

The Access and Processing of Idiomatic Expressions

DAVID A. SWINNEY

Tufts University

AND

ANNE CUTLER

University of Sussex

Two experiments examined the nature of access, storage, and comprehension of idiomatic phrases. In both studies a Phrase Classification Task was utilized. In this, reaction times to determine whether or not word strings constituted acceptable English phrases were measured. Classification times were significantly faster to idiom than to matched control phrases. This effect held under conditions involving different categories of idioms, different transitional probabilities among words in the phrases, and different levels of awareness of the presence of idioms in the materials. The data support a Lexical Representation Hypothesis for the processing of idioms.

Investigation of the processes underlying sentence comprehension has tended, understandably enough, to consider the more simple cases which confront the listener. For example, examinations of lexical processing during sentence comprehension have typically employed words with straightforward and easily derived interpretations. While an important exception to this rule is the study of lexical ambiguities, even this work deals with words which possess relatively well-defined literal meanings. There is, however, a large class of items—including idioms, metaphors, and proverbs—which are extraordinarily common in the language but which have been largely ignored in comprehension research; the few attempts to discover how the compre-

hension device computes nonliteral meanings for these items have found them to be remarkably intractable. The high frequency of occurrence of nonliteral materials alone necessitates that the processes responsible for their comprehension be incorporated in any complete comprehension model. Idioms, in particular, present an important challenge to those interested in building a performance model for sentence comprehension, and the purpose of this paper is to examine some of the basic aspects of how we come to understand idiomatic expressions.

In order to see the nature of the problems that idioms raise for the comprehension device it is first necessary to consider exactly what constitutes an idiom. In its simplest form an idiom is a string of two or more words for which meaning is not derived from the meanings of the individual words comprising that string. Thus, the idiomatic meaning of "kick the bucket" has little to do with the meanings of either "kick" or "bucket"; similarly, the meaning of "by and large" has little to do with the meanings of either "large" or "by."

Thus, by and large, idioms defy traditional concepts of syntactic and semantic analysis,

This work was supported in part by NIMH Grant No. 1RO3MH29885-01 and a Tufts University Faculty Research Award, both to David Swinney. The authors express their gratitude to Carol Hendricks Rodriguez, Peter Kavorkian, Neysa Kaplan, Anita Lubetsky, Shelley Rosenfeld, and Stewart Sherman for aid in gathering and selecting experimental materials, and to Stewart Sherman for help in testing subjects. Requests for reprints should be sent to David A. Swinney, Psychology Department, Paige Hall, Tufts University, Medford, Mass. 02155.

thereby creating problems for both performance and linguistic models. These problems are somewhat amplified by the fact that some idiom strings are not syntactically well formed—often violating both selectional and subcategorizational restrictions which exist in grammatical theory.

Were these problems the only ones, perhaps they could be handled within the confines of existing theory (although that is not at all a certainty). However, most idioms also have the characteristic of ambiguity. That is, grammatical (well formed) idioms, such as “kick the bucket,” have an acceptable literal reading for the word string which *is* derivable by traditional syntactic/semantic analysis as well as an idiomatic reading which cannot be derived from such analysis. The problem is thus one of developing a comprehension model for dealing with phrases which are simultaneously amenable and nonamenable to traditional compositional analysis.

Those few attempts to deal with the problems posed by idioms for linguistic theory have largely revolved around studying the susceptibility of idioms to traditional transformational and semantic analysis. Most of these attempts have focused on grammatical idioms, and two general (and opposing) themes can be discerned. One is that idioms should be treated as a separate, specialized list of items, which have special insertion privileges in the grammar (see, e.g., Newmeyer, 1972; Weinreich, 1967). The other position holds that idioms should be treated as single lexical items which have much the same characteristics in the grammar as do any other lexical items (see, e.g., Fraser, 1974; Heringer, 1976). Fraser has refined this latter proposal with the suggestion that idioms do not form a homogeneous class with respect to transformations which they may undergo. He has categorized idioms on a continuum ranging from those which will not undergo even the most simple transformation and still retain their idiomatic meaning (e.g., “by and large,” “face the music”) to

those which undergo nearly all traditional transformational changes without losing their idiomatic reading (e.g., “throw in the sponge”). While linguistic studies such as this provide some interesting concepts with which to work, it appears fair to comment on the current work by noting that none of it has thus far succeeded in even providing a descriptively adequate account for the class of items comprising idioms; the problems raised by Chafe (1968) and Weinreich (1967) still exist for linguistic theory today. There is not space even to attempt to do justice to the extensive linguistic literature on this topic, but it appears that at this time this work has no finalized contribution to make toward our goal of building an idiom comprehension model.

It has, on the other hand, provided the basis for the two general processing models which have undergone some experimental examination. The first of these, which can be called the Idiom List Hypothesis, holds that idioms are stored in (and accessed from) a special list which is not part of the normal lexicon. Access from this list takes place via what has been called a special “idiom mode” of processing (Bobrow & Bell, 1973) which is seen to differ from a normal, literal processing mode in terms of both manner and privilege of access. While the exact manner in which idioms are located in this processing mode is only vaguely defined, a necessary condition on the Idiom List Hypothesis is that a literal analysis is always attempted on a word string before an idiom mode of processing is undertaken. It is this factor that critically distinguishes this model from other models, and the constraint justifies making the distinction between an idiom (nonliteral) and a literal mode of processing. It is typically assumed that if the comprehension device has already been put into a mode of idiom awareness, then the idiom processing mode might, under these circumstances only, be called into play before a literal processing strategy. Certainly this assumption fits well

with the intuition that one tends to become aware of the presence of a large number of idioms in a conversation once any particular idiom has been brought to one's conscious attention.

The second model, which can be labeled the Lexical Representation Hypothesis, holds that idioms are stored and retrieved from the lexicon in the same manner as any other word. Thus there is no special idiom list nor any special processing mode under this hypothesis. Rather, it is assumed that computation of both meanings—idiomatic and literal—is simultaneously initiated upon occurrence of the first word in the idiom string, much as appears to occur for lexical ambiguities (see, e.g., Foss & Jenkins, 1973; Lackner & Garrett, 1972; Swinney, Note 1). Thus individual words are accessed from the lexicon and structural analysis is undertaken on these words at the same time that the lexical access of the entire string (which is merely a long word) is taking place.

The experimental evidence which currently exists appears, at least on the surface, to favor the Idiom List Hypothesis. Bobrow & Bell (1973) examined this hypothesis by employing a perceptual set paradigm in which subjects either saw a set of four sentences which had only literal interpretations or four sentences which contained idiomatic phrases followed by a grammatically idiomatic (and therefore ambiguous) sentence. Subjects were asked to report which meaning of the grammatical idiom string they first perceived in each of the conditions. Overall, when these were compared to a baseline condition, an increase in the number of "first perceived" idiomatic meanings was reported following an idiomatic perceptual set and a similar increase in the number of "first perceived" literal meanings occurred following a literal set, leading the authors to claim support for a special "idiom mode" of processing. The reasoning behind this conclusion was that if subjects were able to find or to avoid finding the idiomatic meaning of a sentence by virtue

of the presence of a processing set (and in the absence of any specific semantic or syntactically necessitated choice) separate modes of processing must have been called into play.

Similarly, Brannon (Note 2) reported that the reaction time needed to decide that two idiomatic sentences had different meanings was longer than that required for a decision that two nonidiomatic sentences differed (the idiomatic and nonidiomatic sentences were semantically equivalent). Brannon reasoned that decision times should only be longer for idiom strings if they were undergoing processing that was different from (and, perhaps, additional to) the processing normally given to a string of individual lexical items. The data, then, seem to support the hypothesis that a literal analysis is first undertaken, followed by a subsequent idiom mode analysis.

While these studies fairly well exhaust the experimental literature on the perception of idioms themselves, there has been some related work on the comprehension of other nonliteral speech forms. In a study of conversationally conveyed requests, for example, Clark and Lucy (1975) concluded that subjects first comprehend such requests at a literal level and then, only if this interpretation is contextually inappropriate, construct their nonliteral interpretations. This work suggests that the type of processing proposed in the Idiom List Hypothesis may exist, at least for some speech forms.

However, there are at least two factors which militate against accepting the Idiom List Hypothesis without further investigation. The first and most important of these is that the research supporting the Idiom List Hypothesis has relied exclusively upon post-perceptual measures for support of inferences about ongoing idiom comprehension processes. A number of reasoned arguments exist (see, e.g., Cutler and Norris, 1979; Garrett, 1970; Marslen-Wilson, 1976)—not to mention those supplied by one's own common sense—which suggest that moment-to-moment processing is reflected quite poorly

by tasks which are applied after the process of interest has been completed. The second, related, factor is that many such postperceptual tasks appear to reflect merely a reprocessing and/or conscious access of the material under consideration (see Foss & Swinney, 1973, for discussion). In sum, post-perceptual tasks are not necessarily capable of supporting inferences about perceptual processes; any task which measures effects only after they are over runs the risk of reflecting merely the final, conscious, result of such processing and not the processes by which that final interpretation was achieved. Certainly the Bobrow and Bell experiments, which provide the strongest support for the Idiom List Hypothesis, fall into this category. It is not really surprising that the comprehension device decides that the appropriate meaning for a word string is an idiomatic one in the context of a set of idioms, and a literal one in the context of a literal "set." Such a decision, however, does not at all mean that we should assume that two different types of processes are occurring in these circumstances. Rather, we merely need postulate a decision device which considers context in a pragmatic manner and which allows the most likely of several computed meanings to be made available to conscious access.

What is needed in order to better evaluate the candidate hypotheses, then, is a task designed to tap the time course of the access and processing of literal and idiomatic strings more closely. To this end, we employed a task which required subjects to decide whether or not a visually presented string of words formed a meaningful, natural phrase in English. The rationale behind this Phrase Classification Task is relatively straightforward: the task requires that candidate strings be analyzed for sensibleness *as a unit*. If idiomatic meanings are computed by reference to a special idiom list, via some special mode of processing which is instigated following an attempt at literal computation, the phrase classification decisions should take

longer, or at least no less time, for grammatical idioms than for nonidiomatic phrase controls. If, on the other hand, the Lexical Representation Hypothesis holds, decisions made to idiomatic strings should be faster than those made to literal word string controls. That is, under this hypothesis the computation of a literal meaning and the access of a lexical (idiomatic) meaning should be undertaken simultaneously for the idiom string. The access of the lexical interpretation should conclude far more quickly than the access and computation of the relationships among the several lexical items in the literal interpretation of the idiom. Similarly, the lexicalized idiom meaning should also be recovered before the access and computation of relationships between words in the literal control phrase. Hence, if the Lexical Representation Hypothesis is to hold, idiomatic phrases should be consistently classified as "acceptable" more quickly than literal control strings. (Note that the same outcome, again in support of Lexical Representation Hypothesis, would be predicted by a slight variation of this argument. In this, both the lexicalized idiom and the literal analysis are undertaken simultaneously on the idiom string, thus providing two interpretations for that string. Because the control string will have only a single interpretation, one would expect that, on the average, classification responses to the idiom would be faster than those made to the control. That is, any one of the two simultaneously generated interpretations will suffice for a decision in the case of an idiom, but only one is available for the control; thus, in a horse-race model the idiom will win. Again, this result will hold only if the Lexical Representation Hypothesis is correct and both interpretations are accessed simultaneously for the idiom string.)

Finally, as work exists which suggests that the initial portions of a word are often critical for its access (Mehler, Segui, & Carey, 1978; Taft & Forster, 1975), two types of control word strings were created in order to examine

the possible bases upon which a putative lexical interpretation of an idiom might be recognized.

EXPERIMENT 1

Method

Subjects. Twenty undergraduates from Tufts University participated in the experiment as partial fulfillment of an introductory psychology course requirement.

Materials. The experimental materials consisted of 23 grammatical idiomatic word strings which were matched (yoked) with 23 grammatical control word strings. The experimental idiom strings each had approximately equal *a priori* biases toward their literal and idiomatic interpretations, as determined by a pretest given to an independent group of 22 Tufts undergraduates. In this pretest, subjects were shown 100 idiomatic and "filler" word strings one at a time and were required to indicate the first meaning they thought of for each stimulus. Word strings were considered approximately equi-biased if the least frequent interpretation constituted no less than 25% of all reported interpretations (mean percentage of least frequent interpretation for all experimental materials was 42%).

The control strings were constructed by replacing one word in each idiomatic string (either the first or the last word) with a word of the same length (in both letters and syllables) and of equal or higher frequency (Kučera & Francis, 1967, norms). The resulting control strings each formed a literal, grammatically acceptable, and common, English phrase. For 12 of the 23 control strings, the first word of the idiom was replaced (e.g., wrap it up/lift it up), and for the other 11 control strings, the last word of the idiom string was changed (e.g., break the ice/break the cup). See Appendix 1 for a list of the experimental materials.

In addition, 30 acceptable English word strings which were not idiomatic were in-

cluded in the study in order to minimize the ratio of idiom-to-nonidiom strings. Finally, 76 word strings which did not form an acceptable (grammatical) English phrase (i.e., "stranger is during"; "destroy be however") were created. These nonphrase strings were roughly equivalent to the acceptable-phrase strings in both number of syllables and number of words.

Procedure. The 152 word strings were presented in random order on a CRT display screen to each of the subjects. Each string was displayed for 2 seconds, with the interval between successive presentations (following a response) varying randomly between 0.5 and 2.0 seconds. Materials presentation was under control of a PDP8/e minicomputer. The subjects were instructed that they would see a series of short strings of words on the display screen, and that their task was to decide, as quickly as possible, whether or not each of these strings formed a meaningful, "natural" phrase in English. Subjects were to indicate their decision by pressing one of two buttons which were in front of them. Five practice phrases were then given. (No subject had any difficulty in performing the phrase classification.) Response latencies were measured beginning with the onset of each phrase and measurement was terminated by each subjects' button press. Response latencies and classification responses were collected and stored by a PDP8/e computer (see Onifer, Hirshkowitz, & Swinney, 1978, for description of the relevant psycholinguistic software package).

Following the experiment, subjects were debriefed and questioned as to whether they had been consciously aware (during the experiment) of the presence of idioms in the materials.

Results

Mean latencies for correct responses were computed for each subject for each experimental condition. These were subjected to an

analysis of variance with Phrase Type (Idiom vs Control) and Substitution Position (First vs Last Word Change) as main variables. The mean latencies taken across subjects for the four conditions formed by the interaction of these variables are displayed in Table 1.

TABLE 1

MEAN LATENCIES (IN MILLISECONDS) FOR CELLS OF THE PHRASE TYPE \times SUBSTITUTION POSITION INTERACTION

Phrase type	Substitution position	
	First word	Last word
Idiom	918	995
Control	992	1064

The main effect for Phrase Type, $F(1, 19) = 48.5$, $p < .0001$, and Substitution Position, $F(1, 19) = 16.5617$, $p < .0007$, were each significant in analyses employing subjects as the random factor. The interaction of these variables was not significant, $F(1, 19) = 0.02$. $\text{Min}F'$ values (Clark, 1973) computed from independent analyses of variance employing subjects (see above) and items as random factors were also calculated for these data. Again, both main effects for Phrase Type, $\text{min}F'(1, 49) = 4.29$, $p < .05$, and for Substitution Position, $\text{min}F'(1, 58) = 4.06$, $p < .05$, were significant. The interaction of these variables did not approach significance under $\text{min}F'$ analysis.

Error rates were also tallied for each subject in the experiment. The error rates were nearly identical for the four experimental conditions, and the overall error rate in the experiment was extremely low (3.96%).

Discussion

For both of the Substitution Position conditions we find that grammatical idioms are judged to be acceptable English phrases far more quickly than are their matched controls. These data provide clear support for the Lexical Representation Hypothesis; as recovery of any acceptable meaning was suffi-

cient for a positive classification response, and as the access of any single lexical item (the lexicalized idiom) can undoubtedly be accomplished more quickly than the access and computation of the relationships among the several words in a (control) phrase, the results support a model in which idioms are stored and accessed as lexical items. In addition, at least in the most parsimonious account, this lexicalized meaning can be assumed to be accessed simultaneously with computation of the literal meaning for the idiom. Note that if the meaning of an idiom were computed from the individual items in the phrase, or via an idiom processing mode coming into play following attempted literal analysis, then idioms would be recognized no faster than the control phrases.

Lest the argument arise that the reported effects were obtained because our subjects were always in a special idiom processing mode (in which case the Idiom List Hypothesis might still be viable), we performed two additional analyses. In the first we analyzed the very first occurrence of an idiom in the experiment, along with its matched control, for each subject. As presentation was random for each subject, the particular idiom which occurred first differed for most subjects. Also, in a number of cases the control for the "first" idiom preceded the idiom itself. The mean reaction time for the first idiom seen by each subject, as computed across all subjects, was 948 milliseconds and the mean reaction time for the matched control phrases to these idioms was 999 milliseconds. This difference was significant, $t(19) = 2.746$, $p < .01$. Thus, even prior to occurrence of any material which could put the subject into an "idiom mode," the idiom processing main effect is similar in direction and strength to the results found across all materials.

In addition, analysis of the post-test questionnaire revealed that 11 of the 20 subjects admitted to *some* (although often doubtful) awareness of the presence of idioms at some point during the experiment. A separate, post

hoc, analysis was undertaken on the data from "aware" and "unaware" subjects in order to discover whether this distinction was responsible for processing differences which might have been obscured in the original analysis. Substitution Position Conditions were collapsed for this analysis. For those subjects who had some awareness of the presence of idioms, mean reaction time to the grammatical idioms was 936 milliseconds and mean reaction time to the control phrases was 1000 milliseconds. For those subjects who were not aware of idioms in the experiment, mean reaction times to idioms and their controls were 981 and 1062 milliseconds, respectively. The data were submitted to an unweighted means unequal n analysis of variance, in which the main effect for Phrase Type (Idiom vs Control) was significant, $F(1, 18) = 48.43$, $p < .0001$, but both the main effect for Awareness Group and the Awareness Group \times Phrase Type interaction failed to reach significance ($F(1, 18) = 0.572$ and 0.697 , respectively).

In all, it appears there is no basis for believing that our subjects were in any "special" idiom processing mode in this experiment; subjects who might have been in such a mode due to "awareness" of idioms actually performed no differently than those who could not have been in such a mode on that basis. These results appear to substantiate the argument in favor of the Lexical Representation Hypothesis.

A brief word concerning the first vs last word Substitution Position variable is in order. This variable was originally included as a control for the possibility that the initial morpheme in an idiom might be a critical factor in the access of an idiomatic meaning. For example, it was thought that this initial morpheme might be *sufficient* for initiating access of an idiomatic meaning. If true, one might expect a far greater difference in access time between an idiom and its control phrase in which only the final word differed than for an idiom and control phrase in which only

the initial word differed. In any event, such speculation is irrelevant given the obtained data; the idiom/control effect is equivalent under the two substitution conditions and it appears that the absolute reaction time differences obtained between these conditions merely reflect the overall levels of difficulty of the two sets of materials.

Finally, the notion of transitional probability needs to be discussed in relation to these materials and results. It might be felt that the advantage shown by idioms in this task is simply one caused by there being a higher likelihood of occurrence for the individual lexical items forming the idiom phrase than for those items forming the control phrase. In its simplest form, this hypothesis must be wrong; the substitution controls were chosen with care to avoid any such bias (the substituted word was always a word of the same form class and of equal or higher frequency than the word it replaced in the idiom). Thus, under this hypothesis one would expect that the control phrase might actually have developed a statistical edge on frequency of occurrence as a unit. However, it might be felt that it is the *sequence* of items in the phrase that is critical here, not just the summed frequencies of individual items. It can be noted that the notion of phrasal frequency based on transitional probability can be relatively easily explained and motivated by the concept that we have adopted—that they have lexical status. However, we felt that it would be best to examine some evidence which would speak directly to the role of transitional probability between words in these materials.

To this end, a test was constructed in which the 23 idiom phrases, each with its final word omitted, were presented on a typed list to 40 Tufts University undergraduates. These materials were randomly intermixed with 100 "filler" incomplete phrases. The composition of the filler materials was in the same ratio as that of the experimental materials with respect to the number of words in each phrase. The purpose of the filler materials was to de-

crease the possibility of subjects noticing that all the experimental materials could be idioms. In addition, the 11 control phrases which were in the First Word Substitution Position condition were added randomly to this list, again omitting the final word of the phrase. Subjects were required to give the first word that they thought of after reading each incomplete phrase. The number of completion responses which were identical to either the idiom or the control materials was tabulated as a percentage of the total number of responses. The results fell into rather evenly divisible categories. For 10 of the 23 idiom/control pairs the number of completion responses to the idiom and its control was identical. For two cases the control materials had a higher frequency of response completion than their associated idioms. The remaining 11 idiom/control pairs had higher response completion percentages for the idiom than for the control phrases (ranging from 3 to 40% greater). For purposes of comparison, the data were divided into two groups: one group comprised the 10 idiom/control pairs which had equal completion percentages and the other group was comprised of the 10 idiom/control pairs with the largest idiom completion advantage over their respective controls. Another analysis of variance was performed on the reaction time data examining the two Response Completion categories as a main effect. The results are quite straightforward. The main effect for Phrase Type is significant, $F(1, 19)=80.36$, $p < .0001$, as was the main effect for Response Completion category, $F(1, 19)=49.7$, $p < .0001$. However, the critical Response Completion \times Phrase Type interaction was not at all significant, $F(1, 19)=0.0848$. The mean reaction time data for each of these conditions can be seen in Table 2.

In short, there is no evidence that transitional probability *per se* plays a role in the effect reported here for idioms. Rather, it appears that the lexicalization hypothesis is upheld as a strong and viable hypothesis for

TABLE 2
MEAN REACTION TIME (IN MILLISECONDS) TO IDIOM AND CONTROL PHRASES AS A FUNCTION OF RESPONSE COMPLETION (TRANSITIONAL PROBABILITY) EXAMINATION

Phrase type	Response completion category	
	No idiom/ control difference	Idiom advantage over control
Idiom	903	1002
Control	998	1107

the processing and representation of idioms.

It is possible that the findings reported here, while relatively straightforward, still may not constitute the complete and final picture of the processing of idioms. A possible basis for qualification of our conclusions might lie, for example, in examining variations in types of idioms. The concept of linguistic frozenness discussed earlier appears to be the most suitable framework within which to examine this possibility. Recall that Fraser (1974) has categorized idioms into an inclusive hierarchy according to the type and order of transformational operations (with respect to standard transformational generative grammar) which idioms can undergo and still retain their idiomatic meaning. Thus, for example, Fraser has pointed out that phrasal idioms such as "kick the bucket" can undergo very simple transformational change, such as that involved in producing the gerundive ("John's kicking the bucket was tragic"), but will undergo few if any other changes. (For example, it fails to undergo simple passivization: "The bucket was kicked by John" does not retain the idiomatic meaning.) Idioms such as this are mildly frozen. Some idioms, such as "to let off some steam" are considered by Fraser to be completely frozen, while others such as "throw in the sponge" are virtually unfrozen—they can undergo most transformations.

It is worth noting here that no idioms can undergo all transformational operations and

retain their idiomatic meanings. Further, it should be noted that level of frozenness is completely independent of whether or not an acceptable literal reading exists for the idioms—both “grammatical” and “nongrammatical” idioms vary in their levels of frozenness.

It may be that the vagaries idioms display with respect to susceptibility to transformational change are, in fact, merely a result of operations which are allowable on similar, literal, phrases in the language; the fact that language has both properties of plasticity and creativity could account for such generalization in function. However, it may also be the case that the lexical status of idioms is a function of the categorization Fraser has noted. It may be, for example, that the more frozen an idiom is, the stronger its lexical status (in terms of, for example, access or retrieval routes). Overall, whatever the basis for this categorization, the questions of how or whether such derived classifications are reflected in the underlying representation and comprehension of idioms deserve investigation.

The data in Experiment 1 are taken on a set of idioms which represent a wide but uneven spread of frozenness and there are insufficient numbers of certain levels to allow for a reasonable test of the hypotheses just proposed. For this reason, a second experiment was performed, one which explicitly examines the frozenness categorization with respect to lexical status.

EXPERIMENT 2

Method

Subjects. Forty-two undergraduate students from Tufts University participated in this study. Each was paid \$1.50 for participation.

Materials. Twelve grammatical idiomatic phrases which were equibaised constituted the basic experimental materials for this study. (Determination of bias was based on material

pretests and criteria identical to those described in Experiment 1. Thirty-one Tufts University undergraduates participated in the pretest.) Three of the experimental idioms were completely frozen (F_0)—that is, they could undergo no transformation and retain their idiomatic meaning. Another three of these idioms constituted exemplars of the most mild level of frozenness (category 1 on Fraser’s hierarchy). Another three of the idioms constituted exemplars of an intermediate level of frozenness. The final three idioms were taken from Fraser’s fifth level of frozenness (F_5)—that of maximum transformational applicability. Appendix 2 displays the materials used in this experiment.

Each idiom was matched with a control word string which was created by exchanging either the first or the last word in the idiom with a word of greater or equal frequency and of equal length. The resulting word string control constituted a common and acceptable English phrase (refer to Experiment 1 for further details). In addition, 35 word strings which did not form acceptable English phrases were constructed.

Procedure. The procedure was identical to that reported in Experiment 1. Each subject was seated in a booth containing a visual display unit and a pair of response buttons. All idiom phrases, control phrases, and non-phrase materials were presented in random order to each subject. A PDP8/e minicomputer presented all materials with randomly varied intertrial intervals lasting from 0.5 to 2.0 seconds. Subjects were instructed to determine, as quickly as possible, whether each word string was a permissible English phrase or not and to press the appropriate button in front of him/her. Reaction times to these decisions were recorded by the PDP8/e computer.

Results

Individual means were calculated for each of the cells formed by the Phrase Type (Idiom vs Control) \times Frozenness Level variables.

TABLE 3
MEAN REACTION TIME (IN MILLISECONDS) FOR IDIOM AND CONTROL PHRASES AT FOUR
LEVELS OF TRANSFORMATIONAL FROZENNESS

Phrase type	Level of linguistic frozenness			
	Complete (F ₀)	Strong (F ₁)	Mild (F ₃)	Unfrozen (F ₅)
Idiom	960	991	957	940
Control	1086	1068	1043	1041

The mean reaction times for cells of this interaction, calculated across subjects, are given in Table 3. An analysis of variance was performed on these data with Phrase Type and Frozenness Level as main effects. The main effect for Phrase Type was significant, $F(1, 41) = 54.38$, $p < .0001$. This was also significant under *minF'* analysis, $minF'(1, 45) = 5.03$, $p < .05$. The main effects for Frozenness Level and the Frozenness Level \times Phrase Type interaction were not significant; $F(3, 123) = 2.56$ and 1.99, respectively.

Discussion of Experiment 2 and General Discussion

The results of Experiment 2 show no main effect for the variable of Frozenness. Similarly, this variable does not interact with the basic Idiom/Control Phrase Type effect found in Experiment 1 and replicated in this study. This does not mean that linguistic descriptions of frozenness have no perceptual validity, but rather that the susceptibility these idioms show to transformational change is a factor applicable to these items as *lexical items*. That is, acceptance of the Lexical Representation Hypothesis in no way contradicts the notion that restrictions may exist upon insertion or the use of certain idioms (lexical items) in the grammar. However, the results do suggest that idioms are to some extent unified with respect to access, retrieval, and representation in the lexicon.

In sum, then, it appears as though we have

achieved reasonably strong support for the Lexical Representation Hypothesis; idioms appear to be stored and accessed as lexical items, not from some special list that is distinct from the lexicon nor by a special processing mode which comes into play when literal analysis fails. Further, the categorization of linguistic frozenness, which has very real reflections in higher order linguistic organization, does not appear to be a basis for differential access of idioms in the lexicon. It should be noted, however, that the categorization provided by Fraser (1974) which we tested is not the only one available, and there is certainly a lot of mushy theoretical ground around the concepts of "loosely fixed" (unfrozen) phrases.

It appears worthwhile to reconsider, at least briefly, the previous studies on idioms in order to place the current results in perspective. Both the results of Bobrow and Bell (1973) and of Brannon (Note 2) provided evidence which they took as support for the Idiom List Hypothesis, evidence based in the one case on post-perceptual reports of "first meaning perceived in context," and in the other on "speed to compare meanings" (again, a post-perceptual measure). Both results, in fact, make a great deal of sense. However, they do not appear to reflect perceptual processing of idiomatic phrases. The Bobrow and Bell task, as suggested in the introduction, appears to reflect the results of a mental decision as to the most appropriate meaning to be brought to conscious aware-

ness. Similarly, the Brannon experiment required subjects to compare meanings of two idiom strings and of two nonidiom strings. If lexical access is immediately undertaken for a grammatical idiom, and literal analysis is also begun, there would simply be more meanings to compare for two idiom strings than there are for two literal control strings. This situation could account for her finding of longer decision times for idiom comparisons. This suggestion is supported by other data Brannon reports. In a second experiment, she found that it took longer to compare ambiguous nonidioms than unambiguous nonidioms. Thus, in this task, a comparison involving many meanings is shown to take longer than comparisons involving only a few meanings. In all, it appears that post-perceptual tasks are often simply not appropriate for providing evidence about perceptual access; a task which more immediately reflects the processes under question, such as that used in the present paper, often finds results which are undetected by later-occurring tasks.

The hypothesis that idioms are represented lexically has had recent, although somewhat indirect, support. Ortony, Schallert, Reynolds, and Antos (1978) report that comprehension of idiomatic phrases in their idiom sense took no longer (and in fact was often faster) than in their literal sense, a result they took to suggest that idioms and metaphors require no *special* processing for comprehension. Certainly these results fit well with a Lexical Representation Hypothesis.

It is worth noting that the analysis we have applied to idiom processing is the same analysis that we would presumably wish to apply to certain other complicated language forms. For example, one might well expect comprehension of nominal compounds (shrimpboat, hotdog) to occur through lexical representation and access rather than through access from a special "nominal compound list" or from some attempted "nonliteral"

analysis. As a great number of nominal compounds are not "literal" in the sense of their meaning being a simple combination of the literal meanings of their component morphemes, they present a problem similar to that provided by idioms for the comprehension device.

In addition, our lexical representation hypothesis fits well with some relatively recent data on metaphor processing. Harris (1976) has reported finding no difference between latency to initiate paraphrase of metaphoric quotations and equivalent non-metaphoric quotations. He interprets these results as support for a model in which (non-literal) metaphoric processing takes place in parallel with attempted literal processing for metaphor strings. This, then, fits with the hypothesis we have supported—that multiple computations (both lexical and structural) can be simultaneously entertained by the comprehension device in processing idioms.

In all, it appears that idioms, like many other problems which have been relegated to some sort of "special" process, may turn out to have very "normal" processing solutions, at least once we elaborate more fully what the normal system can do.

APPENDIX 1: IDIOM/CONTROL MATERIALS USED IN EXPERIMENT 1

break the ice/break the cup
 pain in the neck/pain in the foot
 see the point/see the group
 out of line/out of food
 lost his marbles/lost his fortune
 on the wagon/on the train
 out of sight/out of shape
 take him for a ride/take him for a beer
 over the hill/over the edge
 in the dark/in the east
 kick the bucket/lift the bucket
 held up/open up
 spill the beans/crate the beans
 see the light/get the light
 put him on/let him on
 cut it out/try it out
 strike out/throw out
 hold on/move on
 pull his leg/lose his leg

bail out/bang out
 drop it/save it
 wrap it up/lift it up

APPENDIX 2: IDIOM/CONTROL MATERIALS FOR EXPERIMENT 2

Frozensness Level (Permissible Operations Cited in Parentheses)

- F₀ (can undergo no operations)
 jump in the lake/dive in the lake
 shake a leg/twist a leg
 fly a kite/fix a kite
- F₁ (adjunction only)
 climb the walls/paint the walls
 kick the bucket/fill the bucket
 chew the fat/melt the fat
- F₂₋₄ (insertion, permutation, extraction)
 break the ice/break the cup
 give her a ring/give her a poem
 see the light/see the group
- F₅ (can undergo most operations)
 bury the hatchet/bury the coffins
 see the point/see the water
 spill the beans/spill the juice

REFERENCES

- BOBROW, S., & BELL, S. On catching on to idiomatic expressions. *Memory and Cognition*, 1973, **1**, 343-346.
- CHAFE, W. Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language*, 1968, **4**, 109-125.
- CLARK, H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335-359.
- CLARK, H., & LUCY, P. Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 56-72.
- CUTLER, A., & NORRIS, D. Monitoring sentence comprehension. W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.
- FRASER, B. *The verb-particle combination in English*. Taishukan, Tokyo, 1974.
- FOSS, D. J., & JENKINS, C. Some effects of context on the comprehension of ambiguous sentence. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 577-589.
- FOSS, D. J., & SWINNEY, D. On the psychological reality of the phoneme: perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 246-257.
- GARRETT, M. Does ambiguity complicate the perception of sentences? In G. B. D'Arcais & W. J. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland, 1970.
- HARRIS, R. Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of the Psychonomic Society*, 1976, **8**, 312-314.
- HERINGER, J. Idioms and lexicalization in English. In J. P. Kimball (Ed.), *Syntax and semantics*, Vol. 9: *The grammar of causative constructions*. New York: Academic Press, 1976.
- KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English*. Providence, R.I.: Brown Univ. Press, 1967.
- LACKNER, J., & GARRETT, M. Resolving ambiguity: Effect of biasing context in the unattended ear. *Cognition*, 1972, **1**, 359-372.
- MARSLER-WILSON, W. D. Linguistic descriptions and psychological assumptions in the study of sentence perception. In R. J. Wales & E. C. T. Walker (Eds.), *New Approaches to language mechanisms*. Amsterdam: North-Holland, 1976.
- MEHLER, J., SEGUI, J., & CAREY, P. Tails of words: Monitoring ambiguity. *Journal of Verbal Learning and Verbal Behavior*, 1978, **1**, 29-37.
- NEWMeyer, F. The insertion of idioms. In *Papers from the Eighth Regional Meeting*. Chicago: Chicago Linguistic Society, 1972. Pp. 294-302.
- ONIFER, W., HIRSHKOWITZ, M., & SWINNEY, D. A mini-processor PDP8/e-based system for investigations of on-line language processing: Automated program for psycholinguistic experiments (APPLE). *Behavior Research Methods & Instrumentation*, 1978, **10**, 307-308.
- ORTONY, A., SCHALLERT, D., REYNOLDS, R., & ANTOS, S. Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 465-478.
- TAFT, M., & FORSTER, K. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 638-647.
- WEINREICH, U. Problems in the analysis of idioms. In J. Puhvel (Ed.), *Proceedings of the Summer 1966 Linguistics Forum at UCLA*. Berkeley: University of California Press, 1967. Vol. 11, pp. 23-81.

REFERENCE NOTES

1. SWINNEY, D. *Does context direct lexical access?* Paper presented at Midwestern Psychological Association, Chicago, Illinois, 1976.
2. BRANNON, L. *On the understanding of idiomatic expressions*. Unpublished dissertation, University of Texas, 1975.

(Received December 28, 1978)