

The recognition and interpretation of idioms¹

Steve Pulman

University of Cambridge Computer Laboratory

Draft 2, Aug 1986

Introduction

Here are some examples of phrases which have been described as idiomatic in both the traditional and the generative linguistic literature:

- 1 John *turned on* the light
- 2 Joe *put up with* the noise
- 3 Bill *took advantage of* her generosity
- 4 He did it *for his mother's sake*
- 5 The old curmudgeon finally *kicked the bucket*
- 6 We got him to *spill the beans*
- 7 That really *put the cat among the pigeons*
- 8 He was chasing *a red herring*
- 9 We feel they are *skating on thin ice* with that project
- 10 He shouldn't have *counted his chickens before they were hatched*

Intuitively speaking, these examples are ranged on a scale from the more grammaticalised, like *turn on*, to phrases which might legitimately be regarded as proverbs rather than idioms, (if such a distinction needs to be drawn). What they all have in common, (although this is not a defining feature of an idiom), is that their meanings are not deducible from the ordinary meanings of their components by the usual rules of compositional semantics. (This is not to deny that they are in some sense compositional). This abnormal behaviour may be for one of several reasons: perhaps the components do not in isolation have a meaning distinct from their occurrence in the idiom, as with *sake* in example 4, or *tabs* in the idiom *keep tabs on*; or it may be that they have several meanings (or one radically underdetermined one), the correct one of which can only be determined in the context of the idiom, as with *up* and *with* and *on* etc in the examples 1 and 2. More characteristically, perhaps, it may be that the components of the idiom have literal meanings, but that these are not what is involved in their interpretation as an idiom. This latter type of idiom is often characterised (metaphorically) as a frozen or dead metaphor: it is certainly the case that someone unfamiliar with the idiom can nevertheless arrive at an appropriate meaning for it by processing it as a metaphor (examples 7, 9 and 10). (Although, as one might expect if they

¹Published version: S. G. Pulman, 1993, *The recognition and interpretation of idioms*, in C. Cacciari and P. Tabossi (eds.) 'Idioms: Processing, Structure, and Interpretation' New Jersey: Laurence Earlbaum Cognitive Science Monographs, 249-270.

are treating it as a metaphor, they sometimes arrive at rational interpretations which are nevertheless not that of the idiom).

Idioms are pervasive in all styles of language use. The problem they present to the theoretical and computational linguist is not the fact that their meaning cannot be worked out by the usual mechanisms, for if it were not for other factors this could be overcome by treating them as ‘big’ lexical items to be looked up in a list in a fairly straightforward way. The problem is that unlike (most) lexical items, (most) idioms have considerable internal structure which seems to interact with the usual productive syntactic and semantic mechanisms of a language in ways that render a ‘look-up’ approach impossible in the general case. Thus a grammatical theory which does not allow for this interaction is likely to be found wanting in crucial respects: the ability of a theory to accommodate the behaviour of idioms has therefore been seen as a major argument in its favour, whether in Standard Theory transformational grammar (Fraser 1970), (Katz 1973), Generative Semantics (Newmeyer 1972), more recent transformational theories (Chomsky 1980), Lexical-Functional Grammar (Bresnan 1982) or Generalised Phrase Structure Grammar (Gazdar, Klein, Pullum, and Sag 1985). The main point of the present paper is to establish that there are some properties of idioms that are not captured satisfactorily by any of the current grammatical frameworks, and that these properties suggest that idioms are in fact not properly to be considered as part of the domain of a grammatical theory at all, but are to be described within some extra-grammatical mechanism probably to be considered part of the (non-) theory of knowledge representation or common-sense inference. If this is correct, then it follows that the succour and support the above mentioned grammatical theories have gained from their accounts of idioms is entirely illusory.

Properties of idioms

First let us review some of the properties of idioms that have figured in recent discussions of them. (Chomsky 1980:150 ff, Bresnan 1982: 46ff, Wasow, Nunberg and Sag 1983, Gazdar, Klein, Pullum and Sag 1985:236-242). The first observation is an old one: that in pragmatically appropriate contexts some idioms can appear in syntactically variant forms:

- 11 John turned the light on
- 12 Her generosity was taken advantage of
- 13 Advantage was taken of her generosity
- 14 He did it for the sake of his mother
- 15 The beans were spilled
- 16 It was thin ice that they were skating on
- 17 Chickens that are counted before they are hatched aren’t a good basis for a bank loan

But some idioms seem fairly resistant to this kind of paraphrase:

- 18 ?*The bucket was finally kicked by the old curmudgeon
19 *It was among the pigeons that he put the cat
20 *He was chasing a herring that was red

The only interpretation these sentences receive is a literal one. It has been claimed (Chomsky 1980:152) that there is a difference between syntactic constructions in this respect: that, in particular, most idioms do not appear in ‘tough-movement’ variants of basic forms even though other transformational paraphrases of them are permitted:

- 21 She is easy to take advantage of
22 * Advantage is easy to take of her

However, it seems likely that this is not a syntactic matter so much as one of finding the appropriate context in which the ‘focussing’ effect of ‘tough-movement’ can seem natural:

- 23 She’s playing hard to get
24 That’s a course which is proving hard to steer
25 That lie is difficult to nail

A second important observation about the properties of idioms is that they are in the main capable of ‘internal modification’:

- 26 I have a theoretical axe to grind
27 He cut through a lot of red tape
28 They were skating on very thin ice

Here it is the components of the idioms which are being modified: 26 means ‘I have a theoretical point to make’, not ‘theoretically, I have a point to make’ (on the most natural reading of the latter, at least). Likewise, it is the amount of red tape that is being quantified, not the action of cutting through red tape. Furthermore, these modifications feel entirely natural and straightforward: there is no sense of ‘playing with words’, humorous intent, or metalinguistic comment. Contrast these in both respects with the following examples:

- 29 He kicked the proverbial bucket
30 I’ll keep a close eye on his progress
31 That turned out to be a very red herring

29 is clearly metalinguistic; in 30, although ‘close’ syntactically modifies ‘eye’ it is felt semantically to modify the whole phrase: 30 means ‘I’ll watch his progress closely’. In 31 most people feel that there is an element of play or humour involved in arriving at the interpretation, although in order to do this some modification of

a component of the idiom is involved: we could paraphrase the result as something like ‘that was an unusually irrelevant diversion’. Now, most people would agree that dealing adequately with examples like 29-31, we have to go outside grammar to some theory dealing with the stylistic properties of language. But for the examples in 26-28, where there is no feeling that anything other than the normal processes of syntactic and semantic modification are at work, the implication is that we must assign independent meanings to the components of the idioms in order that there is something there for the modifiers to modify.

This conclusion is further reinforced by the observation that, contrary to the traditional generative wisdom (e.g. Bresnan 1982, quoting Grimshaw) parts of idioms can serve as antecedents to pronouns and ellipsis:

- 32 He turned the tables on me and then I turned them on him
- 33 They said the tide would turn, and eventually it did.
- 34 *I’ll keep an eye on him and one on her too

(This possibility seems to correlate with that of internal modification, as suggested by examples like 34.) The inference from this observation is that the components which serve as antecedents must be individually meaningful, as is usually assumed in the cases not involving idioms.

A further observation that we can add in support of this general conclusion is that some idioms can be recognised, albeit in something of a stylistically marked way, even when incomplete:

- 35 That’s a case of counting your chickens
- 36 This is just smoke without fire
- 37 That suggestion came from the bottom of the barrel

For examples like at least the second two, to assemble the correct interpretation requires the components to be individually meaningful.

Canonical form theories

Until recently, most linguistic theories, to the extent that they have tried to capture the properties of idioms at all, have adopted one version or another of what I shall call a ‘canonical form’ approach. Starting from the point of view that the idiosyncratic nature of idioms is a lexical property in some sense, and that this idiosyncrasy should only be stated once, rather than having to be repeated for all the variant forms in which the idiom may appear, these theories adopt some level of linguistic representation provided by the theory as the ‘canonical level’ for this statement. That is to say, this level provides a ‘normal form’ to which all different occurrences of the idiom can be reduced: the properties of the idiom can then be stated at this level once and for all, and the normal mechanisms provided by the theory for syntactic variation will ensure that alternative variant forms of

the idiom are related in some systematic way to the canonical form, allowing the properties of all the different forms to be deduced.

For example, in Chomsky (1980) it is proposed that idioms should be catered for by using a special class of lexical rules operating at the level of ‘D-structure’. These rules are sensitive to the presence of particular groups of lexical items and operate so as to reanalyse, or provide an alternative analysis for, independently generated structures. Thus structures like:

- 38 John [VP [V kicked][NP the bucket]]
 39 John [VP [V took][NP advantage][P of][NP Bill]]

will be reanalysed lexically as:

- 40 John [VP [V kicked the bucket]]
 41 John [VP [V took advantage of][NP Bill]]

with the idiomatic meanings being assigned as those of the newly created complex V: ‘kick the bucket’ = ‘die’, ‘take advantage of’ = ‘exploit’, roughly. The original syntactic structures are retained, though, allowing for the observed syntactic variation. The inability of some idioms to undergo certain rules can be described via a mechanism for marking the output of the idiom rules with a list of transformational processes that are allowed or disallowed. Chomsky does not go into any formal detail on the properties of these idiom rules but it seems fair to conclude on the basis of his examples that they have at least the following properties: (i) they operate on constituents, not arbitrary sections of tree; (ii) there is only one rule per idiom. The first property is required if the proposed mechanism for handling the semantics is to be viable: idioms are ‘like lexical items’ - i.e. they are single semantic units. The second is not logically required but seems to be the intent of all canonical form approaches.

Bresnan proposes a mechanism with some similar features, but here the canonical form is a level of lexical representation. Idiomatic meanings, she proposes, are to be stated in the lexicon, associated with the main functor of the idiom: thus one sense of ‘keep’, that which appears in the idiom ‘keep tabs on’ will be as follows:

- keep: V, keep-tabs-on ((subj),(on obj)); (obj form)=tabs
 \Leftarrow *passive* \Rightarrow
 keep: V, passive, keep-tabs-on ((by obj/null), (on subj)); (subj form)=tabs
 (Bresnan 1982:46ff)

The verb will have a special meaning, and will be syntactically marked as requiring the other components of the idiom. Alternative forms, such as the passive version, will be derived via the normal mechanism of lexical redundancy rules, just as for other transitive verbs. The interaction of such rules suffices

to derive all the valid alternative forms: exceptions can presumably be marked as such. Bresnan's approach, given the properties of lexical redundancy rules, seems to share, *mutatis mutandis* both of the properties (i) and (ii) of Chomsky's approach. It is a firm prediction of both theories that all the alternative variant forms in which a particular idiom can appear will be relatable to each other through the syntactic or lexical mechanisms of the grammatical theory: that all the variants can be 'derived' from the single canonical form at which the properties of the idiom are stated.

Both of these theories have their problems. Firstly, as Gazdar, Klein, Pullum and Sag point out, it is not clear how a theory which regards idioms as essentially synonymous with atomic lexical items can assign to them enough internal structure to allow for internal modification and the apparent appearance of some of their components as antecedents. Likewise, our ability to recognise partial or incomplete idioms remains unexplained on such a view.

Secondly, it seems that the predictions of these theories about the variety of idioms are false. Sets of sentences like these are (correctly) not related by regular syntactic or lexical processes in any current theory of grammar:

- 42 He turned the tables on me
- 43 The tables have turned
- 44 Now he's let the cat out of the bag
- 45 Now the cat's out of the bag
- 46 I've got some loose ends to be tied up
- 47 I'm tying up a few loose ends
- 48 A few loose ends need tying up
- 49 He really put the cat among the pigeons
- 50 Now the cat's well and truly among the pigeons

Nevertheless, in 42,43; 44,45; etc. it is intuitively obvious that it is the *same* idiom appearing in each sentence, rather than two separate, accidentally related idioms. However, this fact is not something which can even be described in current canonical form theories, for there is no canonical form at which the properties of the idiom can be represented. At best, there will have to be several different statements for the various syntactic forms in which the idiom appears. While workable, this is a clumsy and unattractive position to be forced into taking.

As well as there being no syntactic canonical form shared by examples like these, it is also true that the components of the idiom are, in some cases, not grouped together under the same constituent. There are other examples of the same thing, where the idiomatic component of a sentence consists, roughly speaking, of the subject and verb, but where the complement can vary:

- 51 The ceiling fell in on Pedro
- 52 The earth finally moved for Ernestina
- 53 Fortune smiled on Bonzo

The partial function approach

A second type of approach to the properties of idioms is taken by Wasow, Nunberg and Sag (1983), Gazdar, Klein, Pullum and Sag (1985). Their proposal is couched within a monostratal theory of syntax (and thus canonical forms are not an option) wedded to a semantic theory which is basically that of Montague (1974). Semantic interpretation consists of translation into a (theoretically dispensable) logical form which has a recursively defined model-theoretic interpretation. In general the meaning of a constituent is either a function or an argument (which may itself be a function). Meanings of complex constituents are built up by applying the meaning of some of the components to the meanings of the others, as function to argument.

Normally speaking, the interpretations of these various functions, and the predicates and constants of Intensional Logic are total functions: that is, they are defined for every object in the domain of the interpretation. Thus (to simplify) the interpretation of something like ‘table’ might be a (function from possible worlds to a) function from objects to truth values: a function that tells you, for every object, whether or not it is a table. Mathematically speaking, we can also have ‘partial functions’; functions that are defined for some arguments, but not others. A simple example of such a function might be *division*: division of X by Y is defined for all values of X and Y except where $Y = \text{zero}$. In these cases the value of $\text{divide}(X, Y) = \text{undefined}$. Thus *division* is a function that is defined for some numbers, but not all. (All except zero, in fact).

The ‘partial function’ approach to idioms advocated by Gazdar, Klein, Pullum and Sag is as follows: we assign to each of the words appearing in the idioms, as well as its usual literal meaning, if it has one, an idiomatic meaning. This is a partial function which is defined for some arguments and not others. It is defined, in fact, only for the arguments corresponding to the other components of the idiom. Thus in the case where all the components of the idiom appear in an appropriate configuration, we will end up with a meaning which is that resulting from the combination of all the components with each other. If some of the components are missing, we will have something that is not fully defined, i.e. not meaningful.

This can be informally illustrated as follows: we assign to ‘turn’ its usual meaning, which we will represent as ‘TURN’ and also an idiomatic meaning, roughly corresponding to ‘seize’. We treat the other components of an idiom in a similar way. To suggest (to us) that the idiomatic components all belong together, we assign them all a numerical subscript (this also serves to distinguish the idiomatic meanings from their literal counterparts: turning the tables is not *literally* (i.e. physically) to seize anything.) Now, the basic idea is that the grammar will assign to a sentence containing the words in question analyses corresponding to all possible combinations of the entries for the words. However, of these combinations, only those which have all the literal or all the idiomatic

senses, combined in the appropriate function-argument pattern, will have a properly defined meaning:

turn: TURN, SEIZE21 ...
the: THE, A21...
tables: TABLES, POSITION-OF-STRENGTH21...
turn the tables =
(i) TURN (THE (TABLES))
(ii) SEIZE21 (A21 (POSITION-OF-STRENGTH21))
(iii) *TURN (A21 (POSITION-OF-STRENGTH21))
(iv) * SEIZE21 (THE (TABLES))
etc.

Combinations of literal with idiomatic senses are not defined: they do not combine into a sensible meaning.

Since on this view the components of idioms are independently meaningful, it is not surprising that they can be modified, can serve as antecedents, and so on. (Those idioms that do not have these properties *are* treated essentially as complex lexical items by the GPSG account: they thus distinguish two subtypes of idioms).

There are, unfortunately, one or two problems with the partial function account. Firstly, it has to be arranged that modifiers of components of idioms are themselves defined in such a way that they can combine with the partial functions assigned as idiomatic senses. Thus for example, the modifier ‘theoretical’ in ‘a theoretical axe to grind’ must be given a special sense which allows it to combine with the idiomatic sense of ‘axe’. It would not do to simply associate the idiomatic sense of ‘axe’ some literal meaning that was already in the domain of ‘theoretical’ (like ‘point’, in the relevant sense), for that would imply that expressions like ‘I want to make a theoretical axe’ should be meaningful (with ‘axe’ meaning ‘point’). (The possibility of this depends on the non-productivity of internal modification: if an indefinite number of different modifiers can be present then they cannot all be listed with the relevant special senses in advance. Although I do not have a proof that this is so it seems highly likely). However, when we encounter such modifiers for the first time in the context of a familiar idiom, there does not seem to be any difficulty over its interpretation. But on the partial function story, this can only be because we have *already* learned a special idiom-combining sense for it. This seems implausible to me.

Secondly, as Gazdar, Klein, Pullum and Sag point out, partial functions will ‘permeate the entire lexicon’ (p 239). Literal verbs will have to be defined in such a way that they do not combine with idiomatic arguments, for otherwise we would be able to say things like ‘seize the tables’ meaning ‘seize a position of strength’. However, *some* literal verbs do seem to combine with idiomatic arguments, provided the rest of the idiom is there too:

- 54 **The light** we *saw* at the end of the tunnel encouraged us to
continue
55 He tried to break **the ice** which *inhibited our conversation*

In order to account for the interpretation of the relative clause, in which ‘the light’ is the object of ‘saw’ (rather than ‘the light at the end of the tunnel’) we need to assume, it seems, a special sense of ‘see’ which is capable of combining with idiomatic senses. The same observations, *mutatis mutandis*, applies to ‘inhibit’, which must combine with the meaning of ‘ice’, (via its anaphoric connection to ‘it’).

Gazdar, Klein, Pullum and Sag are apparently willing to acknowledge this, even though it leads to the prediction that combinations like:

- 56 The light encouraged us to continue
57 The ice inhibited our conversation

should be easily interpretable with the sense of the full idiom. However, there will be a considerable combinatorial explosion problem if this programme is carried through systematically: just consider a phrase like ‘see the light’. This will have many different analyses: a literal one, an idiomatic one involving the special sense of ‘see’ required above; a second idiomatic one in which a complete idiom is present (meaning something like: suddenly have the truth revealed), all possible combinations of these, resulting in structures with no fully defined meaning; combinations of each of these with analyses in which the ‘the’ is the ‘the’ of ‘kick the bucket’ (nb *kick a bucket), or of ‘cat among the pigeons’ (nb *cat among some pigeons), likewise yielding no fully defined meaning, etc., etc.

It is also the case that all the logical constants of Intensional Logic will have to be made systematically ambiguous between total and partial function uses. For example, if the partialness of idiom denotations is to be made to do the work it is required to do, it must be inherited by functions on partial domains. This will happen straightforwardly in most instances, by the normal compositional rules. But in some cases, the components of an idiom are in structures where they are not operating on each other as function to argument, being combined in some other structure. Consider the case of:

- 58 That put a cat among the pigeons
59 Now there’s a cat among the pigeons

Assuming that the relevant parts of the idiom are ‘a cat’ and ‘among the pigeons’, then the rules assumed for these constructions by Gazdar, Klein, Pullum and Sag will deliver structures like these (ignoring intensional operators for simplicity) for the two sentences:

- 60 put (among (the (pigeons))) (a (cat))
 61 be (a [λ (x) cat(x) & [among (the (pigeons))](x)])

In 60, ‘put’ combines with ‘among the pigeons’ to form a predicate which can apply to ‘a cat’. ‘Put’ will pass on the partialness of its argument. But in 61, there is no verb which applies to ‘among the pigeons’ and ‘a cat’ in a similar way. Thus in order to rule out combinations of literal cats with idiomatic pigeons, and such-like oddities, it seems that it must be arranged that ‘and’ is defined in such a way that it only accepts two idiomatic arguments or two literal arguments, and not combinations. This is technically possible, of course, (though the resulting adjustments to the model theory of IL are not trivial) but seems a curiously extreme position to which we should be led by what seemed a simple and intuitive analysis.

Idioms and inference

It seems to me that the way out of the inadequacies and problems of the preceding theories is to acknowledge and make use of the observation that in order for certain types of idiom to be regarded as having a ‘canonical form’, either of the syntactic-lexical type envisaged by Chomsky and Bresnan, or ‘purely semantic’, which is what the Gazdar, Klein, Pullum and Sag proposal amounts to, then we must give up the assumption that this level of canonical form corresponds to anything made available in current theories of grammar as an independently motivated level of representation. The relationship between pairs like 58 and 59, or these:

- 62 John let the cat out of the bag
 63 The cat is out of the bag
 64 He turned the tables on me
 65 The tables have turned
 66 He laid his cards on the table
 67 His cards are on the table

is not anything that could or should be captured by syntactic or lexical rules, or directly by compositional semantics: it is a relationship of entailment. In each case the first member of the pair entails the second. Clearly the same thing is true of more usual examples: the difference here is that there is a syntactic or lexical relationship paralleling the semantic one.

If this observation is correct, then several consequences follow: firstly, if we want to state the facts about an idiom just once - if we regard pairs of sentence like those above as involving the ‘same’ idiom - then that statement will have to be at some post-grammatical level of representation. Presumably in any current grammatical framework where meaning is defined partly in terms of a translation into some interpreted logical form, the mechanism for achieving this will be

meaning postulates of some kind. Secondly, it follows from the nature of some of these entailments that the properties of the idiom will have to be stated on the entailed member of the pairs above, not the entailing member. That is to say, the idiom must be ‘the cat ... out of the bag’, ‘a cat ... among the pigeons’, and not the version with the main verb in.

Let us see how such a proposal might work out in detail. We will assume some grammatical theory providing translations of sentence meanings into logical forms directly. Thus if we envisage such a grammar as used by a parser, the parser will deliver, for a given input string, one or more logical forms, if the input string was covered by the grammar.

The output of the grammar will feed into an inference mechanism of some kind. This inference mechanism will also have access to various meaning postulates governing the behaviour of all the non-logical constants of the logic that we are translating sentences into (at least as many constants as content words in the language, let us assume). Among these will be postulates giving information about the inferences arising from causal verbs like ‘let’, ‘put’, ‘lay’ etc.

68 $X \text{ let } P \rightarrow P$

69 $X \text{ let } P \rightarrow X \text{ cause } P$

70 $X \text{ put } Y \text{ locative-preposition } Z \rightarrow Y \text{ locative-preposition } Z$

71 $X \text{ put } Y \text{ locative-preposition } Z \rightarrow X \text{ cause } [Y \text{ locative-preposition } Z]$

If we now add to these rules similar in form but describing idiomatic interpretations:

72 $\text{cat}(x) \ \& \ \text{out-of}(x, \text{the bag}) \rightarrow \text{secret}(x) \ \& \ \text{revealed}(\text{someone}, x)$

73 $\text{cat}(x) \ \& \ \text{among}(x, \text{pigeons}) \rightarrow \text{disturbance}(x)$

then we will have the formal apparatus in place to enable us to deduce from

John put a cat among the pigeons

via

John caused [a cat among the pigeons]

and

a cat among the pigeons \rightarrow a disturbance

to

John caused a disturbance

(Strictly speaking, we need more axioms governing the behaviour of ‘cause’, but this raises well known problems not relevant to the discussion here). However, notice that we cannot regard our idiom rules as ordinary meaning postulates or inference rules. They are not meaning postulates, because they do not restrict the meanings of the literal interpretations on their left hand side. Nor are they

inference rules, since these require it to be the case that their antecedent is true before their consequent can be regarded as true: inference is the process of going from true premises to true conclusions. But in the case of the rules giving idiomatic interpretations it is precisely the fact that the antecedent is *not* literally true that triggers the idiomatic interpretation.

One way of thinking about this is to assume that literal interpretations of sentences are checked for truth (or plausibility) first, and if rejected, the logical form, and inferences from it are checked against the antecedents of idiom rules. If any match, then the idiomatic interpretation is instead submitted for truth or plausibility checking. This is a process which is in some respects like inference, in that expressions which are logically equivalent are treated as equivalent, even if not syntactically identical (the order of conjuncts, for example, does not matter), but the relationship of the expressions used to their semantic values in the real world is ignored. It is a process syntactically indistinguishable from inference, but without the usual semantic warrant. (There may be some grist for the methodological solipsist's mill here: Fodor 1981)

Notice that the viability of this approach depends on the ability to treat logical forms *syntactically* at some stage. If it were the case that logical forms were dispensed with as merely a convenient but not a necessary stage in going from sentences to model theoretic interpretations, as advocated by Montague (1974), Cooper (1983), among others, then, as far as I can see, it would not be possible to formulate idiom rules operating in the way we have envisaged. Since we would go directly from sentences to interpretations, there would be nothing there to trigger the substitution of an idiomatic interpretation, in the case where the interpretation was false.

Notice also that this interpretation of the recognition of idioms enables us to do without special senses for the words involved in an idiom. The idiom rules can be defined in terms of the literal meanings of the words present. This avoids the technical problems raised earlier, as well as the potential combinatorial explosion of literal-idiom hybrids caused by the partial function approach. All the words, or rather, their translations into logical constants, must be present, and they must combine in an appropriate way to produce a particular logical form for the idiom to be recognised. (Notice that this presupposes a fairly finegrained view of the relation between lexical items and logical constants: that the word is pretty much the primitive element of meaning. This view is argued for in some detail in Pulman 1983).

However, using literal senses of words to trigger the recognition of idioms brings further potential problems. For example, it will be the case that inferences from other sentences which result in the same set of constants in the same logical relationship could also cause the idiom to be found. Thus

74 John put the tabby among the pigeons

given

75 tabby (x) \rightarrow cat (x)

will entail

76 John put the cat among the pigeons

and the above quasi-inference will go through. It is not clear that this is a disadvantage, however, for it is common to hear circumlocutionary versions of common idioms for humorous effect ('then the manure really entered the ventilation system'), which presumably depend on the hearer's ability to carry out these inferences, and also reflect on the speaker's reason for making them do this extra work.

A more worrying consequence is that we might find some idioms accidentally, through the use of other lexical items connected to those triggering the idiom rules via meaning postulates. For example,

77 We watched the chandelier being switched on
might be taken to entail

78 We saw the light

which has an idiomatic reading. One way of circumventing this might be to retain some connection between the actual lexical items used and the idiom rules by developing some kind of indexing (hashing) scheme based on the occurrence of particular lexical items in the input sentence. Thus any sentence mentioning, say, cats and bags, might cause the relevant inference rule to be triggered. On sentences like

79 The cat miaowed as her owner emptied the fish out of her shopping bag
the rule would be invoked, but would not succeed. On sentences like

80 The tabby emerged from the sack with a mouthful of salmon

the idiom rule would not be invoked, even though the sentence might entail a cat being out of a bag. This would suggest that the circumlocutionary idioms mentioned earlier can only be recognised when this mechanism has been overridden: their self-conscious nature suggests that a fair degree of problem solving may be involved in recognising them.

One benefit of the proposed mechanism for lexical indexing is that it would enable us to account for the possibility of recognising incomplete idioms. If a particular sentence makes no literal sense, but the lexical items present narrow down the range of possible idioms to a manageable number, then the retrieval of the appropriate idiomatic meaning will still be possible. If recognition was only by the idiom rules, then nothing less than the complete idiom would suffice for its successful recognition.

A further consequence of the quasi-inferential treatment of idioms is that there can be no stated connection between an idiom and the various syntactic configurations that it can or cannot appear in, at least on the assumption that syntactic properties are not encoded into logical form and thus available for inference. Thus any syntactic variant which produces the same logical form should be capable of giving rise to an idiomatic inference. This is apparently not so, at least for some examples. However, I think it is possible to account for this along

lines suggested partly by Newmeyer (1972) and partly by Gazdar, Klein, Pullum and Sag.

Consider what the pragmatic function of syntactic variation is. Presumably a passive, or a clefted, or a topicalised form of a sentence is chosen, not because a proposition expressed in this way cannot be expressed any other way, but because the demands of the discourse context are such that the information contained in the proposition must be presented in a certain way. At a simple level, this much can be demonstrated by the relative oddity of discourses like the following:

81 Speaker 1: Who hit Bill?

82 Speaker 2: It was Bill who was hit by John

On the assumption, then, that syntactic forms are going to sound odd if there is no context in which the way they present their information is plausible, we can explain the oddity of many syntactic forms of idioms as follows: for a syntactic form to be appropriate, there must be a context in which it makes sense to focus or contrast the discourse entity that the construction does focus or contrast. In the case where there is no such discourse entity, oddity will result. Thus, roughly speaking,

83 *The bucket was kicked by John

sounds odd because it focusses ‘the bucket’. But ‘the bucket’ does not correspond to any discourse entity, and thus there is no context in which this will be appropriate. In other words, as suggested in a different framework by Newmeyer (1972), the sentence sounds odd for much the same reason that

84 *John was died

sounds odd. However, in an idiom like ‘spill the beans’, it is possible to associate (idiomatic) discourse entities with the components of the idiom:

85 the-beans (y) & spill(x, y) → information (y) & reveal (x, y)

and thus possible to find a context in which it makes sense to focus one or the other:

86 The beans were spilled

sounds fine, roughly because

87 The information was revealed

sounds fine. Even cases where in isolation, a syntactic variant sounds peculiar, the idiomatic meaning can be rescued by providing a context in which it makes sense to emphasise the discourse entity that is being focussed. Consider the idiom ‘drop a brick’, meaning ‘make a stupid or embarrassing mistake’. It is very difficult to recover the idiomatic interpretation from a single sentence like:

88 What John dropped was a brick.

But in contexts where the pragmatic function of the pseudo-cleft makes sense, the idiomatic interpretation is perfectly accessible:

89 When he revealed the company finances, John thought he’d dropped a bombshell, but what he’d really dropped was a brick.

(‘Drop a bombshell’ means to reveal some surprising or shocking information). It thus seems legitimate to regard the problem of syntactic variability of idioms to some future theory of information structuring in relation to syntax.

References

Bresnan, J. 1982

The Passive in Lexical Theory, in her (ed.) *The Mental Representation of Grammatical Relations*, Cambridge, Mass: MIT Press

Chomsky, N. 1980

Rules and Representations, Oxford: Basil Blackwell.

Cooper, R. 1983

Quantification and Syntactic Theory, Dordrecht: D. Reidel Publishing.

Fodor, J. A. 1981

RePresentations: Philosophical Essays on the Foundations of Cognitive Science Sussex: The Harvester Press

Gazdar, G., Klein, E., Pullum, G., Sag, I. 1985

Generalised Phrase Structure Grammar, Oxford: Basil Blackwell

Montague, R. 1974

Formal Philosophy, New Haven: Yale Univ. Press

Newmeyer, F. 1972

The Insertion of Idioms, in Chicago Linguistic Society Vol 8, University of Chicago.

Pulman, S. G. 1983

Word Meaning and Belief London: Croom Helm Ltd, and New Jersey: Ablex Publishing Corp.

Wasow, T., Sag, I., Nunberg, G. 1982

Idioms: an Interim Report, in Preprints of the Plenary session papers, XIIIth International Congress of Linguists, CIPL, Tokyo, p 87-96