Mapping Meaning onto Use

Patrick Hanks
(Brandeis University)

Introduction

The job of the dictionary maker is traditionally seen as being to compile an inventory of the words in the language, and to say what they mean, either by paraphrase in the same language, or by translation into a different language. Even this simple statement prompts hedges, however. 'An inventory of the words in the language' cannot mean 'an inventory of all the words in the language.' The vocabulary of any living language is dynamic, and a glance at the rare words in any large corpus will reveal many words that are most unlikely to recur, for example giraffishness and camelization.

Until the mid 1980s, lexicographers were reliant on individually collected citations and introspection for evidence. For this reason, it was not always easy to distinguish between a rare but conventional word on the one hand and, on the other, a dynamic coinage, no sooner coined than forgotten – unless caught by some accident in the fine mesh of a lexicographer’s trawl net. Dynamic lexical coinages are created, more often than was generally acknowledged before the advent of corpus evidence, out of the basic morphology and phonology of a living language. Some of these dynamic coinages have crept into dictionaries, as if they were conventional words. Other items are borrowed wholesale from the lexical systems of other languages, and it is not always easy to distinguish between a foreign borrowing and a term that is more or less fully naturalized.

A Small Infinite Set

For these and other reasons, the notion of an inventory of all the words in a language is incoherent. When we have recorded all the known words, there is always one more. The number of possible words in a language is infinite. But this does not mean that the actual population of words is a very large number. Infinite means ‘unbounded’, not ‘unimaginably large’ (although the two concepts are sometimes confused in colloquial usage). If we abstract away names and errors, and ad-hoc coinages, the number of conventional vocabulary words in the English language is much less than 100,000 by any reasonable definition of ‘conventional’ and ‘word’. More than three quarters of them are
nouns. The lexicon of a natural language is a small infinite set, dominated by nouns.

An even smaller infinite set is the number of meanings associated with each word. In standard dictionaries, most words are given only one or two senses. Even the most complex words are given fewer than ninety-nine senses. And close examination suggests that in many cases, multiple senses in dictionaries represent nothing more than a failure to get the right level of generalization, or confusion of differences in context with differences in meaning.

The advent of large corpora in the 1990s, including in particular the Bank of English and continuing with the British National Corpus, introduced a new perspective. Corpora of 1 million words are simply too small for the patterns of usage of individual words to be identified clearly. But in a corpus of 100 million words, a simple right- or left-sorted concordance shows clearly most of the normal patterns of usage for all words except the very rare or very unusual. Other aspects of patterning emerge through more elaborate computational analysis.

**Defining the Word or the Context?**

Despite heroic efforts by British learners' dictionaries since the 1980s, there is still much to do, lexicographically speaking, in coming to terms with corpus evidence. Old assumptions are slow to die. Dictionaries, even dictionaries for foreign learners, give equal prominence to rare and unusual words and senses, alongside ordinary conventional usage. Sometimes, indeed, the rare and usual words and senses are given prominence at the expense of the ordinary. For dictionaries aimed at native speakers, this may be a correct (or at any rate a defensible) perspective. At least part of the purpose of a native-speaker dictionary is to extend a user's understanding of unusual words. But the fact remains that giving equal prominence to all senses, when they are not equally common, is a distortion. Even when a word really does, indisputably, have more than one sense, generally just one of the senses accounts for over 80% of the uses, while the others are rare. For foreign learners and computational applications alike, lexicographic analysis needs to show a sense of proportion (or, rather, it needs to show the proportions of senses).

Corpus evidence highlights other formerly neglected issues, too, for example the pragmatics of words in use. A spoken corpus shows, for example, that in English conversation the word **right** is often used to signal conversational cooperation or acceptance of what is being said, more often than to express the truth-functional sense "factually true or correct". The expression **if you like** is
used to invite conversational cooperation or acceptance of a new concept or a
new way of expressing an idea. "This is, if you like, a widget" invites
acceptance of the term widget. It does not mean that the speaker is willing or
able to turn the object in question into something completely different if the
hearer takes a dislike to it.

Explanations in traditional dictionaries are driven by truth-functional notions
of semantics, and tend to overlook (or explain very badly) pragmatically words
and function words such as discourse organizers (e.g. anyway, however, all the
same, be that as it may), modal verbs, determiners, and pronouns. But now
lexicographers are taking a long, hard look at the different functions of words in
language in use, rather than pretending that all terms are truth-functional.

Up to now, standard dictionaries have been able to get away with wrongly
focused explanation of terms such as those just mentioned on the grounds that
(in the words of at least one dictionary publisher) "nobody will look them up
anyway" - or because the dictionary buying public is hugely trusting of the
expertise of the printed dictionary, and does not know that anything better could
be expected from the inventorists of the language. (It is as if the inventory clerks
in an electronic equipment store were required to classify all the items in the
store according to the materials out of which they were constructed, without any
reference to the different purposes for which they are used.)

But this is changing. Future developments in corpus-based lexicography will
be driven more and more by the needs of machines and of man-machine
interaction. For activities such as information retrieval, message understanding,
and machine translation, it is essential to identify the different linguistic
functions of the words of a language, to prioritize, and to tease out the apparent
complexities of language, with the aim of identifying and presenting a well-
organized account of the underlying simplicities that make communication
possible.

An important challenge in information retrieval is matching related search
terms – in the words of Greg Notess, "getting away from the tyranny of text
matching". Compiling satisfactory sets of related search terms depends on
distinguishing one sense from another. ‘Seizing a cargo’ is equivalent to
‘impounding a cargo’, but ‘seizing a city’ is equivalent to ‘occupying a city’, not
‘impounding a city’. ‘Toasting a piece of bread’ is equivalent to ‘grilling a piece
of bread’, but ‘toasting a champion’ is not equivalent to ‘grilling a champion’.
Sets of related search terms have to be matched to senses or contexts, not merely
to words.
Existing dictionaries may be guilty of sins of omission (e.g. in accounting for pragmatics and function words), but they are equally guilty of sins of commission. They can make things seem even more complicated than they really are. In part, this is because the structure of a traditional dictionary entry is dictated by meanings not by use. Word meaning (if such a thing exists at all) is extremely vague and unstable. A word can have about as many senses as a lexicographer cares to perceive. Moreover, as Atkins and Levin (1991) point out, it is often difficult or even impossible to map the semantic distinctions made by one dictionary onto those of another. More specifically, human beings have a natural tendency to define the context (without admitting, even to themselves, that they are doing so), rather than focusing on the particular contribution of the word to the contexts in which it occurs. An extreme example, mercifully cancelled before publication, was a proposed definition of *throw* as 'to behave in a wild and uncontrolled manner'. When challenged for evidence, the lexicographer who wrote it pointed to the expressions 'throw a fit' and 'throw a wobbly'.

Less striking examples can occasionally be found in published monolingual dictionaries, having slipped through the editorial net.

Two approaches are possible in attempting a solution to this problem. One, characteristic of 19th- and 20th-century monolingual dictionaries, is to distort the natural syntax of the defining language in order to force it to fit and be substitutable for the word itself and not the context. The other, characteristic of bilingual dictionaries and Cobuild, is to pay more attention to the phraseology of usage, so that a word can be defined in its most natural context(s), or indeed so that the contexts themselves, as a whole, can be explained or translated after being stated explicitly. For this to work effectively for the user, the lexicographer must first group the corpus evidence for each word according to the contexts in which it occurs, and then decide to what extent it is possible to group different contexts together (on the grounds that they express what is essentially the same meaning), and to what extent it is necessary to make distinctions.

With the advent of large corpora, it is possible to be much more precise about the typical contexts in which a word is used, and to associate different meanings with different contexts. The crucial point here is to choose, as an organizing principle for the dictionary entry, context (which is objectively observable and measurable) rather than meaning (which is opaque and depends on the perceptions of the definer). Lexicographers should think first in terms of syntax and context (or, more strictly, syntagmatics), rather than directly in terms
of semantics. They can thus approach meaning indirectly, through syntagmatic analysis, according to a motivated grouping of the evidence. Of course, traditional dictionaries already distinguish noun senses from verb senses, and transitive verbs from intransitive verbs. What is suggested here is that the organizing principle should be both more delicate and richer. The semantic classes of words ('lexical sets': see Hanks 1996) in the co-text surrounding a target word can function as the organizing principle for the lexicographic description of the target word.

Words with a valency of three will not be treated as if they have a valency of only two, so for example the obligatory adverbial of location (e.g. 'down', 'on the table', etc.) of put will not simply be ignored.

Until very large corpora became available, it was not clear that such an approach was even possible. It is now clear that it is entirely practicable.

**Contrastive Analysis in Frame Semantics**

The analysis of different senses of words in different contexts, and of different words with similar meanings, has been greatly influenced in recent years by the theory of frame semantics, developed by Fillmore and Atkins, on the basis of Fillmore's earlier work on case grammar and Atkins’s earlier work on lexical analysis. In a seminal passage, they write:

> A word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning. Speakers can be said to know the meaning of a word only by first understanding the background frames that motivate the concept that the word encodes. Within such an approach, words or word senses are not related to each other directly, word to word, but only by way of their links to common background frames and indications of the manner in which their meanings highlight particular elements of such frames.

(Fillmore and Atkins 1992: pp. 76-77)

These conceptual (semantic) frames are related to syntactic frames, giving actual expression not only to the concept expressed by the word itself, but also to other words and expressions that, typically, are put together with the word when people use it to make meanings. Central to the analysis of a syntactic frame is the notion of valency:

Each lexical item, or idiomatized phrase, can be associated with what can be called its valence description, a description that specifies, in both semantic and syntactic terms, what the expression requires of its constituent and its context, and what it contributes to the structures that contain it. (op. cit.: p. 78)
The term *valence* or *valency* was borrowed (originally by the French syntactician Lucien Tesnières, 1959) from chemistry, where it is used to express the atomic structure of molecules. Applied in linguistic analysis, it indicates the number of other terms to which a term (typically a verb) relates syntagmatically. So, for example, *die* has a valency of one: it is normally an intransitive verb with no direct object or adjunct. *Kill* has a valency of two: it expresses a relationship of the action both to a killer and a victim. *Put* has a valency of three: not only does it relate to the person doing the putting and to the physical object being put; it also relates to the place where the physical object is put.

Where linguists in the European tradition talk of valency or valence descriptions, linguists trained in the American transformational-generative tradition use the term *argument structure*. The two concepts are very similar. However, there is this difference: *argument structure*, a term taken from formal logic, necessarily implies an underlying logical form of a clause or proposition, whereas *valency* does not. For this reason, the notion of valency is both more powerful and more flexible than that of argument structure. It can, for example, deal with performatives such as *I promise* ... It can describe structures without having to transform them into a presupposed underlying logical structure. The valency of a word is simply an analysis of its relationship, actual or potential, to other words in clauses in which it is used. Relating the valence description of a clause to the logical form of a proposition is an additional step, which may or may not be taken. Given that Fillmore and Atkins want to relate clause structure to a fairly delicate system of semantic roles, *valency* (or *valence description*) is a more appropriate term for this kind of lexical analysis. “First, let’s get the syntactic structure clear,” they might say. “Then let’s relate that structure not only to a logical form but also to the whole conceptual framework within which the word exists.” Syntactic structure and semantic structure are independent variables, and relating them is an interesting challenge for lexicographers and linguists alike.

Having introduced the notion of valence descriptions, Fillmore and Atkins go on to present a case study of the word *risk*, showing how the whole context fits together in a broad semantic frame, while different contexts affect the meaning of the word. At a high level of generality, *risk* and semantically related words (including the nouns *danger*, *peril*, *hazard*, and *venture*, and the verbs *gamble*, *invest*, and *expose*) all share the notion of *the possibility of an unwelcome outcome*. But not only are there contrasts among the different words; there are also different meanings arising from different contexts or uses of the same word.
Fillmore and Atkins discuss the semantic frame of risk in two important papers: (1992) and (1994). There is some variation of terminology. In the 1994 paper (p. 367), the relevant frame elements are presented as follows:

- **Protagonist**: the central person in the frame
- **Bad**: the possible bad outcome, or harm
- **Decision**: the decision that could trigger this
- **Goal**: the desired outcome
- **Setting**: the situation within which the risk exists
- **Possession**: something or someone valued by the Protagonist and endangered by the situation
- **Source**: something or someone which could cause the harm.

These frame elements are very similar (but different in interesting respects) to those presented two years earlier:

- **Chance**: uncertainty about the future
- **Harm**: the possibility of an unwelcome development
- **Victim**: the individual who stands to suffer harm
- **Valued object**: the thing that is in danger of being lost or damaged (e.g. one’s money or good name)
- **Situation**: the state of affairs within which someone might be said to be at risk
- **Deed**: the act that brings about a risky situation
- **Actor**: the person who performs the act that puts someone or something at risk.

And four secondary categories (all concerned with Actor’s intentions):

- **(Intended) Gain**: Actor’s hoped-for gain in taking a risk (e.g. you’ve risked your health for a few cheap thrills)
- **Purpose**: Actor’s intention (if any) in doing the risk-inducing act
- **Beneficiary**: the person for whose benefit something is done
- **Motivation**: the psychological source of Actor’s behavior

It is immediately clear that frame semantics is a much richer schema for representing meaning than is used in any current dictionary. It might be objected that this schema invokes excessively subtle distinctions, some of them quite far removed from the valency of risk itself in any normal context of use. However, this would be to miss the point. The point is that, to use risk – or any other word – correctly, one must understand the whole context in which it is used, and know the roles of all the players, and that these roles, whether present or absent in any particular utterance, must be specified as part of the meaning of the word. The players, or rather, more strictly, an idealization of the playing positions, must be understood before meaning can take place.

Any cricket player who has ever read definitions of cricketing terms written by baseball-playing Americans, or any baseball player who has ever read definitions of baseball terms written by cricket players, or indeed anyone who has ever tried to write definitions of the terminology of card games, will
appreciate the point being made here by Fillmore and Atkins, which is that the meaning of a term can only be fully understood in the full context of the relevant "specific schematization of experience". To know what a googly is, the reader must understand the whole semantic frame of the game of cricket, focusing in particular on the different techniques used by bowlers in their attempts to get batsmen out. To avoid solecisms, the definer writing definitions of such terminology will ideally himself or herself be a user of the terms.

Not all of the frame elements are always explicitly present in the text on every occasion when a term is used, but in order to understand the meaning of the term fully, it is necessary to understand that whether realized in actual words or merely implied, all these roles are present.

Where does all this leave the definition writer? Well, in the first place, it reminds her or him that it is necessary to take account of the domain and semantic context in which different senses of words are used. It is often difficult for compilers of dictionaries and glossaries to decide how much knowledge can be assumed on the part of the reader, and how much explicit contextualization is necessary in a definition. A good starting point is to examine the evidence for the various semantic frames within which the words is used, and to draft classifications of the different sets of relevant frame elements accordingly.

Secondly, it reminds the definition writer to take a long hard look at the different syntactic contexts in which each word is used. I shall return to this point shortly.

Fillmore and Atkins note that the noun risk often occurs in the predicate phrase at risk (these children are at risk), with a slightly different meaning from the phrase at the risk of, where the nature of the potential Harm is specified (e.g. at the risk of seeming stupid). As a direct object it occurs most often after the verbs take and run. They comment,

We have found no dictionary that offers an account of the difference between run a risk and take a risk. (p. 85).

They draw attention to this difference: that taking a risk is a deliberate, voluntary act, while running a risk has a broader meaning, encompassing also involuntary actions, as in 1.

1. Puppies in closed cars on hot days run the risk of heat stroke.

But does this distinction belong in a dictionary, and if so does it belong under the headword risk? This is fairly typical of the kind of issues that confront the working lexicographer on a day-to-day basis, analysing corpus evidence. The lexicographer must decide, not only whether such a fine distinction justifies the amount of space that it would require, but if so, where the distinction should be
made (are there other, similar distinctions between *take* and *run*, involving different direct objects?), and above all how much time to devote to agonizing over the issue. A good lexicographical training will alert the trainee to such subtleties, at the same time as insisting on the virtues of a sense of proportion in practical lexicography.

Confirmation of the relevance of the Protagonist’s intention as an important element in the *risk* frame comes from examination of verb uses. The verb is always transitive, but the semantic roles of the direct objects can be very different, as the following examples make clear:

2. *He risked his life [[Valued Possession]] for her [[Beneficiary]].*
3. *He risked his life [[Valued Possession]] by going for a swim [[Deed]].*
4. *He risked being drowned [[Harm, Bad Outcome]].*
5. *He risked going for a swim [[Deed]].*

In 2 and 5, the Protagonist is clearly risking his life intentionally for some beneficial outcome. In 4, it seems equally clear that he is risking his life unintentionally. At the very least, the examples in a dictionary should be chosen to indicate the range of contrasting semantic roles of the typical direct objects and motivations, if any.

**Case Studies**

*Risk* is a sophisticated and complex concept. The rest of this paper consists of two much simpler case studies, illustrating how modern lexicography can approach the analysis of corpus evidence for verbs, adjectives, and nouns, and come up with results that map meaning onto use. Frame Semantics will be taken into account, but the focus will be on more basic syntagmatic distinctions. How many different senses does a given word have, and how can we distinguish one from another?

**Case Study 1: lean**

*Getting started*

The corpus is a 250 million word corpus, consisting of the “balanced and representative” British National Corpus, counterbalanced by American texts from four years of the Associated Press Newswire. Unfortunately there is as yet no American National Corpus. The word selected for analysis is *lean*, which illustrates most of the most basic issues that come up in mapping meaning onto use.
Corpus analysis of the lexicon of a language depends on sampling and concordancing. The first step is to decide what to sample. In the case of *lean* this is comparatively easy: we can start by using a regular expression (*lean*/*) to get from the corpus all the strings beginning with the four letters *l-e-a-n*. In the case of other short words, such as *set* and *go*, the process of sampling calls for more ingenuity, both to exclude words that are not part of the lemma (*setee* and *settle* are not part of the lemma *set*, nor *goat* of *go*) and to include irregular forms of the lemma (*went* has nothing in common, orthographically speaking, with *go*). Some corpus tools packages include a lemmatizer, in which all forms of a lemma are stated in advance. But even a lemmatizer cannot deal effectively with the small number of cases, e.g. *found* and *wound*, in which forms of one lemma are inextricably interwoven with forms of another. For such cases, only manual sorting on contexts (or very sophisticated context-sensitive automatic lemmatization, which comes to much the same thing) can distinguish them. But for most words in English, a fairly simple string search, with wildcards and options, usually gets satisfactory results.

To start the analysis of *lean*, I selected 510 lines at random and eliminated 10 of them, mostly oddities and lines with misprints (for example, errors for forms of *learn*), from which nothing can be learned about the meaning and use of *lean*. A small margin for irrelevant items is always a wise precaution. For practical purposes, samples smaller than 500 lines can often be sufficient to give an overview for purposes of structuring a dictionary entry.

The first step after elimination of irrelevancies is sorting by lemma. The English word *lean* is well established as both a verb and an adjective. Noun senses are more debatable, as we shall see. As it happens, the verb and the adjective have different etymologies: both words are of Germanic origin, and they were spelled and pronounced differently until the Middle English period, when they fell together. The verb is Old English *hleonian*, related to Greek *klimax* ‘ladder’; the adjective is Old English *hlaene*. But from the point of view of the analyst of conventional modern usage, the origins are irrelevant. What matters is identifying the conventions of meaning and usage (linguistic behaviour) that members of a speech community rely on in order to use words to communicate with one another.

Working with a 500-line sample, we sort all the occurrences into different categories, first on broad syntactic grounds (separating adjectives from the verbs), then into more delicate semantic and syntactic frames (e.g. separating ‘lean meat’ from ‘lean businesses’) and finally making more subtle distinctions on semantic grounds (e.g. separating different meanings of ‘lean on someone’,
according to the perceived purpose of the person doing the leaning, i.e. reliance or choice). The sorted results may be seen in Appendix I. It should be emphasized that the level of detail used in categorization of corpus lines is a matter of choice and judgement: even more delicate subcategorization is possible, or different patterns may be lumped together in a single category.

**Analysing Verb Senses**

The sortal categories required differ in matters of detail from word to word, and are dictated, at least in part, by the patterns observed in the corpus rather than by rigid a-priori sets of categories. Nevertheless, some broad a-priori generalizations are possible for most classes of words. The primary subcategorization of verbs will start with the following schema or one very like it, expressing the possible valency roles of the verb:

- **SP:** Subject - Predicate (simple intransitive): *Hoover died; lexicography lives.*
- **SPO:** Subject - Predicate - Object (simple transitive): *Chomsky killed semantics.*
- **SPOO:** Subject - Predicate - Object - Object (ditransitive): *Fred gave Chloe a baby.*
- **SPOA:** Subject - Predicate - Object - Adjunct (transitive with adjunct): *Fred gave a baby to Chloe.*
- **SPA:** Subject - Predicate - Adjunct (intransitive with adjunct): *Fred leaned on the gate.*
- **SPC_S:** Subject - Predicate - Subject Complement: *Chloe is a programmer; Chloe looks happy.*
- **SPC_O:** Subject - Predicate - Object Complement: *Fred made her pregnant; they named the child Ichabod.*

More complex schemas are required for a small number of verbs, in particular those that take a clause (in particular, reporting verbs, but also what Dixon (1992) calls 'secondary verbs').

Further subcategorization is carried out according to the typical semantic arguments that fill each of the clause roles just outlined. So, for example, a very large number of verbs in English are of pattern SPO: sense distinctions are made by grouping the words that fill the S and O roles into lexical sets according to their semantic type, enabling the lexicographer to distinguish *seizing a cargo* from *seizing someone's arm.* A fair amount of art is called for in deciding how to group the words into lexical sets: for example, the lexicographer may have to decide whether *seizing a drug smuggler carrying heroin* is more like *seizing a cargo* or more like *seizing someone's arm,* or whether these should be treated as three separate senses (splitting) or all as just one big, broad sense (lumping).

With this in mind, let us proceed with the analysis of *lean.* It is both a verb and an adjective. The verb is the pivot of the clause. If a word with an unexpected semantic value occurs in a clause, the verb coerces the meaning of
the noun more often than vice versa. Thus, any attempt to interpret an unusual sentence such as *John leaned the sky on the mantelpiece* will typically involve assigning an unusual semantic value to the direct object (sky) rather than to the verb. For this reason and others, other things being equal, it is best to start an analysis of a lexical item with verb senses lexicon with verbs.

What is an appropriate semantic frame within which to consider the verb *lean*, and what are the relevant frame elements? A distinction is often made between verbs expressing an action and verbs expressing a state, but in the case of *lean* the distinction is blurred. In 6, for example, it is not clear whether this should be interpreted as a dynamic act of leaning, or a static description.

6. *He leaned on his spade.*

It is important to note that *lean* is not merely intransitive, but that an adjunct (or adverbial), expressing position, is normally obligatory. (There are exceptions to those, in certain jargons, and we shall come to those shortly.) It shares this characteristic with other verbs of position, e.g: *lie, sit,* and *stand.*

Levin (1993: p. 112) classifies these as “verbs of putting in a spacial configuration”. Levin’s choice of name for this little group of verbs is infelicitous: “putting” implies transitivity, but *lean,* like *dangle, hang, perch, sit,* and *stand,* is a characteristically intransitive, inchoative verb. Transitive uses are accounted for by the causative alternation that is characteristic of inchoative verbs. It is not clear why Levin includes *lay* as a member of this set.

The case roles (frame elements) involved are comparatively simple:

- **Protagonist:** a human agent or a physical object (doing the leaning)
- **Theme or Patient:** the thing that gets leaned somewhere
- **Position:** the direction of the leaning, including the object leaned on, if any.

Gravity is involved in leaning. The semantic frame invokes the interrelationship of physical objects in a world governed by gravity. The protagonist is typically taller than broad. A low, flat thing does not normally lean, unless stood on its side. If the Protagonist is a human being, and the Theme is that Protagonist’s own body, then the motivation is often to take rest or relief. But it would be stretching analysis to identify motivation as a frame element. The motivation for leaning against something or in a particular direction is, as lexicographers like to say, “any of various”: e.g. to get a better view, to be heard, to reach out and grasp something, to stay on one’s feet in a high wind, or to rest. Or there may be no motivation at all.

The Theme is hardly ever expressed explicitly, for the simple reason that (unless otherwise stated), it is co-referential with the protagonist. You *can* lean a
bicycle against a lamppost, but it more typical for a person to lean \([0]\) against something (intransitive), i.e. to lean their own body against it.

Position, by contrast, is always explicit. It is normally expressed in a prepositional phrase, where the prepositional object denotes a physical object. However, object-free uses are also found. You can lean forwards, or sideways, or westwards, without actually coming into contact with anything.

In short, \textit{lean} is a very intransitive verb, much simpler and less dynamic than \textit{risk}. There is much less going on in leaning than in risking.

\textit{Exploiting the Norm}

The ways in which the basic sense of this intransitive verb have been exploited are interesting. Three patterns – phrasal verbs – have become established as separate senses in their own right:

7. \([ [[ \text{Person}] \text{ lean} \{ \text{on} \left[ ^\wedge \text{PhysObj} \right] } \} \) \(= \text{rely on}\)

7 says that if a person is leaning on something that is \textit{not} a physical object, the sense is likely to be that they rely on that thing for some purpose, as in 8.

8. The administration was leaning heavily on its argument that ....

Interpretation becomes rather complex when a person is said to be leaning on another person (see 9). This can mean that the person is putting pressure on the other person, as exemplified in 10. However, ‘leaning on someone’ can also mean relying on them.

9. \([ [[ \text{Person}] \text{ lean} \{ \text{on} \left[ \left[ \text{Person | Institution} \right] \right] } \} \)

10. ... provide benefits only by leaning on them to pay taxes.

To complicate things still further, it must be noted that a person is both a cognitive being and a physical object. It is quite possible for someone to lean on someone else in the most literal, physical sense, as is the most probable interpretation of 11 and 12.

11. Mothers leaned on children for support ...

12. one woman sobbed, leaning on her husband’s shoulder.

The co-existence of both literal and metaphorical senses of particular phraseological expressions is generally overlooked in dictionaries. If a metaphorical sense is recorded, the fact that the same words can also be interpreted literally is generally left to the reader’s imagination.

This is as a good point as any to remind ourselves that concordances may be a very useful tool for the lexical analyst, but they do violence to meaning in text. Meanings are built up as texts go along. In practice, ambiguity (in the sense of genuine doubt as to which of two possible meanings is intended, of the sort
sometimes observable in a concordance) very rarely arises in real texts: potential ambiguities are already resolved by the co-text before the potentially ambiguous word or phrase is uttered.

A third exploitation of the core meaning of lean, verb, is found with the preposition to or towards, with the sense “be likely to choose or consider choosing”. Again, it is hard to identify the criteria that make us choose one sense rather than another: ‘leaning towards someone’ admits both literal and metaphorical interpretations. Not only has this exploitation become conventionalized in its own right, but also it has given rise to a further exploitation in the jargon of American political journalism, as in 13.

13. Historically, women lean Democratic.

A Variety of Semantic Frames

All of the minor, somewhat metaphorical senses just illustrated involve a change of semantic frame. ‘Leaning on someone’, when not meant literally, is part of a social interaction frame in at least two places: with verbs of persuasion and with verbs of dependency. The corpus analyst makes these judgements on a probabilistic basis on the grounds of clues in the immediate co-text. A cluster of clues may add up to a virtual certainty, but they can never add up to a necessary condition for selecting a particular sense. Note, too, how rare these metaphorical exploitations are compared with the overwhelming frequency of the basic, physical, literal verb sense. It would be entirely reasonable to say that the verb lean has only one basic sense, which is exploited both grammatically and metaphorically in various ways in different frames. Postulating four or five independent senses, as large traditional dictionaries do, is merely misleading.

The problem of verbal nouns (and verbal adjectives)

Neglected but pervasive problems are posed by –ing forms in English. In some cases it is impossible to say whether these are verbs, nouns, or adjectives. For example, in 14, the word making is governed by a preposition (for), which implies that it must be a noun; but on the other hand it governs a direct object (beer), which means that it must be a verb.

14. The apparatus needed for making beer.

14a. The apparatus needed for beer making.

14a is no less grammatically well formed than 14, but much more nouny. Here beer, which is logically the direct object of the verb to make, is placed as a noun modifier before the verbal noun making. It may be argued that 14 and 14a are examples of a problem that may trouble text analysts but not lexicographers. These examples do not constitute a case for adding making, noun, as a
dictionary entry. However, before heaving a sigh a relief and deciding that they can safely ignore the problem of -ing forms, lexicographers should note that such global policy decisions have been responsible for some very poor lexicographical practice. The fact that the noun breeding (“the good manners regarded as characteristic of the aristocracy and conferred by heredity” – NODE) is derived as an -ing form from the verb breed does not obviate the need to give it a separate entry in a dictionary. On the other hand, it is less clear the same word in the sense “the mating and production of offspring by animals” needs to be entered in a dictionary. And placing the “good manners” sense as a semantic derivative of the “reproduction” sense is not defensible.

The expression breeding ground was systematically overlooked in standard English dictionaries for many years, presumably on the flawed argument that, because breeding in this expression can be analysed as a verbal adjective modifying ground, it is part of the grammar, not part of the lexicon. But the semantic connection between the two elements is not easily inferred from normal rules of syntax. The syntax is clear, but the meaning is opaque.

To take another example more at less at random, it is not clear why Cobuild records handling as a verbal noun derivative of handle in the transitive sense ("the military's handling of Robert's death"), but not in the inchoative sense ("the car's handling is beautiful"). NODE’s decision to omit all reference to the verbal noun handling is even more surprising. In fact, problems presented by -ing forms are pervasive in the English lexicon. A reappraisal of lexicographic policy with regard to -ing forms, in the light of corpus evidence and users’ needs, is overdue.

In the case of lean, the verbal noun is usually plural (leanings). Neither the form nor the meaning are predictable on the basis of the verb, so this is a lexicographical no-brainer. Leanings (plural) belongs in any self-respecting dictionary of English.

Analysing Adjective Senses

Next in line are adjective uses. In some linguistic analyses, adjectives are regarded as transformed versions of verbs, and in certain languages the verblike quality of adjectives is more obvious than it is in English. The approach to the analysis of adjectives can be very similar to that of verbs. An attributive adjective typically has a valency of one, i.e. its meaning is analysed in relation to the head noun that it modifies.

Predicative adjectives, on the other hand, are more like verbs. For example, glad is typically predicative, followed by a that-clause (she was glad that he had
told her), a conditional clause (I shall be glad if you could help), a to-infinitive (she was glad to hear it), or a prepositional phrase (she was glad about what had happened). As in the case of verbs, a lexicographer doing corpus analysis must decide whether these different complementations should be grouped together, as all expressing basically the same meaning of glad, or whether some semantic difference can be attached to the different syntactic patterns. In the case of glad, grouping seems appropriate. There is only one sense of glad.

However, one very particular pattern is worth noting separately used for a particular pragmatic purpose. A first-person declarative sentence in which glad is governed by a verb phrase containing the modal shall, will, should, or would plus the verb be (e.g. I would be glad to contribute) expresses an offer to do something for the benefit of someone else. Formally, it could be analysed as a prediction about a condition that would bring about the speaker’s future happiness, but of course that misses the point of the utterance. As we have already noted, conventional pragmatics of this kind were neglected in traditional lexicography.

The Adjective lean

Returning to our analysis of lean, we come to an example of another of lexicography’s pervasive and recurring problems. Is the distinction between lean meat and lean muscle in Appendix 1 justified? A fair amount of variation is possible in the level of detail in any analysis, and choices will be determined by the intended application. There is a constant danger of overinterpretation, i.e. of importing subcategorization distinctions into the syntax on the basis of the interpreter’s own beliefs. With these caveats in mind, the reader might argue that, after all, meat is muscle. But then, if we are not distinguishing the sense of lean in lean meat from its sense in lean muscle, is it worth making a distinction between lean muscles, a lean body (containing lean muscles, presumably), and a lean person (i.e. one having a lean body)? Here we have run into examples of what Apresjan (1973) calls ‘regular polysemy’. Adjectives denoting physical characteristics of a body part are regularly applied also to a person having such a body part. Likewise, adjectives denoting abstract qualities are applied equally to behaviour and to people or animals exhibiting such behaviour: a courageous act, a courageous person; a ferocious attack, a ferocious dog.

The semantic frame of an adjective is determined by the head noun whose meaning it modifies. The adjective lean modifies an interesting variety of head nouns. How to group the head nouns modified by lean? There is a continuum
from *lean meat* through *lean muscle* to *lean limbs, a lean person or animal*, and then a metaphorical leap to *lean businesses* and *lean years*.

In all but the last of these, *lean* is used with positive connotations. Lean meat is good, lean muscle is good, and a long lean look has been fashionable for decades. Only lean years are bad, and that expression is of biblical origin. Even the bad adjective *mean* acquires positive connotations when associated with *lean* in the cliché *lean and mean*, which is identified as a desirable goal for businesses and public institutions.

The use of *lean* as a term of approval makes it a particularly ready candidate for metaphorical exploitation. We find not only regular alternations (e.g. *lean muscles, lean limbs, lean people*) but also irregular exploitations (*lean songs, lean clothes, a lean trade fair*). The lexicographer has to decide which of these expression are normal and conventional (and therefore candidates for dictionary entry) and which are exploitations of the norm, ignorable for lexicographical purposes.

**Noun Forms Associated with* lean**

Before leaving *lean*, we turn to some problematic noun forms. Is *lean* itself a noun? And what about derivatives? The orthographic type *leaner* in most cases lemmatizes as the comparative form adjective. However, there is one line, 15, in which it is clearly being used as a noun.

15. I became a wearer of double-breasted suits, a leaner on bars, a discusser of interest rates.

It would be a lexicographical error to treat this single occurrence as sufficient evidence for the existence of an English noun: *"leaner: one who leans"*. The correct analysis is that this is an agent nominalization of the verb *lean*: an ad-hoc product of a grammatical rule, not an item in the conventional vocabulary. The same goes for *discusser* and *wearer* in the same citation. This guideline is, however, only a guideline. The distinction between products of a grammatical rule and lexically distinct conventional vocabulary items is necessarily fuzzy. Indeed, it is not only fuzzy, but also constantly moving. Yesterday’s exploitation may be tomorrow’s norm. Different lexicographers will draw the line in different places, but it is important that all members of a lexicographic should observe team discipline and try to draw it in roughly the same place. With this in mind, it is surprising to see that *“leaker: someone who lets people know secret information”* has been added to the third edition of Cobuild, a dictionary that used to pride itself on the judicious selection of entry words.
The Shorter Oxford English Dictionary (NSOED) is much larger than Cobuild. It contains a sense of *lean, noun*: "the act or condition of leaning; inclination". Leaving aside the rather curious use of the word ‘inclination’ in the definiens, the question arises whether this sense belongs in a general dictionary at all. On the one hand, it is entirely predictable according to the rules of English word formation: any action verb can be used as a noun. (A piece of anecdotal evidence: the question "Did you have a good read last night?" prompted the answer, "No, but I had a good write." ) If leaning is an action, then a noun *lean* is inevitable. But on the other hand, the very fact that it is predictable constitutes an argument against including it in a general dictionary, in the absence of further evidence. If found, it may be regarded as an exploitation of the verb, rather than as part of the lexicon in its own right. Its inclusion in NSOED is justified by four examples in OED:

1776 G. Semple, *Building in Water*. Pressure from either side would give them a lean to the other side. 1850 P. Cunningham, *Handbk Lond*. Leaden coffins piled thirty feet high, and all on the lean from their own weight. 1851 *Jrnl R. Agric. Soc*. The corn has a decided lean in one direction. 1890 Clark Russell *Ocean Trag*. The rounds of her canvas whitened into marble hardness with the yearn and lean of the distended cloths.

This evidence may be considered conclusive, which is no doubt why so many general dictionaries have followed OED in including it. But before joining the throng, a lexicographer working on a general dictionary of modern English might think twice. There is something slightly odd about at least two of the four OED citations. For the compiler of a modern lexicon of English, the real question is whether *lean* is so well attested as a noun in modern English texts that it must be recorded as part of the vocabulary of general, conventional English. It does not occur in our 500-line sample. But we must also beware of the "failure-to-find" fallacy. The fact that we have failed to find something in a sample does not mean that it does not exist. So to resolve our dilemma we must widen the search. Widening the search takes two forms: if we want to make a decision with confidence, first we must forage beyond our sample and skim through all 2164 corpus lines for *lean* and all 411 for *leans*. Secondly, we must be prepared to consider traditionally collected citation evidence, which tends to focus on rare and unusual uses.

For the corpus-based *New Oxford Dictionary of English* (NODE), the clincher that decided the lexicographer to include this sense was a citation from *Know Your Land Rover*, by Robert Ivins (1991):
16. The vehicle has a definite lean to the left.

The conclusion of this detailed discussion of the lexicographical dilemma whether to include borderline senses such as *lean, noun*, "the act or condition of leaning" is that it really *is* a borderline case. Lexicographers compiling a modern dictionary should protect themselves against a natural predisposition to be overwhelmed by OED's sometimes overwhelming historical evidence, and think. Modern corpora may provide enough evidence to offset OED's collection of old citations, illuminating though those may be for purposes of understanding the history and development of a word. The wise lexicographer will resist the temptation to include such cases unless overwhelmed by evidence of present-day usage. ("Overwhelming evidence", be it noted, may consist of no more than a handful of textually well-formed and convincing modern uses.)

The problem of sparse data is a constant one in lexicography, and calls for policy decisions which, in practice, are rarely discussed explicitly. Is the policy of the dictionary to record all imaginable words? That seems like a recipe for disaster, and yet it seems to be the policy attempted by some dictionaries, for example in their choice of undefined run-on entries. Or should it be all words ever found? But, as we have already seen, large corpora contain many, many examples of creative coinages that can hardly be regarded as an established part of the language. All words used by two or more independent authors? All words in conventional usage? If the latter, by what definition of conventional? There is no single correct answer to these questions, but it behoves dictionary makers to think about them and to make policies explicit.

A final remark before we leave "the act or condition of leaning; inclination". Earlier, I questioned the inclusion of "inclination" as part of SOED's definition. The normal meaning of *inclination* in modern English has nothing to do with *leaning*. Having written a satisfactory explanation or definition, the lexicographer should leave it at that, resisting the temptation to have another go at it – though in present-day English dictionaries this guideline is more often breached than observed. Anna Wierzbicka, for one, has argued strenuously (e.g. 1993: p.48) in favour of saying what you have to say and then shutting up:

There is no need to add anything; the simple short definition is okay; on the contrary, it is the longer one that is faulty, because, as Aristotle pointed out 25 centuries ago, in a definition every superfluous word is a serious transgression.

*Names in corpora*

The base form *Lean* occurs not only as a word but also as a name. Human beings recognize names in text with little difficulty, and for this and other
reasons, many dictionaries do not include names at all. Others offer a selection of names of famous people and places, with short identifying descriptions, as part of the cultural heritage of a language, or (in the case of bilingual dictionaries) in order to give translation forms of a name where the forms are different in the two languages. No dictionaries attempt to list all the conventional surnames and place names known to a language, let alone business names and product names. The dictionary would be completely swamped with useless information – useless, that is, for human users.

However, for a computer program processing text, it is a nontrivial task to distinguish proper names from vocabulary words in text, and to know whether a name is a place name, a personal name, or a business name. And for applications such as speech generation and speech recognition, a lookup table of phonemic transcriptions of names can be invaluable. Some lexical database developers now report that over 80% of the items in their databases are proper names, and the percentage is growing. Because names are more or less conventional but do not have a meaning like other words, there can be no more dramatic illustration of the need for lexicography in the age of the Internet to be application-driven. The decision whether or not to collect and store names is determined by the application. And if names are collected, what is said about them is likewise determined by the needs of the application.

In printed texts proper names are usually marked by an initial capital letter, However, this distinction is not reliably observed on the Internet or in email correspondence and is of course completely absent in speech. Corpus-based lexicography designed to support text processing now recognizes named entity recognition as an important part of the task of text analysis. When all the place names and all the personal names in the world have been added to the computational lexicon (always allowing for the fact that new ones are generated from time to time), lexical inventorists will need to turn their attention to names of business corporations, brands, and products. Names of brands and products are a particularly rich source of new vocabulary.

At this point, common sense intervenes and says, “You don’t need to know very much about a name. Wouldn’t it be better to generate the relevant information as part of text processing, rather than storing it in a database?” To do this successfully in a powerful general-purpose natural-language processing system, we would need at least to have procedures for determining that a term is a name of a certain kind, fulfilling a particular semantic role. These procedures would determine information of the following kinds:
a) Is it a place name, a personal name, or a business name?
b) If it's a place name, where is the place, and what class of place is it (e.g. country, city, mountain, river, street)?
c) If it's a personal name, does it contain information about the ethnicity and/or gender of the bearer? Is it the name of a famous person, and if so what is he or she famous for? If we encounter the expression David Lean in text, how do we know that it denotes the film director and not some other citizen?
d) If it's a business name, is it the name of a firm, or a brand or product?
e) If it's the name of a brand or product, what's it for? (E.g., can you eat it, or do you use it to cut holes in metal?)

These are the main broad categories about names, but of course there are a myriad other subcategories of information about names that can be of great importance, constituting a database of conventional shared knowledge. For example, if the answer under (c) is “He is famous for making films,” we might add supplementary questions such as “What films did he make?” Thus, the lexicon that includes names becomes not merely a list of vocabulary words, but also a collective cultural index.

As it happens, only three of the uses of Lean in our sample are proper names. To understand one of them, it is necessary to know that David Lean is a film director. This information may be available from a standard encyclopedic dictionary, which would be less likely to include anything about Lin Lean Lim or Gangsta Lean. In computational text processing, the identification of such strings as proper names in general texts is likely to be dependent on coercion from the local context for many years to come. On the other hand, Gangsta Lean is already stored in many Rappers' databases, available through on-line resources. (At the time of writing, Gangsta Lean gets 4570 hits on Google.)

To summarize, therefore: it seems desirable to use the enormous power of computers to coerce named entity recognition dynamically, but also to use names as pegs on which to hang attributes and values for as many names as possible in a hierarchically organized knowledge representation system. Here the role of the lexicographer merges into that of a traffic manager.

Case Study 2: How to tell a fish tank from a battle tank

The title of this case study is a tribute to Mike Lesk's influential paper (1986), 'How to tell a pine cone from an ice-cream cone'. The lemma lean has very few noun uses, so to study the mapping of the meanings of nouns onto their uses, we must look elsewhere. The word chosen for illustrative analysis is tank (noun senses only).
The first point to notice is that quite different techniques are required for the analysis of noun senses and verb senses. Verbs and adjectives, on the whole, respond well to syntagmatic analysis, and this is also true of nominalized forms of verbs such as *distribution*, which has the same argument structure, mutatis mutandis, as the verb *distribute.*

*Tank*, however, is not a nominalization. It is one of many referring expressions, which refer to physical objects in the world around us. In fact, it refers to at least two classes of material object: 1) a type of military armoured vehicle; 2) any of various types of storage containers for liquid or gas. It is obvious enough to a human that an armoured vehicle is quite different from a storage container. The task for the computational lexicographer is to provide explicit clues to help an NLP (natural language processing) program make the same distinction, and to compute the implications accordingly.

The relevant clues for noun meanings are different in kind from those that indicate the relevant sense of the verb and adjective. Some of the clues (in italics and underlined in Appendix II) are indeed in a syntagmatic relationship with the target word, for example verbs of motion such as *rolled* and *rumbled.* Others (in italics but not underlined) are in a freer relation with the target word: they are what J. R. Firth called “collocations” as opposed to “colligations”.

Typical colligations of the military vehicle sense of *tank* include:
- subject of a verb of motion: e.g. roll, rumble, break through, drive over
- subject of a privative verb of motion: e.g. be stuck, abandon
- subject or object of a military destruction verb: e.g. fire at on, strike at, destroy, bombard
- modification by a cardinal number (military analysts like to count the number of items in a force available to a commander)
- modification by a term denoting a nation or other military power (e.g. the U.N.)
- modification by one of a specific set of brand names, e.g. *Leopard, M1, Merkava, T-72*
- modifying terms denoting military units such as *squadron, division, force.*

Typical free collocations of *tank* in this sense include:
- infantry combat vehicles, armoured fighting vehicles, armoured personnel carriers, warplanes, planes, ships, aircraft carriers, submarines
- artillery, mortars, weapons, machine guns, ammunition
- bombard, mortar fire, heavy shelling
- army, troops, soldiers, force (noun), general (noun), enemy, rebels
- battle, maneuvers, manoeuvres
• camouflage
• headquarters

Terms that are typically colligations may also occur in the co-text as free collocations: for example, someone else, not the tank, may be doing the firing.

Expressions such as these provide clues (not certainties) that the kind of tank in question is a military vehicle, not a habitat for fish or a storage container for fuel. The effect is cumulative and probabilistic, not digital and deterministic. It is perfectly possible to strike at and destroy two hundred heavily camouflaged fish tanks with artillery fire. But less likely. With each additional clue, the likelihood of a particular interpretation increases. And when ordinary people read ordinary texts in an ordinary, linear fashion, expectations are pre-set from the outset. Indeed, they are set by the domain before the reader reads even the first word of a text: the very act of picking up a sports magazine engenders certain expectations about what kind of things will be said in it.

Texts containing tank in its military sense are particularly rich in both collocational and colligational clues. By contrast, clues for the container sense of tank are rather sparse. The colligational clues consist mostly of a pronominal modifier:
• Fuel tank, water tank, storage tank, ballast tank, oil tank, gas tank, etc.

The density of collocational and colligational clues seems to correlate with the specificity of meaning. The military vehicle meaning is more specific than the container sense.

Free collocational clues in the latter sense include names of species of fish, in particular exotic aquarium fish. This brings us to another question: is a fish tank the same sense of tank as fuel tank? At a very general level, both are containers, but they are used in very different semantic frames. There is no single “right” way of dividing up the different uses of this term. If we separate fuel tanks from fish tanks, should we go on to distinguish the small fuel tanks on a vehicle from the vast storage tanks in which fuel is stored in bulk? Does the term fuel tank have two senses? The answer to this and similar questions will be determined by the application (the kind of dictionary being compiled, its intended use, the space available), not by anything in the language. Finding the right level of generalization is an essential part of the lexicographic task.

A term with a very broad meaning, such as tank, has a wide variety of uses in different contexts, and lends itself to the coinage of multiword expressions, both as modifier and head, some of the resultant terms having specialist meanings:
• Towing tank, breeding tank, single species tank, flotation tank, greenhouse tank, septic tank, think tank
• Tank engine, tank car, tank barge, tank cutter

Again, the decision about which of these to define in a dictionary will be determined by the application. Traditionally, lexicographers have, quite rightly, favoured entering multiword expressions whose meaning cannot be easily deduced from the parts. Thus, a breeding tank is a tank for breeding something in (as it happens, fish); a flotation tank is a tank for floating something in; a greenhouse tank is a tank in a greenhouse. These terms can, perhaps, be safely left out of the dictionary. The term fuel tank is one of the most common N+N combinations, but entering it in a dictionary would be silly. Who would benefit? It would entail disregarding the most basic type of modifier-head semantic relation, while opening the floodgates for an unlimited number of other noun phrases in which the term denoting the thing contained modifies a head noun denoting the type of container. Frequency is not a good criterion for dictionary entry of N+N combinations. On the other hand, it is a very good criterion for indicating typical syntagmatic behaviour of compounding nouns, especially if accompanied by a specification of the type of semantic relation(s) involved. Dictionaries might at least indicate typical modifier+head relations for certain noun entries such as tank. Among the monolingual English dictionaries, only Collins English Dictionary (CED) attempts to do this, and it is constrained by the fact that it is not corpus-based (despite occasional imaginative claims to the contrary by its publisher).

On the other hand, terms such as septic tank, think tank, tank engine, and tank top are opaque without some kind of explanation, and are therefore good candidates for dictionary entry.

It is a slow and time-consuming business, wading through corpus evidence looking for significant evidence, especially as, ideally, the whole corpus will be searched, not just a small random sample. A lexicographical shortcut for corpus analysis, particularly useful for the analysis of nouns of this kind, is the statistical test (mutual information) described in Church and Hanks (1989, 1990). Table 1 shows the words most associated with all occurrences (10,547 – far more than even the most dedicated lexicographer could read through ) of the term tank(s) in the entire AP corpus for 1992-96. Mutual information (MI) selects associations on the basis of statistical significance, not mere frequency. The lexicographer can use a table like this to begin the task of sorting out the different semantic frames and senses.
Table 1: MI scores for words most associated with 'tank(s)' in AP 1992-96.

<table>
<thead>
<tr>
<th>Associated word x</th>
<th>MI score of ( tank(s) + x )</th>
<th>Co-occurrences of ( tank(s) + x )</th>
<th>Total occurrences of ( x ) in the corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>30,000-gallon</td>
<td>10.75</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>t-55</td>
<td>10.67</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>t-82</td>
<td>10.64</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>t-72</td>
<td>10.64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>M1-a1</td>
<td>10.64</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>67-ton</td>
<td>10.64</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>400-gallon</td>
<td>10.64</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>1,500-gallon</td>
<td>10.64</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Merkava</td>
<td>10.58</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>side-mounted</td>
<td>10.54</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>t-80</td>
<td>10.53</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>sidesaddle</td>
<td>10.43</td>
<td>61</td>
<td>53</td>
</tr>
<tr>
<td>t-54</td>
<td>10.37</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>20-gallon</td>
<td>10.22</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>pressurizing</td>
<td>10.12</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>M1-a2</td>
<td>10.11</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>septic</td>
<td>10.03</td>
<td>106</td>
<td>68</td>
</tr>
<tr>
<td>70-ton</td>
<td>9.98</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>m-1</td>
<td>9.83</td>
<td>107</td>
<td>61</td>
</tr>
<tr>
<td>761st</td>
<td>9.36</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td>Pressurize</td>
<td>9.28</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>m-60</td>
<td>9.21</td>
<td>51</td>
<td>19</td>
</tr>
<tr>
<td>Briarcliff</td>
<td>9.05</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>million-gallon</td>
<td>8.90</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>propane</td>
<td>8.89</td>
<td>386</td>
<td>115</td>
</tr>
<tr>
<td>overfilled</td>
<td>8.70</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>rocket-launchers</td>
<td>8.64</td>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>Treads</td>
<td>8.40</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td>ballast</td>
<td>8.31</td>
<td>90</td>
<td>18</td>
</tr>
<tr>
<td>self-propelled</td>
<td>8.26</td>
<td>52</td>
<td>10</td>
</tr>
<tr>
<td>howitzer</td>
<td>8.22</td>
<td>203</td>
<td>38</td>
</tr>
<tr>
<td>howitzers</td>
<td>8.18</td>
<td>296</td>
<td>54</td>
</tr>
<tr>
<td>turret</td>
<td>8.15</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>transporters</td>
<td>8.13</td>
<td>34</td>
<td>6</td>
</tr>
<tr>
<td>155mm</td>
<td>7.98</td>
<td>151</td>
<td>24</td>
</tr>
<tr>
<td>punctured</td>
<td>7.96</td>
<td>250</td>
<td>39</td>
</tr>
<tr>
<td>turrets</td>
<td>7.95</td>
<td>45</td>
<td>7</td>
</tr>
<tr>
<td>ruptured</td>
<td>7.76</td>
<td>552</td>
<td>75</td>
</tr>
<tr>
<td>armored</td>
<td>7.75</td>
<td>4016</td>
<td>544</td>
</tr>
<tr>
<td>Leopard</td>
<td>7.65</td>
<td>166</td>
<td>21</td>
</tr>
<tr>
<td>rumbled</td>
<td>7.52</td>
<td>383</td>
<td>44</td>
</tr>
</tbody>
</table>

[Footnote: Thanks are due to Richard Sproat of AT&T Bell Laboratories for help with the table.]
Other statistical tests are in use in lexicography, in particular t-score. T-score favours high-frequency function words, so is especially useful to EFL lexicographers in compiling dictionaries showing the natural colligational phraseology associated with a target word. MI favours the less common content words, and so is particularly useful in identifying free collocations.

Conclusion

This paper has presented the main points of corpus analysis for lexicographical purposes of content words: verbs, adjectives, and nouns. It shows how the corpus analyst can take the syntagmatic structure of a word and relate it to the meaning or semantic structure.

It suggests that several aspects of traditional lexicography, including pragmatics, function words, -ing forms, treatment of names, multi-word expressions, and sense distinctions, can benefit from recent developments in cognitive linguistics, corpus linguistics, and frame semantics. It offers a reminder that lexicographers have to make choices and suggests that the better informed they are in these areas, the better the choices they will make.

It proposes that for verbs and adjectives at least, syntagmatics, rather than (or rather, in tandem with) perceived meaning, should be the organizing principle of the dictionary entry. A similar, but slightly more complicated case, is made for the organization of nouns: different kinds of nouns require different kinds of treatment. For referring expressions in particular, any syntagmatic arrangement needs to be supported by classification of the statistically significant collocations found in the immediate environment of the word being analysed.

Syntagmatic relations in texts are observable, measurable, and testable. Perceived meaning is private and depends on the private beliefs of the perceiver. If perceived meaning is the organizing principle of a dictionary, examples can often be found to illustrate the perceptions of the writer, but that does not mean that the writer has achieved an appropriate level of generalization.

Syntagmatic arrangement does not obviate the need for lexicographic art and judgement in the arrangement of entries. Decisions still have to be made about how to group uses, how to word definitions, and what to leave out. But syntagmatics requires working with palpable material, which can be measured and objectively evaluated, as distinct from the unsubstantiated assertions of introspection and ad-hoc reactions to isolated citations.
References


Lesk, Michael (1986): ‘Automatic Sense Disambiguation: how to tell a Pine Cone from an Ice-Cream Cone’ in Proceedings of the SIGDOC’86 Conference, ACM.


APPENDIX I: Corpus sample: lean*

NAME

other good purchase. One of David Lean 's timeless epics, it speaks of every(\Warner Bros. 20. " Gangsta Lean " DRS(Capitol " Platinum)< MODE
document made public Monday. Lin Lean Lim, the paper 's author, said

VERB 1, S P A: [(PERSON) lean [ADJUNCT[ Position | Direction]]

ealized. Over the sea bass, Angus leaned across the white linen tablec
e rickshaws moved together and he leaned across and punched his brothe
was nothing lazy about the way he leaned across the table towards her,
thing, could n't it " Dangerfield leaned across and whispered, with th
 go away. Jessica, seat belt off, leaned across Rory 's chest, drove h
 e made a move to open her door he leant across and slammed it shut. Wh
ght to my heart; I should like to lean against that tree between the o
 t I have made it " He had come to lean against the foot of the baniste
ld man on the s-bahn. The manager leaned against the wall, arms folded
ted his collar philosophically, leaned against the hardier Theatre b
axactly a week from now. As and he leaned against the railings he sudde
r word. Trembling with shock, she leaned against a sundial and re-ran
 loose-box and, coming out again, leaned against the closed door with
er he took a very deep breath and leaned against a wall. He had been o
He put his arm around her. She leaned against him comfortably, and he leaned against the arm of Morris's armchair, and lit his pipe. He leaned back in his seat, smiling ironic, holding their drinks in mid-air. She held her breath. Halfway down she stopped, leaning against the arm of Morris's armchair, and lit his pipe. He leaned back puffing at it wondering if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She as thinking " Susan broke off and leaned back against the seat, her hair armchair, and lit his pipe. He leaned back against the wall and sat back. You need to sheet out control as you if passing seconds. Making herself road steppe of her Slavic soul. I held their drinks in mid-air. She
d fiercely. Then, standing up and elf over the canoe's side; first executed quickly. Do not hold the airs, and I gave her a cigarette. re was a man in the driving seat, . There " You see " It 's awful " " Of course " he agreed smoothly, r the last five years " he asked, trong hands behind his dark head, 's slimmed down and laid back. He ame is played. As they answer she he guard stands by the door. Bill luggage. She folded her arms and rom the paperwork on his desk and e " Ain't fair, eh " Aldo Isidori , put his fist to his mouth, then Hal. I had his chair, too, but I w hours into Atlanta from Madrid, was behind it " my kids " he said, the number and shape of debates. the horn of the saddle, Joel Berg day to night " Guy Sterne came to sler, who wouldn't say how he was e music and eyed a handsome youth distance, shows Charles and Diana . When it stopped, Delaney had to Russians " I nodded. The bald man drove her from Scotland " Agrippa angering her mind, she turned back, to St Paul 's and found Ralemberg o great effect in England, was to what happened, but they seemed to ain. But why must you go out " He ave we here " their father asked, he got a steel knee on top of it. arder at controlling himself. She Brown Simpson 's mother, Juditha, er octaves. Graffman, twisted and three sides in a broad loop. She ecrcy and privacy, was forced to nd complains of a slight ache you Raise your knee to chest height. Lean forward and place your opposite began Henry. Yes " The policeman leaned forward. It does seem almost d a cracked skull and " The Woman lea deceitfulness. Miss Thorne " he ed in a soft, concerned voice. He r before. What 's up mate " Billy leaned forward over the steering whe leaned forward further and punched h eard of a surprise " The pale golden head lea formation in very particularly. He " The voice paused again. Manescu leaned forward, tried to get some im leaned forward towards the grey and e old man was smiling. Tuan Ti Fo leaned forward and took her hands in and then, drawn inexorably, she leaned forward across the table and on of yours " Indeed I was " Vass leaned forward in his seat, reached
e, and as he got into her car she leaned forward and kissed his cheek, what is the matter " My master leaned forward and in curt, clear to the beginning of the book " She leaned forward, hands clenched on her knees, I see what you mean " Lindsey leaned forward to lift the cup to her at home " He frowned slightly and leaned forward, her fingers gently p
es, I what is the matter " My master leaned forward to take her hand. No one leaned forward. Then when I saw you relaxed and where a wind-blown standard leaned forward again, his weight drifted out to the field. Wait " Perdita leaned forward and stared down in di
es, I see what you mean " Lindsey leaned forward a little, a cigarette leaned forward and judiciously stub
y, what is the matter " My master leaned forward, hands clenched on her d the table. Every time one of us leaned forward for the salt the flam
es, I see what you mean " Lindsey leaned forward, her fingers gently p
y " Her hands were shaking as she leaned forward and took her hand. An
es, I see what you mean " Lindsey leaned forward and took her hand. An
, and as he got into her car she leaned forward and kissed his cheek, what is the matter " My master leaned forward and in curt, clear to the beginning of the book " She leaned forward, hands clenched on her knees, I see what you mean " Lindsey leaned forward to lift the cup to her at home " He frowned slightly and leaned forward, her fingers gently p
door opened and a guy got out and n a while " Francis said. The man d the castle 's steep rampart. We
He did n't say anything else. He do we get up there " Beth asked, ddling up to the Liberals is like
I hope " you do n't mind a man " turned to advantage " Lachlan was rned to come back for me. I stood
striped breeches and greasy caps y round the piazza, the Brigadier hey are placed, either hung up or e. Alina turned to Liston. He was e a punnet. Loyonel, he 's there, s hand was suddenly up above her, ee warnings in the last round for ome funny reason my. Well you 're hat record. There was no place to

Mothers ar Claude, 72, of the exhibit. He e my baby boy " one woman sobbed, e stage level, sitting, lying and of approval among fellow fighters aceful people " one old man said, Mr. Simpson and a female that was wly through seaside Mirador Park, s there 's a line of 200. Domingo oks, art and the rolling cart she propel or lead the person. Don't fer the time bein', thanks " Fred ting more naturally to the gusts, bowls of candy. We in the gallery her rebel chiefs. Victoria Garcia on a central street, as residents brochure includes a photo of him ould fly " But police said he was he car came alongside me. Coombes m from behind. It was Tom. He was ) of water that 's almost boiled. ca n't blame her. And, he used to Did you see that " What wee car " ing-tip to his scratched hand, he going to be all right " Catherine our point, have n't we " The man ok here " At this point Maidstone erked open. Nice " he smiled. She anyone else until our ambassador kage of the bed again. The guests d made as if to turn back. Pascoe f feet at once, and instinctively o. The rest you know " Mandeville study of the dead man. Athelestan ything would be tangled " Corbett ginning to stop sweating when Jan the soldier lay on the grass; he ive up " she said to herself. She " We dumped Carol face down and I think perhaps you 're right " He ypocrite " Harriet says drowsily,
ue and handed it to Denis. Denis, leaning over, half-raised from the c
king up here anywhere " he demanded, leaning over to pick up the towel he
ot restlessly to his feet and was leaning over the green wooden railin
do the stores to get something and leaning over the counter to get some
the head, and handed it to Sally, leaning over so I could get close to
an shakes me. Right " he says. He leans over me, puts his big fingers
ee nearly touched, as the two men leant over the table to hear what th
anage " Thanks " he grinned, then leant over and kissed her on the che
the proximity of her body as she leant over the desk. First we took t
the crossword or something my dad

I built my fence. The one I ' m
leans heavily on its argument that
leaned on his sartorial knowledge wh

VERB 2, S PA: [[Person]] lean {on [[PhysObj]]} (= rely on)
, like it or not, is likely to be leaning fairly heavily on the US for
thing from that document, but it leans heavily on this which is the R
ly player-manager, at Swansea, he leant heavily on former Liverpool cl
tement, which he has hinted would lean heavily on allegations that pol
action. The administration was leaning heavily on its argument that
decided on " Big City Blues " to lean more on her playful, youthful v
or Burton ' s the tailors. I later leaned on his sartorial knowledge wh

VERB 3, S PA: [[Person]] lean {on [[Person | Institution]]}
(= put pressure on)
d you 're very much aware you 're leaning on him. " I could feel the w
fits can provide benefits only by leaning on them to pay taxes. And th
t that is like " Young said. Cash leaned to the country side of rockab
30, a police officer and Democrat leaning to Dole. You don't feel ther
here goes to the Rhino set, which aste in politics is often said to lean more toward international than
certain that the decisions he is leaning toward are iron-clad " But ion was reached, but the scholars democratic rule. Only 10 percent leaned toward dictatorship.
new editor, said many journalists leaned toward accepting the go
8 presidential race. He initially leaned toward conservative Texas Sen
community movement. Those purged leaned toward the Soviet Union and w
her lover 's wife says she was leaning toward a guilty verdict. The e. One legislator who said he was leaning toward supporting the measur
, Raymond Seitz, said Clinton was leaning toward sending a fact-finder delegates, both her own and those leaning toward Charest. She is doing
rches. The news agency has been leaning toward Boris Yeltsin during slators who had been undecided or leaning toward supporting the trade aid " that in Sicily the Mafia is leaning toward Forza Italia and newl
the Energy Department generally leaning toward ending the ban. Depar
said Wednesday that officials are leaning toward the suicide scenario. ly identical measure in Idaho was leaning toward defeat but remained t
oional 21 senators were opposed or leaning toward voting no, and 32 wer
partment sources Wednesday to be leaning toward allowing more ships, ave to enforce it. The states are leaning toward requiring that oyster Boris Yeltsin has showed signs of leaning toward increased state contr er runway. The plane came to rest leaning toward its left wing and wit
lyy 11 accident, investigators are leaning toward blaming a fire starte said they would vote for, or were leaning toward, Clinton to 35 percen ler, has 50 years on Tsai, and is leaning toward Clinton " though I wi
o lives in Newton, Mass. , and is leaning toward the Democratic ticket voters who had been undecided or leaning toward Clinton. But they sai
ht 's collision, and the NTSB was leaning toward human error, Black sa
etwork support. The company won't lean towards Kode Computers ' tradit igence contacts with Iran; Israel leaned towards favouring Iran in the
course. The resultant design has leaned towards a course more heavily anage. I guess I do have a slight leaning towards comedy, but I find i

thing, put a VCR on the roof, and leaned a ladder against the outside
ledges for books " says Clodagh. Leaning a book with a lovely cover a
ed. She did n't wait an instant. Leaning her bike against the gate, s
t eased himself into an armchair, leaning his cane against the wall. A
You 're right, Dick " Martin now leant his gun gently against the thi
ad " Now think on " Stanley said, leaning the discarded wheel against
k there. Inside, the soldiers lean their M-16 rifles in a corner b
iping the mattresses down and had leant them up against the wall to dr

Lexical set [[BodyPart]] = {head | back | arms | body}
to the bushes in a wild leap; and leaning both hands on the pommel of
over a cushion so that she could lean her damp, aching head against i
Maggie 's colour improved and she leaned her head against Sarah 's sho red to make its ideals a reality. Leaning her head tiredly against the other funeral. He stood aside and leaned his back against the wire pal
ared in Felicity's cheeks. Keach leaned his little forearms on her de
out from the side of a building, leaning his back against the buildin
r was forthcoming. The little boy leaned his face to one side as if he
Leaning his tired body on a cane, Ch
melody. They end near each other, leaning their heads together. The th
VERB 7, SPOA: [[Person]] lean {bulk | weight} [ADJUNCT
[Direction]]
Iser is just being polite when he leans his bulk forward and says "I
hands on the rim of the sink and leaned her weight forward. Glancing
her ankles between his knees and leaned his full weight against her.
VERB 8, SPO: [[Person[Politician[American]]]] lean [[Political
Party]]
[[Political Party]] = [N | Adj]
e exceptions. Historically, women lean Democratic, but Dole cannot who
asha-d vs. Gov. Pife Symington-r. Leaning Democratic. ARKANSAS Gov. Ji
th 262 electoral votes "that now lean his way. Among them, Kohut note
VERB 9, SPO: [[PhysObj]] lean
und shifted underneath it. It now leans about 13 feet off the perpendi
oosen in wet weather and begin to lean. larch-lap produces a universal
an before. He speaks more slowly, leans more eagerly, so as to offer m
UNUSUAL AND JARGONISTIC VERB USES
cott mckenzie is also among those leaning in favor of an independent r
of rated power, therefore one can lean. When climbing at full throttle
[Full text: The Old Wives' Tale "Don't lean below 5,000 feet" resulted
from the fact that above 5,000 feet, full throttle at red-line rpm
gives less than 75 per cent of rated power, therefore one can lean.
When climbing at full throttle, the 2,300 rpm or so of your fixed-pitch
propeller probably delivers less than 75 per cent.]
VERBAL NOUN: leaning
school at which I was CFI taught leaning in Exercise 4 and throughout
VERBAL NOUN: leaning(s)
in "Can you discern an editorial leaning or tendency in the work they
of us regardless of our political leaning" said former Democratic Par
as been quiet about his political leanings and did not give any hints
e a justice not for his political leanings as much so as for his abili
servative and isolationist in its leanings. Asia is a vital u. s. inte
and that it could make money. My leanings had always been towards the
isms "War: Sukarno 's Fascist leanings He grew ever bolder in what
ner party. Indeed, the ideological leanings of readers ' papers had mor
f a broad survey of the political leanings of America ' s religious gro
hat he perceived to the communist leanings of Kennedy and other govern
and blacklisted for his communist leanings, Seeger recently went on a
te People. Buchanan denies racist leanings that some critics have draw
endence were accused of communist leanings, the report said. The unive
ange. I guess I do have a slight leaning towards comedy, but I find i
members of the Committee who had leanings towards a particular style
they didn't know of her political leanings when she was nominated for
es some of his more inaccessible leanings to tell a straight tale wit
affirmative action. Those liberal leanings were balanced, however, by
The Pisa of duomo, baptistry and leaning tower " but to many of its 1 Archaeologists digging under the Leaning Tower of Pisa have found an
The leaning tower of Pisa has stopped it

ADJ 1: lean [[ Human | Animal | Body Part]]
ic gene that correlates with how lean an animal will likely be " said for him again, and his arms felt lean and hard through the sleeves of e sun. To make himself especially lean and bronzed for his meeting wit hirt the Frenchman 's body looked lean and hard, and she guessed he mu Chertro, her fedpol friend of the lean and grinning face. I was please re prowling round each other like lean and hungry panthers waiting for ef cattle], it is growing longer, leaner and taller, especially during Africa " Hopcraft is middle-aged, lean and lanky, with a long, narrow ust have left them all fitter and leaner at the end of the day. The ca
The second man, who was tall and lean, came running from the glade in hard to forget " and more: tall, lean, dark and almost painfully attr p. Buckmaster ran a hand over his lean face. It 's essential at least
He is not much older than her, a lean figure, his jeans patched, not middle-aged man, loose-limbed and lean, got out of the driving seat an nt " He glowered at her, raking a lean, hard hand through his hair. Ye a horse, sometimes fat, sometimes lean, is all the while the same hors a square shape that concealed his lean jaw and fallen cheeks as much a ned, Hollywood star. His build is lean, like Grant 's or Gary Cooper ' body. The sheer size of the dark, lean, muscular male body made her thr Schuylkill River. The Biglins are lean, muscular and athletic, but the to account for the fact that many lean people were smokers or had life ds, Jeffrey Meek has that whippy, lean physique associated with black likely to develop cataracts as the leanest, researchers said after trac Sabrina, her eyes sparkling in a lean tanned face, likes to talk to h t middle-aged women should be far leaner than the standard recommendat rumming. She thought how long and lean they were, a true artist 's han the big-eyed cutesy seals. Or the lean-flanked wild and free horse. Or y-five years old, scruffy, fairly lean, with a young ungainly beard on way for the castle. His face was lean, with a distant, other-worldly

ADJ 2: lean [[ Meat]]
ttles of milk, and half and half. Lean bacon. Mustard. Peanut butter. be slim. MENU BREAKFAST 1 oz(25g) lean ham (all fat removed) served wi rs might include fish, poultry, lean meat and vegetarian dishes, but k fish " both white and oily very lean meat poultry " but without the oyce to prepare a special diet of lean meat, fish, pasta and protein. and pulses, unsweetened cereals, lean meat, poultry and fish. Look fo c knife. He cuts small fillets of lean meat the size of quails ' breas ins, fresh fruits and vegetables, lean meats, fish and poultry, and sk low-fat sausage lovers, there's a Lean Poultry Sampler ($45) from Aide n(no skin) or I ounce lean ham or lean roast beef(remove any fat), and Public Voice. Even if pork may be leaner than some beef products, the e Simmental, and the meat is very lean. There is also some double-musc

ADJ 3: lean [[ Muscle]]
gen balance and limit the loss of lean body mass. Exercise should grad xercise to increase the amount of lean muscle tissue and decrease the by surprise when twelve stone of lean muscle powered into him at snak problem. Exercise helps preserve lean muscle mass, an important benef ugh protein have deterioration of lean tissue, immune response and mus
at while the U. S. forces will be
leaner, America will not abandon the
economy by keeping inventories
smaller paychecks and create a
services to taxpayers and create a
third, a new kind of government, a
and production systems
will remain at their currently
rming Russia's large army into a
1993, the Big Three have become
ill and make the benefits package
described the White House as a "
that because of the move towards
iness efficiency these days means
peration in Europe designed for "
to more profitable retail sales.
and his ministry can survive as a
health subcommittee guarantees a
post-cold War era. She foresees a
leader United Nations concentrating

ced. Businesses are, as they say, lean and competitive again
ays are in keeping with Ibanez's
most effective bloodsuckers are
ld have to take steps to become "
infal tradeoff " Sinai said. The
s. The only way IBM can be really
ploy to pay for the $350 million. This
we must be careful not careless, lean not mean " We can't avoid our r

ative Office " he said. In these
ng work from them and there was a
le in the late 1980s and ear-
el euphoria. Richard was having a
lean time. He was in a state of perp-
dealers. ANALYSIS Hooray for the
lean times " Our new world industry
And did they, they must have had a
lean year for some charities that ha
how did it manage to survive the
year since 1818, except for the
takes blind faith to survive the
alls the tourist boom after three lean years a " sort of miracle " She

nd glowing caverns; did it have a
le and hungry look, deprived, excl-
ings like " Wasting My Hate " are
lean, aggressive and spitting. Netfi-
trim at just above the knee, or lean and layered for longer lengths.
the couture silhouette is longer, leaner and more feminine than ever,
een too large and expensive for a
m a thinner list of
earer of double-breasted suits, a
earer on bars, a discusser of inter
NOUN: leanness

20 kg/m. The association between leanness and fertility has been exte
it earlier and began to develop a leanness and the agility to cope wit

NOUN: lean

s no fat and the wife who eats no lean. For a man so rooted in his dom
nd meat for potting: Take as much lean of boiled ham as you please, an
ght be followed by seven years of lean. To move in such landscape was

NOUN: lean-to
d just outside the window under a lean-to, beside a stack of peat. Tha
frequently found is the barn with lean-tos built against it to provide
El Salvador's hills, in tents or lean-tos, eating little more than co
ed body out of a manure pile in a lean-to near the calving shed. Last
60 mph. Four campers rescued from lean-to on Appalachian Trail. RHODE
hot proved to be too close to the lean-to roof to give a habitable roo
allow ventilation at the apex of lean-to roofs the new Abutment Venti
nessman who had recently added a lean-to scullery and installed water

APPENDIX 2: Corpus sample: tank | tanks

NAME
n City school superintendent Don Tank said Foster would not have been hi
NOUN: [[Armored Combat Vehicle]]

ounded by Israeli soldiers and
strength. Saddam lost some 3,000
craft, 645 Leopard-series Combat
osnla, gives orders to strike at
 triumphanty into their city on a
ussia has already scrapped 1,141
S. contribution includes 43 M-60
outains. The Syrians used their
e deployment of crack troops and
, is the largest manufacturer of
or rifle division supported by
s of fifty-seven of the precious
la which policed the zone, used
China develop its advanced 780
s, the general would pass in his
, slightly radioactive dust from
ugh the sale of shoddy fuses for
evo was bombed by at least 68
vehicles, a cross between a light
own, but the Serbs renewed their
, he stumbles onto a camouflaged
rael pounded south Lebanon with
o buy the parts for Chinese-made
Israeli force led by three dozen
de Iraq. They want to move their
or so to destroy several hundred
de recently to upgrade its M-48
compound containing Soviet-built
order dispute, Saddam sent Iraqi
t the Serbs, who have warplanes,
sun Ilbo newspaper said about 10
tanks, 1,860 armored fighting vehicles a
tanks, 1,043 infantry combat vehicles, 1
tanks, 1408 " Heavy shelling into city,
tank 50 years ago. He described the livi
tanks, 608 armored vehicles and 2,709 ar
tanks, 80 armored personnel carriers, 84
tanks against the leftists without compu
tanks against civilian pro-democracy pro
ammunition and medium caliber ammu
and artillery had been halted by 6
tanks and ten of the even more precious
tanks and heavy machine guns to force th
tank and the M9 and M11 missiles, which
tank and throw cartons of cigarettes to
tank and aircraft cannon rounds made of
tank and artillery shells. Six other peo
tank and mortar rounds for 45 minutes Su
tank and an armored personnel carrier. A
tank and artillery fire after the first
tank and overhears Bosnian Serb national
tank and artillery fire today, the third
tanks, and Thailand agreed, Kraprayoon t
tanks and three armored personnel carri
tanks and self-propelled artillery deep
tanks and weaken the Republican Guard fo
tanks and supplement them with the Germa
and artillery, which Aidid had agr
tanks and soldiers storming into Kuwait.
tanks and artillery. Using U.S. military
tanks and anti-aircraft guns were positi
ment forces destroyed three Serb service most of the 2,500 battle\textit{dent state. Although many of the key Khmer Rouge headquarters. Two .S. combat troops, more than 100 of 300 soldiers, backed by six Ira a new government. Troops with an-guided munition for targeting \textit{infantry, they lack the heavy have about 60,000 men, plus 700 roops prepared their \textit{M1 Abrams troyed or \textit{Captured 30 opposition round, carry large cargo such as \textit{Iraq-Turkey border said Turkish April 16, 1945: Some 6,800 Soviet ivanko. Serb forces, backed by \textit{Pentagon reported a shifting of roops, backed by more than 100 s of thousands of planes, ships, d with Kuwaiti soldiers and U.S. other heavy weapons. Several m view is that only by rolling h tanks due four months ago. The \textit{Resisting " Hundreds of \textit{Fragg}N. hardware in. The \textit{Leopard} 163 h as those designed to penetrate \textit{ur} surfew Thursday, disappeared. Mussolini and Hitler \textit{"aircraft, roops fought off the rebels with \textit{''s call for attacks only against tance as roops backed by battle. Fang \textit{'s legs were \textit{Crushed by a of Serbian shells, a lone, aged that would allow him to blow the on \textit{a southern coast Friday, but SLA responded with artillery and \textit{Libyan army has 28 brigades, 42 \textit{eastern Dorsale area. Ferociousack Monday by Serb artillery and and Electric officials said the small force of 300 soldiers and source said the army also wants helped the Chinese improve their nd were raked with automatic and ate Friday in a battle involving king out of our window we saw US just two days before \textit{Soviet Army al guilt " said Karl Hoffmann, an anse company that also produced ear and Israeli troops backed by priests began work after \textit{Soviet through the armour of an \textit{enemy brother, Yoram, a corporal in a to reduce the readiness of three officials said Saddam ordered three uns rattling in my right ear and tanks and killed 50 Serb soldiers, but i tanks and the 250 aircraft that survived tanks and \textit{field guns are obsolete}, they tanks and an \textit{armored personnel carrier w tanks and an \textit{aircraft carrier} to the Eas tanks and artillery destroyed several re tanks and two helicopters, surrounding t tanks and \textit{armored cars} battled \textit{protester tanks and} \textit{Bradley fighting vehicles for tanks and \textit{downed} four helicopters and on tanks and Patriot \textit{missile batteries and} tanks and special forces headed toward t tanks and 1.6 million soldiers open a ma tanks and \textit{mortar and artillery positions} tanks and \textit{artillery} by \textit{Saddam \textit{''s army}, B tanks and \textit{armored vehicles}, stormed the tanks and \textit{trucks}. They planted the victo tanks and \textit{armored vehicles} already in Ku tanks and \textit{armored personnel carriers} \textit{run tanks and marching soldiers} \textit{forward} into tank appears to be modeled on U.S. armor tanks are stuck in Croatia, he said. Bjo tanks are inside the city. The army is \textit{tanks, armed} with 105mm guns and thermal tank \textit{armored}, which carry their own sensor tanks, \textit{armored vehicles} and three helicco tanks, \textit{armored personnel carriers} and mi tanks, \textit{armoured vehicles}, small arms and tanks, \textit{artillery and helicopter gun ships} tanks, \textit{artillery and mortars}, to include tanks, \textit{artillery and warplanes}, moved to as he pushed a fallen friend out of tank \textit{awaits} its daily orders. This is th away, but he held his fire " and a tank \textit{barrages} and mock air raids kept te tank \textit{barraged} on suspected guerrilla hid tank \textit{battalions}, 48 \textit{mechanized infantry} tank \textit{battles rage} on the southern \textit{Soviet} tank, Bosnian radio reported. It said v tank \textit{bowled over} power lines, knocking o tank \textit{broke through} rebel lines and ente tank, \textit{cannon fire} by the \textit{Israelis} and \textit{cannons}, \textit{machine guns, mortars and tank \textit{coming over the Bridge} of the Amer tank \textit{crashed through} the wooden pallisa tank \textit{crewmans} in the \textit{Panzer Lehr Divisio tank \textit{Decoys} during the \textit{Persian Gulf War} tank \textit{deployed} in the \textit{West Bank and Gaz tank \textit{destroyed} the \textit{Prague Spring reform} tank, destroying equipment inside, \textit{Kill tank \textit{Division}, was traveling in a milita tank division} that would be asked to fi tank divisions} with more than 30,000 sol tank \textit{driving over} my head " Bole had le
U.S. jets "bombed" abandoned Iraqi tanks during maneuvers on the fourth ann.
tanks facing the White House hit the front wall.
tanks firing in some suburbs, and
fire and two Cobra attack helicopters fire out across the barren field.
tanks fired on the Russian White House.
tanks fired 200 howitzers and tank shells
forces south of Bir Hakeim in an attempt.
tanks advancing west of Kiev captured
forces attacked German units trapped
guns and ships. At home, many Amer
had surrounded Khodzaly, and that
tanks had returned to their bases by Wed
 tanks, holding nuclear weapons and battl
tanks. I don't want to make an example of
in southern Lebanon today, wounding
tanks in the Bosnian army to 100, one:
tank, infantry, commando, artillery, air
tanks into the security zone, said secur
tank is challenged by "troops" at a ro
tanks knocked over a lantern, igniting t
lined up as a backdrop to our cere
tanks lumbered down the streets.
A senior
manned by Russian and Ossetian soldiers,
"memorial" to the men of Exercise Tig
might not be up to a real fight in
tanks. Mikhailov said legal proceedings
 tanks, missile launchers and other heavy
tanks mounted with 105-mm guns and other
near the key suburb of Stup closed
Tanks now patrol the city centre, and i
on the streets. Yeltsin announced
on a main road. An hour later, oth
tanks. On Sunday morning, they reported
opened fire on the parliament building.
tanks put an end to the Prague Spring in
tanks. Railings survive where a hazard w
 REGIMENTS and a Russian motorized di
repulse an intense German assault
roll into Chechnya, 10,000 to 40,0
rolled into Budapest. It is almost
rolled in the next afternoon" fire
rolled through Beijing streets in
rolled into Chechnya in December 1
Tanks found hit government positions on
rumbled through the market at Char
sent by Saudi Arabia. Mujaheddin g
shells and fire in October, is now
shells. There was no word on casual
surrounded by about 10 helmeted sol
surrounded the last remaining separa
tank. This was a potentially decisive a
tanks to put down" protests in March 1991
submarines, as the United State
 track as the aircraft swung towards
Tanks, transport trucks, and jeeps with
tanks Fraps and great banks of boulder cl
 tanks, trucks, artillery, ammunition and
tanks, warships and air power, to rescue
le for refitting and sending them on a March 1991 protest, and army
cross the river, huge T-72 battle
. Three days later, 4,500 Soviet
tanks was apparently never told of the c
tanks were dispatched to restore order.
were being prepared. Their young c
tanks were rolling toward the city cente

Noun: [[Container]]

ic generators and fuel and water
The urine was pumped into a waste
e flames around the 5,000-gallon
emptied the ruptured
ally in a lagoon and in improper
waste that is in 51
a few weeks to establish in the
at the experience of a flotation
ies carrying the fuel and water on the
er other waste water from a storage
7, was pumping liquid waste from
ad is structurally rather like a
black. Both wore face masks and
is no power to dump it into the
so congested he had to consume
, without the need for a bulky
your swimsuit. Some have holding
to empty seawater from ballast
to the sewer system. The rest use
, and electricity from the three
30,000 gallons of solvents in 20
pipe that led from the storage
site's 149 aging single-walled
g maintenance work on nine-story

Noun: [[Container]] (for storage of liquid or gas)

set up in six towns. Two ammonia
and gas build-up in the storage
tanks capable of holding about a million
gas tanks. Excelsior reported. In Progreso,
tank exploded during a morning shift in a
storage tank that exploded Monday afternoon apparent!
tank facility. In Hannibal, in northeast
tank for the third time in eight years,
tanks have to be empty of oil before we
storage tanks. It's a very serious safety conce
tank moved because it wasn't bolted down
tank on fire in Wood County, near the Oh
tanks on July 26 was broken up after com
tank supplying cooking and heating gas t
storage tank that exploded Tuesday morning conta
tank that was about half full. An effort
tank. The deaths came six months after a
storage of oil in the storage
used the explosion in a nitrogen

Noun: [[Container]] (for fuel on a vehicle or plane)

ntary heater, quirky reserve gas
as taken up by a huge extra fuel
he canister is part of the fuel
y finished attaching a new
r was found in the plane's

Mapping Meaning onto Use 195
also unusual in that the two wing fuel tanks and fuselage near the explosive, under the main fuel from the balloon’s propane fuel a section above the center fuel to add five gallons to the fuel d. While some of the center fuel ent over the “sidesaddle” fuel d the distributors open the fuel shafts, engine cases, seats, gas d ran with her before the petrol t gasoline from a punctured fuel g jet fuel: If it’s in the wing sulation around his propane fuel ging tests on the safety of fuel of an airliner, such as the fuel miles on gas in its half-gallon don’t “top off” or fill their n the vehicle’s body and frame. lows fuel to bleed back into the wage, and hookups for extra fuel e has emptied the nation’s fuel vers of ice on the external fuel fisherman carries a rod and gas cars, not to fill their gasoline ct collisions because their fuel yres and the oil, and filled the prevents ice from forming on the recovered 80 percent of the fuel tanks and a more durable thermal insulat tanks are linked, but the carburettor i tank are still missing, all of which cou tank as being a natural cause. Well, w tanks. Ash, 47, was one of several ballo tank design of its older pickup truck mo tanks, estimated to hold up to three mon tanks, exhaust pipes and mufflers. That tanks exploded. The BMW went first, its tank fed the flames. The victim’s parent tanks, finer if it’s in a 55-gallon dru tanks. He said he didn’t insulate his fu tanks in GM pickup trucks made between 1 tanks in the wings.

It is not permitted to drive on or tanks. Mike Overly, a spokesman for the Tanks on later models are located within tank or the petrol pipe is routed too cl tanks that would allow Endeavour to stay tanks. The embargo was used as a means o tank. There was one last-minute technica tank to the boat landing, preparing to s tanks. Travel to Italy is forbidden. Any tanks were unprotected outside the frame tank. When he returned to the parking-lo tank when it’s filled with super-cold f tank, which exploded in the air.

NOUN: [[Container]] (for water - often with fish in it)

e clownfish you list in the same or dying, isolate it in a small remove the fish to an isolation give me any information on the overs drown the father in a fish 2-million-gallon rehabilitation and cichlid pellets. His breeding and duckweed cleared from other our catfish may regard the whole . Check them carefully in low Ph g:F to 85deg:F. I keep my angel ucing the purple moon angel your nfish and put it in a quarantine hose connected to a small water species tank. A single species at I should be able to obtain my ded in very large numbers in any many problems. To set up such a oke down for one of the concrete e of a toilet roll. The aquarium u remove the mops, make sure the "The exhibit " an 18-inch-deep serts I’ve had a 20-gallon marine h keeps the tank very clean. The gen Biolife filter in our office eed neon tetras in his community 's called the feed-and-expansion tank. A pair of either common clowns or tank. A dead snail decomposes very rapid tank and treat it with a proprietary ant tank and feeding requirements of these f tank and set fire to the mother. The cam tank, and backers of the 19-hour trip fr tanks are all full. Jay sells many of hi tanks are rapidly disposed of; and lettu tank as his territory and attack anythin tanks as not all are fully glazed. Examp tanks at 80 6deg:F, as this suits variou tank became overstocked. Did you check f tank, but he died. I exchanged the wras tank, but it was too weak, he said. Marv tank can be taken to mean one of two thi tank, complete with cover glasses, for tank containing large cichlids. The cich tank containing only the larger species tanks containing contaminated water stor tank could have been built from scratch tank cover is replaced as killies are go tank filled with baby sharks 18 inches t tank for almost one year. I’ve two clown tank gets plenty of natural light during tank has been running for a few weeks no tank. His letter shows that fish will sp tank. However, recent changes in the wat
An essential part of all aquatic tanks is the tank's evolution. In the early 1950s, R. A. Fraser, a retired U.S. Navy officer, created the first fish tank in the United States. The tank was a simple, small aquarium with a few fish and some plants. Over time, the tank evolved to include larger tanks, more diverse species, and advanced technology. Today, fish tanks are a popular hobby and are used for both decorative and scientific purposes. They are designed to mimic the natural habitats of the fish they house, with careful attention paid to water quality, temperature, and lighting. Fish tanks can be simple or highly complex, with additional features such as filters, heaters, and pumps to maintain optimal conditions for the fish and their environment.
atogenic Studies, a military think tank. The armed forces are giving the ci
think tank. The study by the Pacific Research
think tank. The party of power doesn't need to
think tanks. The Democratic Leadership Council
think tank, was tried in July and sentenced to
think tank, where lively minds, well versed
think tank with ties to moderate Democrats sai

**towing tank**

**tank engine**
as lived and breathed Thomas the Tank engine ever since he wrote the firs
cluding an ingenious Thomas the Tank Engine. But the real fun is going o
as the narrator of " Thomas the Tank Engine and Friends " an animated ch
oes around the track. Thomas the Tank Engine is another favorite of the p

***
d Tanks No 3, an RSH 0-6-0 side tank, and Lamport, a Bagnall 0-6-0st,

**tank car**
hey were concerned that 14 other tank cars loaded with liquefied propane

**tank barge**
the tugboat Emily S was towing a tank barge, loaded with nearly 35,000 b

**tank cutter**
make medium to large holes. A tank cutter has an adjustable arm which

Noun [[ Garment]]: tank suit, tank top

e, 1910-style navy-white striped tank suit. These turned into clean, ead of the signature outfit of a tank top and short shorts. If the EEOC h
were graphics for swimsuits and tank tops. The provocative handkerchief

UNDECIDABLE ON LOCAL CONTEXT

. Sunday, punching a hole in the tank and sliding along a metal railing. it for the big escalation of the tanks and the gas " Edwards said. Agent
at point. Errs as far as the drop tanks are concerned, does that have any
on the same rail tracks, as the tank comes up over the curve it will til
said it was something to do with tanks but nobody knew Rumours. Yeah. Mh
ile disconnecting a valve in the tank, he was overcome by fumes. With the
for bursting apart, turning the tank into a giant pressure cooker. The S
med at limiting disasters if the tank is given less attention than some o
utes earlier, a large tree and a tank of some type had floated past. I th
overcome by gas after entering a tank on the ship and, other crewmen died
st site. A mesh fence around the tank protected it from flying debris, bu
y invisible glued seams. On some tanks the panels are curved. Another tan
ill occur. In the meantime, gang tanks whose residents misbehave will los