Idiomatic phrases differ in the extent to which they can undergo syntactic operations and still retain their idiomatic interpretation. Fraser’s (1970) thorough description of this phenomenon distinguishes six levels of syntactic “frozenness”; for example, he assigns build castles in the air to level 0 (completely frozen), whereas let the cat out of the bag appears in level 5 (virtually unfrozen):

(1) *Castles were built in the air all the time by her improvident husband.
(2) *He built castles most outrageously in the air.
(3) The cat was let out of the bag that evening by their over-excited child.
(4) He let the cat most decidedly out of the bag.

The mental representation of idioms is not differentially affected by variation on this dimension; idioms are stored and accessed as unitary lexical items, irrespective of their level of syntactic frozenness (Swinney and Cutler (1979)). Syntactic frozenness is apparently not correlated with differing manner of mental representation.

The present study was undertaken to examine an alternative possibility, namely that syntactic frozenness may be correlated with the length of time that the idiom has been present in the language. To investigate this, the sample of idioms of

This research was supported by the Science Research Council. Thanks to C. J. Darwin, M. Deuchar, and P. N. Johnson-Laird for useful comments.
differing levels of frozenness that Fraser (1970) used was simply checked against the Oxford English Dictionary (1933; henceforth OED) to ascertain the earliest attested citation for each one.

Note that this procedure will yield only an approximate estimate. First, Fraser’s list of examples is a very small subset of the total population of idioms in the English language: and moreover, as Fraser points out, his assignment of a particular idiom to a particular level represents the state of affairs in his own dialect and may not hold for some or most other speakers of English. Second, the OED’s written sources are obviously more numerous in later centuries than in earlier, so that the chance that a particular form present in the language will occur in the available written corpus is rather smaller for earlier times than for later; in addition, the type of literary text available to the lexicographer may change across the centuries, and some types of text may be more likely to contain idioms than others. There is no reason to believe, however, that the variations across literary sources are confounded with syntactic frozenness.

Fraser’s (1970, 40–41) lists of “representative examples” for each of his six frozenness categories yielded 131 idioms. Of these, 106 were found to have entries in the OED, with the earliest citations for their unequivocally idiomatic readings ranging from the thirteenth to the twentieth centuries. (All the idioms could also take a literal reading: idiom-only strings—e.g. cast a slur on—all belong to the completely frozen category.) A further eight were located in forms minimally variant from those given by Fraser (e.g. make the best of a bad job rather than Fraser’s ... bad deal). Fourteen were not listed in their idiomatic sense at all, and a further three which were listed had no dated citation.

The proportions of earliest citation dates across centuries for each level of syntactic frozenness are given in table 1. Level 0 contains those idioms Fraser holds to be completely frozen, undergoing no syntactic operations at all; level 1 is almost completely frozen (that is, idioms of this class can undergo only minimal transformation); and so on through level 5, which contains those idioms least restricted in the operations they can undergo while retaining their idiomatic sense. (Fraser claims that no idioms are completely unrestricted.)

Frozenness and age are not perfectly correlated; but there is a reliable tendency for the more frozen idioms to have been longer in the language. An estimate of correlation (Kendall’s tau) reveals that the distribution shown in table 1 differs significantly from that which would be expected by chance alone: \( \tau = .17, z = 2.79, p < .005 \). If the date distributions are collapsed across the two most frozen, the two mildly frozen, and the two least frozen levels, respectively, the median earliest citation date occurring in the dictionary for levels 0 and 1 is
Table 1. Earliest citations given by OED: proportions of each frozenness category occurring in each century

<table>
<thead>
<tr>
<th>Fraser's level:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Century</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13th</td>
<td>.05</td>
<td>.08</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14th</td>
<td>.12</td>
<td>.05</td>
<td>.03</td>
<td>.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15th</td>
<td>.04</td>
<td>.09</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16th</td>
<td>.21</td>
<td>.16</td>
<td>.41</td>
<td>.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17th</td>
<td>.26</td>
<td>.20</td>
<td>.05</td>
<td>.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18th</td>
<td>.05</td>
<td>.16</td>
<td>.14</td>
<td>.07</td>
<td>.13</td>
<td>.16</td>
</tr>
<tr>
<td>19th</td>
<td>.32</td>
<td>.16</td>
<td>.18</td>
<td>.50</td>
<td>.34</td>
<td>.37</td>
</tr>
<tr>
<td>20th</td>
<td></td>
<td>.07</td>
<td>.03</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not listed</td>
<td>.11</td>
<td>.08</td>
<td>.05</td>
<td>.28</td>
<td>.06</td>
<td>.16</td>
</tr>
<tr>
<td>No citation</td>
<td></td>
<td>.07</td>
<td>.03</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Number of examples given by Fraser) (19) (25) (22) (14) (32) (19)

1631, for levels 2 and 3 1705, and for levels 4 and 5 1801. It might be claimed that the 14 idioms not listed in the dictionary are of recent origin and should therefore be assigned to the twentieth century; if this is done, the median earliest citation dates shift to 1660. 1777, and 1829, respectively.

Thus, imperfect as the present analysis must be, it suggests that those idiomatic expressions which are least susceptible to syntactic operations are also those which have been in use in the language in their idiomatic form for the longest time. There are several possible reasons why this should be so. On the one hand, it may be that syntactic freezing is a gradual process that afflicts idioms over a period of decades or centuries. If this is the case, then one might expect to find that idioms which are highly frozen today occur in a less frozen form in texts from earlier times. Limited support for this hypothesis can be found even in the OED, which cites instances in which *build castles in the air* (Fraser’s level 0) was used in passive form in 1575 and 1630, and *burn the candle at both ends* (Fraser’s level 1) with an inserted parenthesis in 1848. On the other hand, syntactic frozenness may occur when the meaning of the idiom is no longer obvious, the original literal reference having become obscured by time. (Recall that idioms lacking any literal reading at all are assigned to Fraser’s level 0.) On this hypothesis, *kick over the traces* and *let off steam* would belong to Fraser’s level 0 because nowadays we travel by horseless carriage and electrified railway.

References


A CASE AGAINST THE MORPHOPHONEMIC–ALLOPHONIC PRINCIPLE
Yoshimitsu Kanai, Gunma University

1. Koutsoudas has proposed the Morphophonemic–Allophonic Principle as one of the universal principles of rule application.¹

(1) The Morphophonemic–Allophonic Principle
A morphophonemic rule application must always take precedence over an allophonic rule application. Given a rule \( A \rightarrow B / C \_ \_ \_ D \) applicable to a form CAD, the application of this rule is morphophonemic if there are strings of the form CBD which could be derived from a source other than CAD; otherwise the application of the rule is allophonic. (Koutsoudas (1980, 32))

Koutsoudas (1980, 33) also claims that the distinction between morphophonemic and allophonic rules is precisely Kiparsky’s distinction between neutralizing and nonneutralizing rules.

(2) Given a phonological process \( P: A \rightarrow B / XC \_ \_ \_ DY \), \( P \) is neutralizing if there are strings of the form CBD in the immediate input of \( P \); otherwise \( P \) is non-neutralizing. (Kiparsky (1976, 169))

Here I would like to discuss an example from the Tohoku dialect of Japanese which suggests that Koutsoudas’s distinction between morphophonemic and allophonic rules does not correspond precisely to Kiparsky’s distinction and that the Morphophonemic–Allophonic Principle does not always predict the precedence of rule application.

2. The Tohoku dialect of Japanese has the following rules (Inoue (1968), Miyajima (1961), Muraki (1970), Ohishi and Uemura (1975)).²

(3) a. Syllabic Nasal Formation
\([ + \text{ nasal}] \rightarrow [ + \text{ syll}] / \_ \_ \_ \{C, #\}

b. Nasal Insertion
\( \phi \rightarrow \left[ \begin{array}{c}
\text{n} \\
\text{n}
\end{array} \right] / V \_ \_ \_ \left[ \begin{array}{c}
\text{d} \\
\text{g}
\end{array} \right] V \)

¹ This principle was first proposed in 1973 at the annual meeting of the Kentucky Foreign Language Conference, Lexington, Kentucky.
² The details of the rule forms and the quality of high vowels are omitted for the sake of clarity.