% of the time, a rather high 14% is marked as negative, with regard to voice 86% of the verb phrases are active and with regard to aspect, non-progressive and non-perfective are predominant (85% and 86% respectively). A relatively high 70% of the finite groups are in singular and all but one express the indicative mood and are in the past tense 60% of the time; the commonest person by far is the 3rd (84% of the time). Finally, some sort of modal meaning is expressed in 19% of the data.

The three distinct lexical units which were identified on the basis of SHED's collocational relation to its noun complement as well as semantic preference are the following (in bold characters):

SHED+TEAR	SHED+'light'	SHED+BLOOD
	LIGHT	SWEAT
	GLOW	
	RADIANCE	
	BRIGHTNESS	
	RAY	
	REFLECTION	
	TWILIGHT	
	SMILE	
	SHADOW	
	DARK	
	DARKNESS	

Figure 1⁴

Establishing semantic preferences among the patterns of co-selection of SHED, by virtue of identifying and grouping together its noun collocates on the grounds of shared semantic components, was rather straightforward. Fifteen different nouns co-occurred with SHED, of which 11 are one-offs. Most significant statistically is the noun object TEAR, which is the unique noun collocate of a distinct semantic

² The grammatical categories used in the analysis are taken from M.A.K. Halliday's Functional Grammar (1994). A summary of the findings for all 3 verbs is given in the form of tables at the end of the paper.

³ 'Verb group' is used here in the traditional grammatical sense and does not take into account the semantic compositeness of the verb and its noun collocates.

⁴ It has to be made clear that this is not meant to be a representation of the full meaning of the unit, but merely a schematic one consisting solely of its lexical core. All noun collocates are given here in the form of the lemma, however, it is acknowledged that those occurring only in the singular or plural form would be more accurately represented in italics.

unit accounting for 59% of the data. The second most common unit, with a frequency of 45%, is the one involving nouns whose meaning would justify assigning the label ‘light’ to their set: *light/-s, glow, brightness, radiance, ray, reflection, twilight, shadow, dark* and *darkness*. It should be observed that LIGHT is the core member of the set, whereas the rest are one-off cases (apart from *glow* which occurs 3 times) and are in either a hyponymic or an antonymic relation with it. The last unit, with a frequency of 17%, is the one invariably composed by SHED+*blood*⁵. The remaining two nouns found in the data, *smile* and *sweat*, have quite distinct meanings from the rest. However it could be argued that they may still be viewed as reflecting semantic preference, though on the basis of a rather more remote semantic similarity to other paradigmatic realisations this time. Namely, *sweat* could be grouped together with ‘blood’, since they both relate to human physiology, while *smile*, although appearing out of context to be semantically associated with ‘tear’ (carrying common reference to human feeling being facially expressed) should rather be grouped with ‘light’ on account of a simile being made contextually: ‘*she doffed her mask with a swift gesture and shed upon me a smile like a sunbeam*’.

In terms of number, 68% of the noun heads are not marked for plural and in the majority (67%) are preceded by a modifier. Specific deictics and non-specific deictics are employed with a similar frequency. A characteristic feature of the delexical frame SHED+noun appears to be the use of partitive structures, mostly complex numeratives. It is not particularly high statistically (14%), but it does not seem negligible either, if compared to the frequency of use of adjectives (12%) or classifiers (2%). Quantification is often being done through nominal expressions which have a descriptive force (e.g. ‘*its pool of light*’) and whose function in most cases is rather an ‘intensifying’ one: ‘*rivers/showers/floods/buckets of tears*’, ‘*streams/a drop of blood*’. The important thing is not the quantity they refer to but the emphasis the add: ‘*without shedding a drop of blood*’. If we compare an expression like that -say ‘*buckets of tears*’- with the possible paraphrase using a single verb -‘*cry a lot*’, the difference in expressivity and forcefulness is evident. The image evoked in the former constitutes a hyperbole and the tone of the writer could be described as humorous and/or self-sarcastic (the context supports this kind of interpretation: ‘*I was sixteen and shed buckets of tears when he left*’). Nevertheless, after examining individual delexical structures involving the same verb as well as structures with different delexical verbs, we have come to believe that generalisations about the internal structure (and consequently the overall semantic function) of multi-word semantic units bearing the same delexical verb can be misleading, and each unit has to be analysed individually in order to establish safely its typical realisations and overall function.

To illustrate this point, we will now briefly comment on the colligational behaviour of the unit SHED+TEAR -which is the most typical realisation of the frame SHED+noun, with respect to its contribution in defining the overall meaning of the unit. TEAR is used in the plural 16% of the time, while all the other nouns listed above are singular (with the exception of one instance of *lights*). It is hardly surprising, as TEAR not only is a countable noun (while *blood* is not and *light* only when denoting ‘source of artificial light’), but most importantly refers to a concrete entity that is typically perceived as ‘physically produced in more than one unit at a time’⁶. A closer inspection of the relevant concordances showed that 9 out of 16 times where singular is used, it is combined with negative polarity of the verbal group. Negative polarity also appears to be of a relative importance in determining the structure of the unit, since it is present in about 20% of its instances (in fact, it accounts for 2/3 of all instances of the SHED+noun frame where non-positive meanings are expressed). Negation seems to be emphasized in cases where the deictic *no* is used or postdeictics like in the phrases *not+SHED+ a single/ the smallest tear*. These phrases could be viewed as carrying semantic prosody by being more overtly evaluative/ subjective on the part of the narrator or ‘commentator’ of the events, expressing disapproval, surprise etc. and are meant to influence the judgment of the reader (which clearly distinguishes them from possible paraphrases such as ‘*did not cry at all*’) part of the narrator -or ‘commentator’ of the events, expressing disapproval, surprise etc. and be meant to influence the judgment of the reader.

⁵ It should be noted that -as is often the case when one tries to categorize language in use with respect to meaning/function- there are some borderline cases where, although the noun object is one of those discussed so far, arguments in favour of a delexical interpretation of the expression appear rather weak: ‘*it <=the Spanish soil> has been bought at the highest price and with the most precious coin: the Spanish blood shed here*’, ‘*I think that the tears that the doctors <...> had in mind were things like torn ligaments rather than the kind shed by Paul Gascoigne*’ or ‘*the scarlet colour of her lips reminded some Christians of the blood shed by Christ at his passion*’. The nouns in the above examples are felt to be much more ‘concrete’, an interpretation that is accentuated by (or is partly due to) the clause structure, namely non-finite v-ed in a restrictive qualifying relation with its understood subject.

⁶ It not being the most typical realisation is, in fact, surprising, and the most likely reason for which SHED+*tear* feels intuitively more idiomatic.

On the other hand, the use of singular in the 7 examples with positive polarity seems to add an extra nuance of meaning, that is aspectual: it minimises the duration of the activity, it presents it as lasting for a short time⁷ e.g. *'shedding another tear of relief'* (through the use of *another* reference is made to individual instances of the activity), *'The last clod of earth fell on the coffin. Michael finally shed a tear'* (here in line with the given aspectual meaning is the fact that the Agent restrains himself from indulging fully in it). This implied aspectual meaning can contribute contextually to a further meaning being carried by the expression, this time pragmatic, i.e. either the presentation of the activity as non-genuine, e.g. *'Hoffenberg shed a very small, very pathetic tear'*, or the downplaying of the importance or the extent of the activity, e.g. *'I've shed a tear or two myself'*. Finally, we could mention the fact that pronominal substitution of this head is observed on 5 occasions, whereas pronominalisation does not occur with the other nouns entering the construction; the concrete reference of *tear/s* to a physical object, might be used as a (partial) explanation.

Turning to the unit SHED+'light', it might be of interest to note that LIGHT appears to be typically expressing a figurative meaning -only in 7% of its instances is it used literally- and it accounts for all the SHED+noun structures with a metaphoric meaning in our data. Exceptionally, the noun *dark* is also used metaphorically in one example, where by being compared with LIGHT is readily perceived as expressing an opposite meaning: *'The philosophy does have grey areas, there are not always clear differences between right and wrong, but there is a pattern to it, there is an on-going process which I think sheds light and dark over our lives'*. Similarly, the unit SHED+*blood* typically refers to physical death, but can also have an alternative meaning as the following example illustrates: *'<she> sustained a nasty bruising on the forehead and some blood was shed', 'he undoubtedly shed blood in fighting other prisoners'*, where we find a more general reference to bodily harm (which also implies suffering as a consequence). In addition to this, on certain occasions it roughly corresponds to 'die', the disambiguation between this meaning and 'kill' being achieved co-textually through a colligation, i.e. the use of a possessive deictic in the expressions intended to carry the former meaning⁸.

Nevertheless, in order to be consistent with our treatment of these verb and noun combinations as lexical items, it would be more fitting to describe the whole unit as being used metaphorically and not only the noun collocates⁹. In relation to metaphoricity, Sinclair (1998:22) makes briefly a relevant point when discussing the word *budge*, that seems to be supported by our data, namely that the distinction between literal and non-literal uses 'hardly seems necessary', since both the literal and figurative aspects of the meaning of other words in the wider context are highly relevant too for the overall interpretation. Effectively, a literal or metaphorical interpretation is ultimately determined by adding up (or balancing out) contextual clues, although literal as well as metaphorical elements might be simultaneously present, display a degree of subtlety, and interact in a way that the intended reading of a particular piece of language might remain inconclusive. For instance, SHED+TEAR appears to us as not necessarily implying 'secreting drops of fluid' but potentially designating merely 'an emotional state of sadness', however the co-text does not seem to provide readily concrete -i.e. linguistically encoded- evidence for this. Similarly, the unit SHED+*blood*, it is noteworthy that although it typically refers to physical death, it can have an alternative meaning as the following example illustrates: *'<she> sustained a nasty bruising on the forehead and some blood was shed', 'he undoubtedly shed blood in fighting other prisoners'*, where we find a more general reference to bodily harm (which also implies suffering as a consequence). In addition to this, on certain occasions it roughly corresponds to 'die', the disambiguation between this meaning and 'kill' being achieved co-textually through a colligation, i.e. the use of a possessive deictic in the expressions intended to carry the former meaning¹⁰.

With SHED+'light' as well as with SHED+'blood' a second participant can be said to be implicated in the predicate's valency as an Affected entity. This is either suppressed linguistically or expressed by means of a prepositional phrase functioning on the surface level as a Spatial circumstantial element in the SHED+'light' unit, and as a noun qualifier (prepositional phrase introduced by *of*) in the SHED+'blood'

⁷ This semantic trait is by no means restricted to this particular delexical structure.

⁸ This by no means implies that we consider it synonymous with the verb DIE, the difference lying in the exact nature of the subject's agency among other things, but this is a complex issue that will not be dealt with here.

⁹ As far as units that have both literal and (standard) metaphorical realisations are concerned, this would raise the issue of polysemy vs. synonymy.

¹⁰ This by no means implies that we consider it synonymous with the verb DIE, the difference lying in the exact nature of the subject's agency among other things, but this is a complex issue that will not be dealt with here.

unit (e.g. *'He preferred to die, rather than shed the blood of his would-be-murderers'*). A third participant occurs occasionally as part of the SHED+'blood' unit, actualised as a prepositional phrase introduced by *for*, which, although it functions as a Circumstance of Support, is semantically very close to the role typically expressed by a Beneficiary. Lastly, the structure SHED+TEAR is also accompanied quite often by a circumstantial element (*for/over*+noun) that either refers to a person or to an event or situation which at semantic level is the entity causing a particular instance of the activity. It is also important to assess the typicality as well as the centrality of colligational tendencies such as these for each unit separately¹¹, and incorporate this kind of information in an eventual definition.

We will now be turning to the other two delexical verbs; however space limitations oblige us to restrict ourselves to a much briefer commentary.

2.2 CAST as node

214 concordances of CAST were considered instances of its delexical function and were kept for analysis, which makes a frequency of 43%, if we take into account the number of concordances initially retrieved. The analysis of the verb groups showed them to be 62% of the time finites, 99% of them have positive polarity, 88% are active and, in terms of aspect, they are typically non-progressive and non-perfective (79% and 91% respectively). The non-finites express the indicative mood 96% of the time and are marked for past 72% of the time; they are most frequently singular (74%) and as for person, the 3rd occurs 93% of the time.

The following units were identified by examining the collational patterning of CAST with respect to its noun complements and the semantic preferences realised:

CAST+ 'glance' GLANCE LOOK EYE	CAST+'shadow' SHADOW SHADE LIGHT GLOW BEAM ILLUMINATION RADIANCE REFLECTION SHEEN COLOUR	CAST+'doubt' DOUBT SUSPICION ASPERSIONS BLIGHT SLURS SCORN	CAST+'mind' MIND THOUGHTS
			CAST+SPELL
			CAST+VOTE

Figure 2

Detailed analysis showed that the most common unit (accounting for 36% of the data) is CAST+'shadow'. The words realising its semantic preference denote not only the actual form of energy but also the result of its succeeding to or not physical objects: *shadow/-s, shade, light/-s, beam, illumination, glow, radiance, sheen, reflection* and *colour* (*beam* is the only noun referring to a unit of light)¹². The next most frequent semantic unit is CAST+'glance' (accounting for 30% of the data), under which the nouns *glance/-s, look/-s, eye/-s*. We have also included here the noun *smile*, as it is semantically associated with them on the basis of its reference to bodily-facial activity, although the rest have to do with the human sensory system and it does not. In fact, at a closer inspection, all 4 nouns can be seen as carrying reference to human communication with either the environment or other people: the first two refer to the input and the third to the output of this kind of interaction. 15% of the data instantiate the unit CAST+'doubt', whose collocate constituents are *doubt/-s, suspicion, scorn, slurs, aspersions* and *blight* (with *doubt/-s* being clearly the core member of the set). The next two units CAST+SPELL and CAST+VOTE occur 6% and 7% of the time respectively. Last comes CAST+'mind' with a frequency of 4%.

The units involving nouns from the 'light' and 'glance' sets are those assigned most frequently intrinsic qualities in the form of Epithets. An interesting trait of the adjectives modifying the 'glance' set is that they are most often descriptive of 'a state of mind or emotion', e.g. *furtive, anxious, wistful, curious*

¹¹ For instance, out of the colligational features mentioned here the co-occurrence of an adverbial denoting Place with SHED+;LIGHT' is of much greater significance than the Circumstance accompanying SHED+TEAR, as it occurs 91% of the time while the latter 16% of the time.

¹² The reason for which we gave 'shadow' as a label and not 'light' is that collocates expressing the first meaning are twice as frequent. In a more fine-grained description, these nouns should better be presented as constituents of two distinct units in antonymic relation with each other.

etc. +glance/-s, fond, icy, sly, worried etc. +look/-s, hopeful, envious, critical etc. +eye/-s. A subset could be distinguished comprising adjectives with reference to speech acts: *apologetic, questioning, speculative, doubtful*. Other than that, they express ‘frequency or duration’, e.g. *frequent, quick, lingering etc.*, ‘direction’, e.g. *sidelong, backwards, retrospective etc.* and in 2 instances they have some kind of evaluative meaning: *sweet, wan +smile*. The modifiers used are highly significant with respect to the construction of the meaning of the extended unit in question, as they specify it in most cases as an act of transfer of feelings, opinion etc. of the Agent with ‘eyes’ being an instrument of communication rather than perception. Modification does not seem particularly significant in the case of the other units, where epithets are far less common: the nouns of the ‘spell’ group are modified by an epithet 30% of the time, those of ‘doubt’ 10%, the ‘vote’ ones 7% and finally for the members of the ‘mind’ group Epithet is never realised. Once again this seems to support the claim made earlier that subsuming lexicogrammatical features of individual units under a single description by virtue of them sharing the same delexical verb can distort meaning identification and misrepresent it or under-specify it in a definition.

Let us now turn to a few collocates, *feeling (of desolation), ripples, eddies* and *spray (of rainwater)*, that we found problematic in terms of categorisation, in order to demonstrate a different point. The first one is a word with a rather general meaning, which despite being specified further in the co-text through the qualifier, cannot justify the noun’s occurrence in terms of semantic preference (needless to say that a single occurrence would not justify positing a distinct unit of meaning). The same holds for the other three that arguably have a more specific lexical meaning. However, they neither cause a problem of interpretation in their co-occurrence with CAST, nor do they feel particularly unidiomatic. On the basis of having come across cases like these repeatedly in a current study we are carrying out of a great number of delexical structures, we maintain that lexical combinations of each particular delexical verb with a noun –which we call a ‘delexical frame’, share certain aspects of meaning with each other, even when they do not constitute realisations of the same unit and this often seems to originate in the ‘prototypical’ meaning of the delexical verb. Moreover, some of the units display a closer ‘semantic affinity’ than others, which makes it very difficult at times to draw the line between classifying structures as realisations of the same unit or to identify them as distinct units. Consider for instance the obvious meaning relation of the structures CAST+SHADOW and CAST+LIGHT (that we grouped together in Figure 2), or a less straightforward one connecting CAST+‘shadow’, CAST+SPELL with CAST+ripples CAST+eddies, i.e. designating some kind of energy transfer and the resulting effect of this transfer (the energy is supernatural in the case of *spell* and physical for the other two nouns. A clear component of the common semantic core of all the structures seems to be a general aspectual meaning of causation, which highlights the Agent’s role and moving from top to bottom would reveal more specific meanings present in subgroupings of structures¹³. The nature of these semantic relations is a very interesting issue but too complex to be discussed in more detail here.

We will close this section with a brief comment on the overlap that structures might display with respect to the nouns forming part of their lexical core. Our view is that their parallel existence can only be explained by them having a different function. Using the expressions CAST+LIGHT and SHED+LIGHT as an example, we can first mention that there is a clear difference of frequency –the latter being almost 4 times as frequent (and accounting for 39% of the delexical occurrences of SHED). Besides, it has 97% of the time a metaphorical meaning (i.e. ‘make clearer’), while CAST+LIGHT has a similar meaning in only 32% of its instances. In addition to this, half of the examples of SHED+LIGHT have modal meanings expressed, which could be attributed to the metaphorical meaning of the expression involving ‘an intellectual quest for knowledge or truth’, which in turn is inherently associated with a degree of uncertainty and subjectivity. On the other hand, CAST+LIGHT is much more heavily premodified -and in that respect less ‘fixed’-, relying rather heavily on adjectives as well as predicative elements for the construction of its overall meaning in context. It could be claimed that it is ultimately their discourse function -directly relating to textual distribution- that differentiates them: CAST+LIGHT is typically used in descriptive accounts of the natural environment (very common in fiction), whereas SHED+light typically refers to cognitive processes (characteristic of academic language). In effect, this point highlights the importance of examining the distribution and function of delexical structures across genres and registers in order to discover their intertextual, interdiscoursal functions, that are ultimately the purpose for which they are selected.

¹³ Wierzbicka (1982), in her discussion of the construction ‘Have a V’, argues for the existence of subtypes which must be defined in terms of semantic domains.

2.3 THROW as node

The verb THROW was found to be used delexically with a very low frequency: only 5% of its instances were delexical structures. Nevertheless, 18 different lexical items occurred as noun complements, 11 of which were one-offs. The majority of the verb groups are finite (88%) and active (92%) and primarily non-progressive and non-perfective (83% and 97% respectively). Finite clauses are marked for past reference 64% of the time. they are exclusively in the 3rd person and predominantly singular (74% of the time). Modality is encoded in 11% of the data.

THROW+‘light’ LIGHT SHADOW GLOW	THROW+‘look’ LOOK GLANCE SMILE	THROW+‘utterance’ HINT EXPRESSION ACCUSATION QUESTION
THROW+‘party’ BALL MATCH	THROW+‘doubt’ DOUBT SUSPICION	THROW+‘attack’ PUNCH BROADSIDE

Figure 3

THROW co-occurs with a wide range of noun collocates, some clearly forming sets on semantic grounds. The unit THROW+‘light’ was found to be realised in our data by the nouns *light*, *shadow* and *glow*, referring to brightness (or its absence); THROW+‘look’ by the nouns *look*, *glance* and *smile*, all three referring to a bodily function relating to human communication; THROW+‘utterance’ by *question/-s*, *accusations*, *expressions* and *hints*, where reference to interpersonal interaction is made as well, although in this case communication takes place through verbalised messages; THROW+‘party’ by *party*, *ball* and *match*, which refer to social events related to leisure and entertainment; THROW+‘doubt’ by *doubt* and *suspicion*; and lastly, THROW+‘attack’ by *punches* and *broadside*¹⁴. There are few more nouns co-occurring with *throw* in our data, which we were not able to assign to one of the previous groups, i.e. *doubt*, *spell*, *salto-mortal*, *lock*, *ball*. We could not stipulate them as distinct units either, as units have to be defined probabilistically, which necessitates sound evidence of routinization.

An important point could be made here, namely that a particular (noun)lexeme not only can form part of different extended units in its co-occurrence with different delexical verbs, but can also belong to different units among those formed with the same verb by virtue of its polysemy. On the other hand, not all of its instances are necessarily realisations of extended units (this observation, however, concerns only non-deverbal nouns). Take *ball* as an example: different senses of the lexeme are observed in ‘*pitchers who threw too many gopher balls*’ and ‘*he throws lavish balls (= ‘dances’) at his clubs*’; in ‘*Pitchers like Jesus and Socrates were experts at throwing a curve ball*’ it is understood as conveying the action performed, thus it is considered to enter a delexical structure and form part of the predicate, whereas in ‘*I got up and went to throw the ball out to the wing*’ it refers to the actual physical object and is discounted. It should also be stressed that co-text might dictate interpreting the meaning of a structure in a quite unexpected way, and could mark the structure as not being an instance of the unit that the noun in question typically realises (see for instance the comment made in section 2.1 regarding SHED+smile).

THROW has a frequency of 12.7 per million words in the Bank of English, which makes it by far the most frequent out of the three verbs examined here -CAST occurs 3.5 times per million words and SHED 1.6. It could be speculated that the overall frequency of a verb (i.e. irrespective of its meanings/functions) is a correlate of its productivity with respect to the range of delexical structures it forms: the more common it is in language in general, the more likely it is to collocate with nouns and build the core of various units extending at the level of a clause. Moreover, we see the facility with which a delexical verb lends itself to the creation of phraseological verb-noun combinations, i.e. the number of discrete units it forms part of, as indicative of its ‘degree of delexicalisation’. We take delexicalisation to refer not only to the process of the semantic weakening of a lexical item but also to its resulting state, which can be presented as a cline along which different verbs occupy different positions. In this line of

¹⁴ *Broadside* is used elliptically as the noun head: ‘with [...] and her eight eighteen-pounders she <=ship> could throw a broadside of no less than 456 pounds, more than the gundeck of a line-of-battle ship’, where an act of violence is referred to (a noun like *shot* or *attack* is substituted for).

thought, THROW is the most delexical out of the three verbs, SHED is the least delexical, while CAST comes somewhere in the middle. Certainly, the description of the use of a verb characterised by a relatively high degree of delexicalisation, needs to be grounded in more extensive data, so that more patterns can surface and observations regarding strength of attraction be made possible, a prerequisite for any linguistic description in probabilistic terms.

4. Delexical structures as instances of functionally complete units of meaning

We believe that reconciling the paradigmatic and syntagmatic axes of patterning and allowing for descriptive flexibility of meaningful choices in language, makes the model of the lexical item as put forward by Sinclair a powerful descriptive framework. His proposed new kind of lexical item though, in order to be of real value, has to be defined in a comprehensive way and to be as independent as possible. Extended units involving a delexical verb co-occurring with a noun, due to the fact that they are phraseological units at the level of the clause, are to a certain extent self-contained and thus in principle would lend themselves more easily to such an analysis of meaning. We believe that descriptions of delexical uses of verbs restricted mainly to listing possible or probable noun collocates are of limited usefulness and are doomed to fail to capture their overall meaning potential. The meaning of these extended units is constructed through the interplay of their colligational and collocational patterning and is complete -and thus becomes their linguistic function- only when more abstract connotational meanings are specified. These prosodies¹⁵ are meanings acquired through habitual usage of the units in question in recurring contexts, and they can be located at different points on the semantics-pragmatics continuum. We see delexical verb+noun units as carrying primarily interpersonal meanings of various kinds, reflecting the way the speaker (or writer) position themselves with respect to the hearer (or reader) or other entities involved in a given situation. Taking into account our theoretical standpoint according to which meaning is function, their stylistic values, established intertextually and relating directly to their currency in particular genres and registers, are also integral part of their meaning. Moreover, our view is that these linguistic expressions are an inherently subjective way of conceptualising events, carrying conventionalised aspectual characteristics of various kinds. It could be maintained that all these semantic components are indispensable parts of a functional definition of their meaning, one that would merge 'the two dimensions, the contextual one and the functional one' (Tognini-Bonelli, 2002:79). Besides, only defining them as functional units would eventually allow, on one hand, differentiating structures that exhibit a partial lexical overlap, i.e. share items of their lexical core with certain objectivity, as well as specifying with precision the meaning components that are shared between units comprising the same delexical verb.

We would like to close this section by commenting on an important characteristic of the extended units of meaning comprising delexical verbs, which can be readily observed with respect to the units we tentatively posited above. A unit of this type might be realised not only through one or more typical noun objects, but also by some other 'peripheral' nouns (consider for example the one-off collocates of SHED). Peripheral status is assigned to them on the grounds of low frequency and possible difficulty in allocating them to a lexical set (realising semantic preference), and the delexical structures they form can vary with respect to markedness, conventionality etc. It seems that the noun has to fulfil certain semantic criteria though, i.e. semantic closeness to a 'core' noun. In other words, the selection appears to be done paradigmatically not in an arbitrary way but in compliance with a meaning already expressed by an existing unit, hence reflecting semantic preference. It could be argued that the implication of this for the nature of these lexical units is that besides the greater or lesser degree of flexibility at the level of optional constituents (which by definition is part of the nature of an extended unit of meaning), they often display structural variability with respect to their internal constituents at the level of their core.

5. Conclusion

Our aim has been to demonstrate how the model proposed by Sinclair can be applied to the description of delexical structures, which lie in the largely unexplored grey area between idioms and collocations. Delexical verbs enter units, which can be variable in their lexical realisations and which might generally be

¹⁵ Stubbs (2001:66) uses the term 'discourse prosodies' but suggests that 'pragmatic prosodies' might be preferable.

considered conventional in their semantic patterning, but in fact carry meanings beyond a straightforward denotational core that need extensive, as well as detailed, corpus studies in order to be pinned down. This makes the relations used for defining the structure of the extended unit -collocation, colligation, semantic preference and semantic prosody-, a very fitting way for specifying their function. Given the limited space, we did not attempt any complete formal definition of the units discussed, but we hope to have given a glimpse on how all four categories of co-selection can be used for this purpose. In our opinion, verb delexicalisation is a far more pervasive phenomenon than is generally suggested, and delexical verbs, which we see as verbs with a 'vague' or 'general' meaning, rather than 'semantically empty', can form part of a very wide variety of distinct semantic units, which also seem to carry certain common semantic components.

References

- Algeo, J 1995 Having a look at the expanded predicate. In Aarts, B, Meyer, C (eds) *The Verb in Contemporary English*. Cambridge, Cambridge University Press, pp 203-217.
- Allerton, D 1982 *Valency and the English verb*. London, Academic Press.
- Cattell, R 1984 *Composite Predicates in English. (Syntax and Semantics 17)*. Sydney, Academic Press.
- Curme, G 1931-1935 *A grammar of the English language*. Boston, D.C. Heath & Co.
- Dixon, R 1991 *A New Approach to English Grammar, on Semantic Principles*. Oxford, Clarendon Press.
- Jespersen, O 1933 *Essentials of English Grammar*. London, George Allen & Unwin.
- Leech, G, Svartvik, J (1994, 2nd ed) *A communicative grammar of English*. London, Longman.
- Live, A 1973 The take-have phrasal in English. *Linguistics* 95: 31-50.
- Nickel, G 1968 Complex verbal structures in English. *IRAL* VII: 1-21.
- Nickel, G 1978 Complex Verbal Structures in English. In Nehls, D (ed) *Studies in Descriptive English Grammar. (Studies in Descriptive Linguistics 1)*. Heidelberg, Julius Groos, pp 63-83.
- Poutsma, H 1926 *A Grammar of Late Modern English*. Groningen, P. Noordhoff.
- Renský, M 1964 Nominal tendencies in English. *Philologica Pragensia* 7: 135-150.
- Sinclair, J 1991 *Corpus, Concordance, Collocation*. Oxford, Oxford University Press.
- Sinclair, J 1996 The Search for Units of Meaning. *Textus* IX: 75-106.
- Sinclair, J 1998 The lexical item. In Weigand, E (ed), *Contrastive Lexical Semantics*. Amsterdam and Philadelphia, Benjamins, pp 1-24.
- Stein, G 1991 The phrasal verb type 'to have a look' in modern English. *IRAL* XXIX: 1-29.
- Strang, B 1970 *A History of English*. London, Methuen & Co.
- Stubbs, M 2001 *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford, Blackwell.
- Tognini-Bonelli, E 2002 Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach. In Altenberg, B, Granger, S (eds) *Lexis in contrast: corpus-based approaches*. Amsterdam and Philadelphia: Benjamins, pp 73-95.
- Tognini-Bonelli, E 2001 *Corpus Linguistics at Work*. Amsterdam and Philadelphia, Benjamins.

List of tables

Table 1

	FINITENESS		POLARITY		VOICE		ASPECT		MODALS
	Finite	Non-f.	Pos.	Neg.	Act.	Pass.	Non-Pr.	Non-Pf.	
THROW	88%	12%	100%	-	92%	8%	83%	97%	11%
CAST	62%	38%	99%	1%	88%	12%	79%	91%	6%
SHED	53%	47%	86%	14%	86%	14%	85%	86%	19%

Table 2

	MOOD		TENSE		NUMBER		PERSON		
	Indicative	Imperative	Non-Past	Past	Singular	Plural	1st	2nd	3rd
THROW	100%	-	36%	64%	74%	26%	-	-	100%
CAST	96%	4%	28%	72%	74%	25%	3%	4%	93%
SHED	99.5%	0.5%	40%	60%	70%	30%	13%	3%	84%

Table 3

	THROW		CAST		SHED	
Number of nouns	18 > 11 one-offs		30 > 16 one offs		14 > 10 one-offs	
SINGULAR	75%		69%		70%	
PLURAL	25%		31%		30%	
PREMODIFIED	75%		85%		68%	
	Pre-deictics	7%	Pre-deictics	5%	Pre-deictics	0.8%
	Spec. deictics	14%	Spec. deictics	31%	Spec. deictics	34%
	Non-Spec. d.	42%	Non-Spec. d.	43%	Non-Spec. d.	37%
	Post-deictics	4%	Post-deictics	3%	Post-deictics	1%
	Numeratives	8%	Numeratives	10%	Numeratives	21%
	Epithets	31%	Epithets	50%	Epithets	17%
	Classifiers	11%	Classifiers	0.5%	Classifiers	2%
POSTMODIFIED	8%		5%		2%	
	Prepos. phr. as	33%	Prepos. phr. as	80%	Prepos. phr. as	100%
	Qualifiers		Qualifiers		Qualifiers	
	Clause as Q.	33%	Clause as Q.	20%	Clause as Q.	-
	Adjective as Q.	33%	Adjective Q.	-	Adjective Q.	-