

Why do idioms occur in natural languages?

Erik-Jan van der Linden
Tilia Cordata IT B.V.
ej@tilia.nl

Idioms are constructions in natural language which, from a formal point of view can be described as follows. An expression, *kick*, with a syntactic type, $(np \setminus s)/np$, and meaning, **kick** combined with its argument, *the bucket*, with a syntactic type, np , and meaning, **the bucket** forms an expression *kick the bucket*, with syntactic type $(np \setminus s)$, but with meaning which is not equal to the application of the function to its argument, **kick(the bucket)**. The rudimentary formula which represents the meaning of *kick the bucket* could on one hand be **kick(the bucket)**, but on the other hand it could also be **die**. The latter is problematic for any compositional theory of the relation between form and meaning. There is quite some literature in the field of formal and computational linguistics about this matter, and about the formal representation of idioms (see van der Linden 1993 for references)¹. However, this will not concern us in this paper.

Where linguists have paid little attention to the question *why languages feature idiomatic expressions*, it is a matter considered important by at least one broad-minded logician².

One suggestion in the literature is that the metaphorical nature of idioms is of importance: metaphors are closely tied to the experience of language users, and thus provide easy ways for language users to express themselves. Another suggestion concerns the role of idioms in discourse: idioms more often occur at the end of paragraphs, and thus seem to be used to mark the end of a part of the discourse. Moreover, it has been suggested that idioms are less formal than non-idioms.

The explanation provided in this paper is linked to a conjecture about the evolution of language. Language changes over time are due to the fact that its primary function, communication, requires verbalisation of *new* concepts, objects, etc. about which ideas and thoughts are to be exchanged between speakers. Although the principle of compositionality gives speakers opportuni-

¹It has been argued that idioms are too hard to deal with in our current formal frameworks (Frank Veltman, personal communication).

²Johan van Benthem, personal communication. Expecting extensive discussion on the formal part of it, Johan van Benthem's first question in our first discussion of my research concerned this basic question. The same question can be raised about for instance, negative polarity items, as he explained.

ties to formulate an infinite number of new expressions on the basis of a lexicon and a set of grammar rules or grammatical principles, this does not seem to be sufficient: lexical items are added to the lexicon, and sometimes even grammar rules or grammatical principles are changed. Still, the conjecture here is that although there are many different ways to formulate new concepts, a language user will apply the Gricean maxim of manner, and therefore he will come up with an new expression that is perspicuous, brief, and unambiguous. We'll refer to this as the *principle of perspecuity*. Moreover, it is conjectured here that language users apply a *principle of recycling*, which states that attempts should be made to re-use the means available in language. Adding new words or grammatical principles is not efficient for communication, since it requires additional inferencing on the part of the hearer, and maybe even negotiation between language users whether the new words or rules are permissible.

The following examples illustrate this.

In order to interpret the sentence in (1), the reader should take into account that a rule has been added to English which states that changing the order of the parts of a noun phrase denotes an extremely pleasant event.

(1) Johan celebrated his birthday fiftieth with a where many friends and colleagues were present party.

It is hard to expect that such constructions will become common in natural language.

Formation of new concepts is another example. Rather than using new lexical items like (3) or (5), the first parts of the compounds in (2) and (4), which already existed in natural language, were combined with a technical term that was in use - the second part of the compound, and the expression as a whole was assigned a specific meaning³.

(2) modal logic

(3) dallogomic

(4) dynamic logic

(5) dicolnamicy

Idioms are just when further step: they re-use the vocabulary *and* combinatorial principles of natural language to form new expressions. So, *kick, the* and *bucket* are already part of the vocabulary including syntactic types, and so were the syntactic principles which combine these. The only addition to the lexicon is a specific meaning for this expression.

³Maybe there is a fundamental difference here between logicians and physicists: for every new particle that is discovered in physics a new item is added to the physicist's lexicon, whereas logicians tend to re-use their existing terminology.

One final remark is in order. There is a significant interest in the similarities between natural languages and programming languages by logicians (van Benthem 1989). It is interesting to note that designers of artificial languages (logic, programming languages) hardly make use of idiomatic constructions, but try to stay close to a compositional framework⁴. This way, they implement the principles of perspicuity and recycling. These principles are so strong, that as soon as one single new grammatical principle is added to any artificial language, it is no longer considered the same language (C becomes C++, logic becomes modal logic, etc.).

References

- Van der Linden, Erik-Jan (1993). *A Categorical, Computational Theory of Idioms*. Ph.D. thesis. OTS Dissertation Series, Utrecht.
- Van Benthem, Johan (1989). 'Semantic Parallels in Natural Language and Computation'. In H-D Ebbinghaus et al., eds. 331-375.

⁴Although one could argue that a construction such as `IF THEN ELSE . FI` is in fact an idiomatic, non-compositional expression.