

SOME LOGICAL CONSEQUENCES OF MAKKAI'S "IDIOMATICITY AS A LANGUAGE UNIVERSAL"

by William J. Sullivan

INTRODUCTION

1. On defining the idiom

Adam Makkai (1972:121) defines the (lexemic) idiom as a "polylexonic lexeme which is made up of more than one minimal free form or (morphological) word, each lexon of which can occur in other environments as the realization of a monolexonic lexeme." Thus, *kick the bucket* and *bite the dust* are idioms:¹ *kick the bucket* has nothing to do with the physical act of kicking; no actual bucket is involved; and therefore, there can be no deixis involved, which is what *the* normally signals.² Conversely, *kith and kin* 'acquaintances and relatives' is a pseudo-idiom, because *kith* occurs nowhere else and with no other meaning in English. The lexemic idiom provides the central datum for consideration herein.

Makkai also defines a sememic idiom, a member of a second idiomaticity area, to account for expressions like *don't count your chickens before they're hatched*. The sememic idiom differs from the lexemic in two important ways. First, it is at a higher (more abstract, closer to meaning) linguistic level, a point to which I return in section IV below. Second, the function words in a sememic idiom (e.g., *before*) have exactly and only their normal language function. The consequences of this observation are extensively discussed by Makkai (1972:128-134), and I do not treat them further. They represent an interesting subset of the logical consequences of the lexemic idiom, rather than forcing a consideration of anything new. Thus, the general considerations of the lexemic idiom cover the logical consequences of the sememic.

2. Idioms and universality

Makkai (1978) explores in detail the occurrence of idioms in several unrelated languages (French, Hungarian, Thai, etc.). All these languages have idioms of both types. Though the case is not closed, it seems likely that idioms

William J. Sullivan is Associate Professor of Linguistics at the University of Florida.

occur in all dialects of all languages.

If stratificational theory is right about the structure of language, then all linguistic universals must be logical (and probably neurological and biological) in nature. They must also have semiotic and cultural/anthropological function. A great deal could be said in these areas. But however interesting this may be, I concentrate exclusively on the logical relations and their consequences.

3. The linguistic structure of the lexemic idiom

Before beginning a consideration of the logical structure of the idiom, one must outline its linguistic structure in more general terms.

Consider *kick the bucket*. It is a lexemic predicate (or verb phrase, in post-Chomskyan terms). This implies that it feeds into the lexotactic relations of the clause. Makkai's definition suggests (to me) that the lexemic idiom feeds into the morphotactics at the level of a morphemic phrase (but cf. Makkai 1978:444). This is a possibility. However, for a number of structural reasons that do not affect the logical discussion (see section I.2), I prefer the lexotactic connection.

Semantically, *kick the bucket* is open to two very distinct interpretations. The literal or unmarked interpretation as a lexotactic construction is found in a discourse like (1). The idiomatic interpretation is found in (2), with or without the portion in parentheses.

- (1) John's foot is sore because, when the electricity failed last night, he went for the candles and _____ standing in the hall.
- (2) (When the electricity failed last night, the intensive care unit shut off and) John _____. That's why he's not here now.

These facts imply two things. First, the literal interpretation is available, but only in appropriate discourse contexts. Second, the idiomatic interpretation is therefore preferred; it takes precedence over the literal (see section I.3).

Finally, consider the sememic (linguistic semantic) relations of idiomatic *kick the bucket*. Its subject is a patient, not an agent, just as is the subject of *die*. It cannot be passivized or otherwise affected by what transformationalists have called "movement" rules. But it can occur with a wide spectrum of tenses. Some tenses, e.g., the progressives, may not occur. This is possibly because they violate some pragmatic or cultural restrictions, but their lack is a realizational or discourse-determined feature of *kick the bucket*, not a feature of the idiom itself. The idiom is, in short, a unitary sememic predicator, not a construct, semotactically indistinguishable from *die*.

Semantically, its relationship to meaning is also basically indistinguishable from that of *die*. A great deal can be said about cultural restrictions on the occurrence of *kick the bucket*: (3) is all right; (4) is simply inappropriate.

- (3) This is my cat's third litter this year. Thank God most of the kittens kicked the bucket last night.

- (4) (to a weeping widow at the funeral parlor) I was so sorry to hear your husband kicked the bucket yesterday.

But while such considerations are important inputs to usage, their relationship to structures is less critical and can be taken care of in a number of different ways. The minimum necessary structure is given in section I.3 (see figure 5, page 147).

As Makkai (1972:122) indicates, the idiom does not differ from a chain of ordinary morphemes in its morphological and phonological relations. Thus the linguistic relations outlined above are all that need be considered.

4. The content side of language

Idiom structure is a part of the content side of language. The content side of language deals with the structures involving the meaning of messages. The stratal organization is summarized in table 1.

Table 1 focuses on the tactic portion of the structure. These tactic patterns are connected by realizational relations. Idiom structure is located in the realizational (interstratal) portion of the content side of language, in three areas: the hypersememic, the semolexic, and the lexomorphemic realizations.

TABLE 1

Semantics	{ Gnostology	General information store; encyclopedia knowledge
	{ Semology	Discourse structure, logical structure of predicates, deep (sememic) case assignments
Grammar	{ Lexology	Phrase, clause, sentence structure; prepositions and lexemic cases
	{ Morphology	"Words," morphological cases, etc.

This bare outline will suffice for the description that follows. One further comment must be included for clarity: the gnostology is not itself a part of language. It does provide inputs to language during encoding and receives outputs from language during encoding.

5. The idiom's logical relations

In this study I translate the general considerations outlined above into the specifically logical relations used in stratificational theory. I show what relations are necessary to describe idioms and where, in linguistic structure, idioms are located. This leads to a general prediction of the process of idiom formation and of the degree of semantic transparency of different types of idioms. I conclude by listing the universal consequences of idiomaticity in logical terms, showing that the humanistic and logical universals involved in idiom structure and formation are completely compatible.

I. THE LOGICAL STRUCTURE OF THE IDIOM

1. The idiom as a realizational relation

Consider that *kick the bucket* is indistinguishable from *die* in its denotata (semantic relationships) and its semotactic relationships. It is obviously different morphologically and syntactically. The relationship of the sememe to the alternative semons (*die* and *kick the bucket*) is expressed by a downward OR, given in figure 1. Here the semantic referent or meaning is given in Latin. The reason for this is that *die* is not the "meaning" of *kick the bucket*. Both are realizations of the same "meaning," which is different from but related to both of them. Hence, the Latin.

The semon *die* is related to the left-hand branch of the OR in figure 1. Its relation to the grammar (syntax/lexotactics and morphology) is a simple, one-to-one realization. But the relation of the right-hand branch is composite—the one-to-many AND realization symbolized by the triangle in figure 2.

The three sememes *kick*, *the*, and *bucket* are morphologically and phonologically identical to the lexemes *kick*, *the*, and *bucket*. They may be said to neutralize with these lexemes. This is expressed by the upward OR relations at the bottom of figure 3, which repeats the relations of figures 1 and 2.

Figure 3 displays the minimum relationships necessary to describe this idiom (see Makkai [1972:151, figure 41]). However, it is necessary to examine the level of this node in linguistic structure in much greater detail than has heretofore been the case.

2. The locus of the lexemic idiom

The AND node in figure 3 can be labeled *kick-the-bucket* and can be identified as the idiom. This is simply a convenient label, which could as well be applied to the line leading into this AND from above; it has neither theoretical significance nor model-specific consequences. What is significant is the identification of the two strata between which the idiom is located.

Makkai's definition implies that these idioms are between lexemic and morphemic strata, as remarked above. This is indeed possible. However, I identify its locus at the level of semolexic relations. There are three reasons for this. First, it is not clear that English has separate lexemic and morphemic strata, a necessity if *kick the bucket* is located at the level of lexomorphemic realizations. Second, the evidence cited above for locating *kick the bucket* requires only that it be below the semology and above the bottom of the morphology. Either semolexic or lexomorphemic realizations fit this requirement. Third, and most important, *kick the bucket* is a syntactic phrase (predicate) as well as a morphemic phrase. This is clear from the fact that unmarked information focus is realized by sentence stress on *bucket*, not *kicked*. Such accent placement can only be accounted for in a general fashion by the location of the idiom on the level of semolexic realizations, connecting (downward) to the lexotactics. Ordinary clause accent relations then apply.

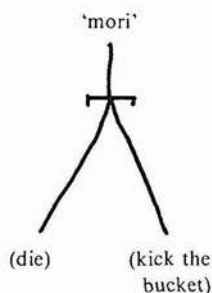


FIG. 1

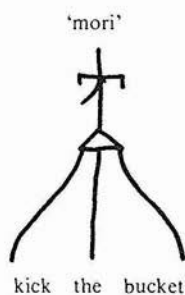


FIG. 2

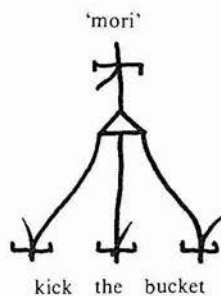


FIG. 3

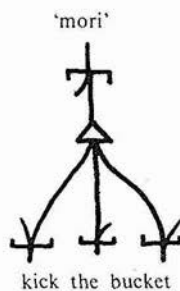


FIG. 4

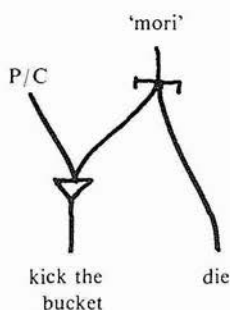


FIG. 5

The structural implications of clause-level stress suffice to require inclusion of *kick the bucket* among semolexic realizations. But there are some additional refinements of logical relations of idioms that become possible if the lexemic idiom is placed on this level (see also Makkai [1978:443ff]).³

3. Some additional considerations

If *kick the bucket* is on the level of semolexicemic realizations, then it is not necessary to order the AND node in figure 3. The lexotactics of the predicate will automatically supply the correct order to the lexemes.⁴ This modification is included in figure 4.

The pragmatic and cultural considerations that tell us when it is appropriate to express 'mori' as *kick the bucket* must also be accounted for. Knowledge of the pragmatic and cultural restrictions is applied to the encoding process in addition to the semological relations. This is expressed logically by an upward AND node related to the appropriate realizational relation. This modification of figure 1 is included in figure 5, where P/C indicates the connection to a set of pragmatic and cultural considerations outside linguistic structure.

Finally, consider the interpretation of (5).

(5) John kicked the bucket yesterday.

In isolation, it is most likely to be interpreted idiomatically. This suggests a preference in decoding for the relationship from the upward OR nodes at the bottom of figure 4 to the downward AND node representing the idiom. This preference is indicated by ordering the upward OR relations. This is included in figure 6, along with the modifications given in figures 4 and 5.

In sum, figure 6 represents the logical structure and relations of Makkai's lexemic idiom.

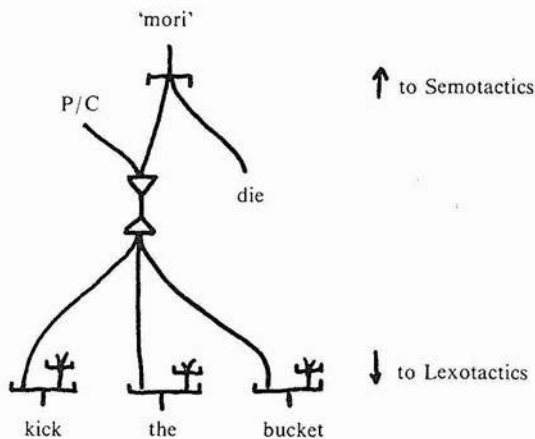


FIG. 6

II. AN ASIDE ON THE PROCESS OF IDIOM FORMATION

1. The creation of a metaphor

Louisiana, Mississippi, and Tennessee were, during the 1820s, a jumping-off place for emigrants from the United States to the Mexican province of Texas. In that region, somebody often left town on his way to Texas, and when his absence was questioned, he was said to have *gone to Texas*. The expression became so frequent that it was even abbreviated (orally) to *GTT*.⁵

Most of the emigrants were pioneers looking for a new start. This description fits William Barret Travis and Davy Crockett. Travis saw Texas as a place where a young man with nerve might start close to the top, in his case, as commander of the garrison at the Alamo. Crockett, after pioneering central and western Tennessee and the United States House of Representatives, had time at age fifty for one more fresh start.

Thus, *GTT* had its literal meaning and a fixed connotatum: 'to pioneer, to make a fresh start.' In short, it had become an expression whose literal and metaphorical interpretations were simultaneously recognized.

2. From metaphor to idiom

Another group of *GTT*-ers included those who were in trouble because of their politics and resultant involvement in illegal dueling, like James Bowie, and those who were outright criminals. In their case, Texas was attractive for its lack of an extradition treaty with the United States. They also were *GTT*, but their "fresh start" was for getting away from their pasts, rather than for pioneering. In their case, the use of the expression *GTT* was purely metaphorical. It was actually a fixed metaphor or sememic idiom.

The following generation's use of *GTT* shows that it had become a lexemic idiom. The geographical area of *GTT* increased in scope to include the entire trans-Mississippi west, while the pioneering purpose lost out to the escaping-the-law purpose. By the Civil War, the expression simply meant 'on the run from the law'—a lexemic idiom just like *kick the bucket*.

Sadly, at least to a romantic, this idiom is now lost, mainly because of our loss of the frontier. The process of idiom formation, however, continues.

3. From idiom to lexemic sign

Lexemic signs provide the direct input from the lexemic stratum to the morphotactics. Lexemic signs and their morphotactic parallels, morphological words, give evidence of structures that were lexemic or sememic idioms to former generations of English speakers. Consider *understand*. It is formally (i.e., morphologically and phonologically) indistinguishable from the preposition *under* followed by the verb *stand*. Of course, it is semantically distinct from them both. There is no possibility of deducing (or deriving) the meaning of *understand* from the meanings of what appear to be the constituent morphemes. There are no constituent morphemes, only an ordered AND relation-

ship, each part of which neutralizes with a lexemic sign which has other, direct relationships to meaning. This is represented in figure 7, which contains a structure exactly parallel to that of figure 4. The ordering on the AND might be dispensable, as it was for figure 4. But its presence allows the *under* to bypass the morphotactics. This is desirable because *under* is not a productive or even a typical verb prefix. In short, the semantic uniqueness and the formal identity with a certain strong verb produces a lexomorphemic analogue to the idiom.

However, *understand* gives no impression of having been a phrase. Recall that a lexemic idiom looks like a phrase with meaning unpredictably related to its shape. Similarly, a morphemic idiom should look like a word resembling a phrase, but with meaning unpredictably related to its shape. The process from lexemic idiom to morphemic is frozen in progress in the spelling of *hot dog*. Phonologically and morphologically, *hot dog* is a single word. Placement of accent, non-predicability of the *hot*, etc., make this clear. Its relation to meaning, however, shows it to be a lexemic sign rather than a phrase: it is not a dog, nor is it made from dog meat, nor is it necessarily hot. But we still write it as two "words," i.e., as a phrase. In both *understand* and *hot dog* types of expressions, it is clear that the language has lexomorphemic analogues to semolexic and (hyper) sememic idioms.

4. The process of idiom formation

The examples given above allow an overall summary of the process of idiom formation. Perhaps predictably, the summary is couched in stratificational terms. For a less technical, less model-specific discussion, see Makkai (1978:443-445), paying special attention to the multiple reinvestability principle. In any case, the terminological differences have no theoretical significance; the divisions below are chosen to parallel the steps detailed in sections II.1-II.3 above.

- (a) Frequent use of a particular expression with more or less fixed and unique form-meaning relationship; resulting in

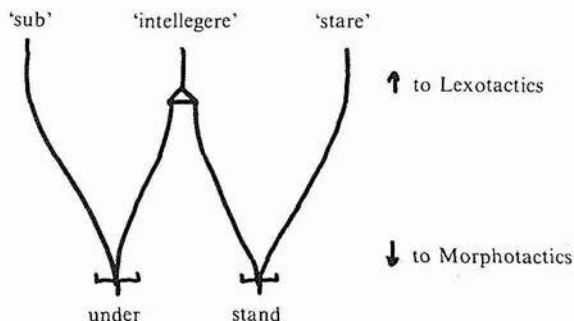


FIG. 7

- (b) Replication of the semotactic relationships of the expression in an AND node in the realizational portion of the network; resulting in
- (c) Two ways to reach the denotatum, via the semotactics or via the replicated realizational node;
- (d) Spread of the environment of usage of the replicated expression; resulting in
- (e) Relation of the AND node to a growing set of connotata in the general semantic store;
- (f) Replacement of the original (composite or deduced) denotatum by one of the connotata, resulting in a sememic idiom;
- (g) Shift of the AND node downward to the semolexic realizational pattern, resulting in a lexemic idiom;
- (h) Shift of the AND node downward to the lexomorphemic realizational pattern, resulting in a morphemic idiom or composite lexemic sign.

This is the maximum pattern, taking generations to complete the process from (a) to (h). The process, once started, is not inexorable. To the contrary, it can be broken or arrested at any point. The cause of this break is not immediately obvious to me. Perhaps there is a maximum size for the AND, beyond which it cannot shift downward to the next stratum (see Makkai [1978:444, last paragraph]).

Idioms, like any other lexical item, can be lost by passing out of current usage. This happened with *GTT*. Many things can happen across generations. But at least some expressions can make it all the way through to (h), as *hot dog* has.

5. The semantic transparency of idioms

In the stratificational organization of language, semology is more closely related to "meaning" than lexology, which is more closely related to "meaning" than morphology. The more intervening strata, the more diversifications and neutralizations there are likely to be in the relational path between a particular idiom and "meaning."

The stratificational prediction, therefore, is that a sememic idiom should be more transparent than a morphemic one. Indeed, many sememic idioms are completely transparent. I can recall first reading *Don't count your chickens before they're hatched* in about third or fourth grade. I immediately "translated" it into (i.e., related it to) the Irish expression *There's many a slip 'twixt cup and lip*. A few years later, while I was reading selections from Robert Burns, the chickens and the lips were joined by *The best laid schemes o' mice and men gang aft agley*.

The same can be done for lexemic idioms that are graphically suggestive, e.g., *bite the dust*. But this requires much more care: *eat dust* refers to a cattle drover riding behind the main herd, picking up strays. Surely the formal resemblances between *eat dust* and *bite the dust* are much more obtrusive

than any formal resemblances between chickens, lips, and mice in the sememic idioms cited. Nonetheless, the lexemic idioms are less "transparent." That is, they are somehow semantically farther apart and show less mutual relatedness, because they are less directly related to "meaning."

Try to come up with a transparent morphemic idiom; it can be done for English, but only on an ad hoc basis. Explain *understand* and *withstand* in a systematic and integrated description. Try *understand* and *undergo*. And so on.

This is not to suggest that people do not try to make all idioms transparent. In many cases, clever folk etymologies arise, etymologies that are more convincing than the historically true ones. These folk etymologies can be useful in teaching idioms to non-native (or even developing native) speakers of English. Such a teaching device should not be contemned.

But in general, observations about the transparency of idioms, whether sememic, lexemic, or morphemic, parallel the predictions made within the stratificational view of language.

III. UNIVERSAL CONSEQUENCES OF IDIOMATICITY

1. The logical relations

Figure 6 suggests directly that upward and downward unordered OR nodes, upward and downward unordered AND nodes, and upward ordered OR nodes are necessary to describe the logical structures of idioms. Figure 7 adds downward ordered AND nodes to this list. If we include downward ordered OR nodes, which occur frequently in all tactic patterns, and upward ordered AND nodes, which seem not to be necessary, we have the inventory of nodes posited by stratificational theory.

Insofar as idioms are universal across languages, therefore, these logical relationships are universally applicable, even necessary, for the description of language. In other words, idioms fit right into the relational network theory of language known as stratificational theory and give positive reinforcement to its basic assumptions. Idioms are fully integrable into a stratificational description of language.

Ironically, this is the diametric opposite of the effect of idioms on transformational theory. The major claim of Fraser (1970) is that idioms are not affected by transformations. Now stratificational theory relates levels (strata) via realizational relations. Transformational-generative theory relates levels (components) via transformations. Idioms are integrable into the set of realizational relations, but not into the set of transformations. The consequences of these observations are worthy of discussion, some of which is given in Makkai (1972:47-57). Idioms are important, and they must not be swept under the rug—or into the lexicon—and forgotten.

2. The stratification of language

One other consequence of idiom structure should be mentioned. The evidence adduced for idioms in English suggests a division of the content side of language into three strata: sememic, lexemic, and morphemic. I say "suggests" because I am not sure what should be considered sufficient evidence for defining a stratal boundary. Detailed research into the idiom structure of non-Indo-European languages, combined with research on alternations, neutralizations, and tactic specifications is necessary to show how much stratification is universal, how much is language-specific.

IV. CONCLUSIONS AND POSSIBILITIES

1. Recasting concepts into logical terms

The idiom is a broad concept, stretching from proverbs and even institutionalized or often-quoted metaphors to words like *hot dog* and *understand*. It is so broad, so apparently unbounded a concept that the methodology of its description in logical terms is not immediately obvious. However, the process of description detailed above shows that idioms can be described by a certain configuration of logical relationships with a regular, patterned relation of these configurations to sememic, lexemic, and morphemic strata.

There is a real advantage to this process. Our normal descriptions of such concepts are given via language. However detailed the descriptions may be, they are necessarily circumlocutions. This is not bad, but it makes comparison of different but related concepts difficult, since there is no way to insure that two circumlocutions will follow exactly the same path. Translation of the concepts into logical terms avoids this difficulty, because the comparison can then proceed on the same (logical) basis in each case.

But we must be careful in logical analysis to avoid the trap of reductionism. Makkai has indeed avoided reductionism; in fact, Ruhl's (1977) criticisms of Makkai suggest that Makkai has gone too far in the other direction, namely, that he has tried to give an explicit account for more than is necessary. In this Makkai has erred, if at all, in the direction of redundancy, rather than taking a chance on leaving important considerations untouched (but cf. Makkai [1975]; Ruhl's thinking is correct, though the criticisms do not really apply to Makkai). In fact, redundant storage seems to be a characteristic of human (i.e., real) linguistic systems. Thus, even Makkai's original work is not seriously lacking in this way. In fact, his original work has made this study possible, straightforward, and even easy.

2. Humanistic and logical universals

In the introduction to this paper I remark that "linguistic universals must be logical . . . in nature." This comment proceeds directly from my interpretation of the stratificational view of language. Many who know of my linguistic pro-

clivities have expressed doubts that humanistic and logical considerations are compatible, let alone intertranslatable. Idioms, their structure, formation, and use, with their relatives metaphor and metonymy, are surely a humanistic function of language, and apparently a universal one. I show herein that idioms are consistently describable in the terms of logical relations. This type of description is effected by an intertranslation/interpretation, thus demonstrating that the humanistic and the logical are perfectly compatible. In the cases of explanation and explication, they are usually complementary and never in conflict. In fact, we need both. We cannot afford to ignore any of our possible tools. The humanistic approach is millennia old and rightly revered. Makkai and stratificational theory have shown us how to make the best use of the logical approach.

NOTES

1. Of course, Makkai (1972) makes a much more precise classification than this. See also Makkai (1978, sections 2.2.2 and 2.3).
2. Makkai (1972:123) deals with this differently. Our arguments are complementary.
3. One further non-structural reason—as opposed to a justification—for this locus exists; I discuss it in section II.3 below.
4. I consider this an additional advantage to the semolexic locus, not a structural justification of such a choice.
5. Professor Howard Lamar of Yale University (now Dean of Yale College) first put me on the trail of this expression.

REFERENCES CITED

- Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6:22-42.
- Makkai, Adam. 1972. *Idiom Structure in English*. The Hague: Mouton.
- . 1975. The metaphorical origins of idiomaticity. Robert J. DiPietro, ed., *Georgetown Working Papers in Linguistics* 11:10-59.
- . 1978. Idiomaticity as a language universal. In Joseph A. Greenberg, Charles A. Ferguson, and Edith Moravcsik, eds., *Universals of Human Language*, vol. III, pp. 402-448. Stanford, California: Stanford University Press.
- Ruhl, Charles. 1977. Idioms and Data. In J. DiPietro and Edward L. Blansitt, Jr., eds., *Third LACUS Forum*, 1976.