Comprehending Idioms Cross-Linguistically

Heather Bortfeld

Department of Psychology, Texas A & M University, College Station, TX, USA

Abstract. Speakers of three different languages (English, Latvian, and Mandarin) rated sets of idioms from their language for the analyzability of the relationship between each phrase's literal and figurative meaning. For each language, subsets of idioms were selected based on these ratings. Latvian and Mandarin idioms were literally translated into English. Across three experiments, people classified idioms from the three languages according to their figurative meanings. Response times and error rates indicate that participants were able to interpret unfamiliar (e.g., other languages') idioms depending largely on the degree to which they were analyzable, and that different forms of processing were used both within and between languages depending on this analyzability. Results support arguments for a continuum of analyzability (Bortfeld & McGlone, 2001), along which figurative speech ranges from reflecting general conceptual structures to specific cultural and historical references.

Key words: figurative language, idiom processing, cross-linguistic research, conceptual structure, metaphor, bilingualism

Recently, a German friend remarked that our university's new president "had set the accent" with his changes about campus. Despite being unfamiliar with the phrase, I understood the meaning intended by its use. As a direct translation of the German idiom *Akzente setzen*, "to set the accent" means to set a course or to indicate direction. Unfortunately, the figurative meanings of other languages' idioms are not always so easily understood. The same friend once commented that a colleague drove "with a monkey's tooth" (from the German *mit einem Affenzahn*), intending for me to understand that the person drove extremely fast. As a native English speaker with only rudimentary knowledge of German and limited knowledge of German culture and history, I was stumped. Why is the first idiom understandable when literally translated into English while the second is not?

Idioms are fixed phrases that mean something other than what a literal interpretation of their individual words would indicate (Fraser, 1970; Katz & Postal, 1963; Weinreich, 1969). For example, when the phrase *spill the beans* is interpreted figuratively in American English, it means to reveal a secret or divulge information; it does not mean to inadvertently pour legumes on the floor. Despite what appears to be a significant difference between the stated and intended meanings, adult native English speakers have little difficulty understanding this phrase when used to indicate either sense.

Idioms are pervasive. In their original (1975) *Dictionary of American Idioms*, Boatner, Gates, and Makkai presented over 4,000 expressions occurring in English alone. A more recent edition of this same dictionary (Makkai, Boatner, & Gates, 1995) contains over 8,000 English language idioms. This is not to say that idioms have increased in number in recent years; rather, it shows that people are increasingly willing to acknowledge these phrases for the major role they play in daily language use. Yet despite idioms' ubiquitous nature, early models of language comprehension were based on only the most literal language. As a result, these models had difficulty accounting for how people are able to understand figurative language. Since then, idiom processing has itself inspired much research (e.g., Cacciari & Tabossi, 1988; Gibbs, 1992; Glucksberg, Brown, & McGlone, 1993; Swinney & Cutler, 1979). And because metaphor is considered the basis for much of figurative language, the debate regarding the nature of idiom representation, storage, and access reflects

DOI: 10.1027//1618-3169.50.3.217

specific disagreements about the nature of metaphor comprehension.

Metaphor Comprehension

Many models have been developed in an attempt to explain how people understand idioms and, by extension, metaphors. One possibility is that metaphorical structures established through prior experience guide people's comprehension of idiomatic language (e.g., Gibbs, 1992; Lakoff, 1987, 1990). According to this view, many idioms are motivated by people's implicit knowledge – acquired over time – of the "conceptual metaphors" that underlie the phrases' figurative meanings (discussed by Lakoff, 1987, 1990). An example of an underlying conceptual metaphor is: ANGER IS A HEATED SUBSTANCE (from Lakoff, 1987), which may be expressed in English with phrases such as *blow your stack* and *lose your cool*. Gibbs (1992) argued that the figurative meanings of such idioms are at least in part motivated by the metaphorical mappings between source and target domains that constitute people's conceptual knowledge.

Another hypothesis worth considering is Jackendoff's *thematic parallelism* argument (Jackendoff, 1983; Jackendoff & Aaron, 1991), which holds that the use of certain words, phrases, and other grammatical constructions to describe propositions in two domains (e.g., "spill" to describe both physical and communicative dispersal) arises from the parallel conceptual structures of these domains. According to this view, idioms do not reflect metaphorical correspondences, but rather the parallelisms between the domains. In the same vein, the *structural similarity* hypothesis (Murphy, 1996, 1997) posits chains of similarity that can motivate extension of terms from their literal to figurative senses.

Alternatively, processing algorithms or conceptual schema may be constructed on-line and made available for access during idiom comprehension (Gentner & Wolff, 1997; Glucksberg et al., 1993). Although models within this class differ substantially from one another, they all generally take issue with the notion that pre-existing structures guide idiom comprehension as it occurs on-line. The predominant argument stemming from this work is that conceptual structures are constructed anew and made accessible in working memory during the comprehension process (e.g., Glucksberg & Keysar, 1990; Glucksberg, Keysar, & McGlone, 1992; Glucksberg et al., 1993). Specifically, the *attributive categorization* view holds that people infer a category based on its vehicle (where *b* is the vehicle of implicit similes such as *a is like b*) and apply the properties of that category to a topic (where *a* is the topic of *a is like b*). In this way, a new (attributive) category is created, allowing the topic and vehicle (for example, of an idiom) to be combined in a metaphorical relation.

These views predict different forms of processing for metaphor – and idiom – comprehension. Most notably, one set implies that the mappings between conceptual structures and figurative language are well-established and that these mappings guide the comprehension process, while the other set suggests that mapping takes place as speech is unfolding. While both sets of views describe how lexical form might get linked to conceptual structure, what neither set addresses is how processing might differ for well-established, familiar phrases relative to those that are heard for the first time. Idioms are demonstrably variable – both within and between languages. What does this variability mean for the type of processing used to comprehend different phrases at different stages of familiarity?

One way to view how we deal with this variability is in terms of what is salient at the time. Giora's *graded salience* hypothesis (1997, 2002) claims that figurative and literal language use is governed by a general principle of salience. Salient meanings (e.g., conventional, frequent, familiar, predictable, enhanced by prior context) are processed first, and parallel processing is induced when more than one meaning is salient (e.g., Blasko & Connine, 1993; Gibbs, 1980, 1985; Giora, 1997). This hypothesis suggests that the standard processing model for figurative language is inadequate and should be revised such that salience (defined as encompassing conventional, familiar, and frequent interpretations) guides processing rather than the traditional literal-figurative dichotomy. Consequently, it would not be the figurative versus literal split, but a salient versus nonsalient continuum that guides interpretation. The graded salience argument provides a working description for why certain interpretations may take precedence over others and it is consistent with research showing the powerful influence that characteristics like discourse context have on language comprehension (e.g., Gerrig & Bortfeld, 1999). However, this view does not entirely account for the variability inherent in specific classes of figurative language, as is the case with idioms.

Idiom Analyzability

Linguists initially characterized idioms as noncompositional (e.g., Weinreich, 1969; Katz & Postal, 1964), meaning that a phrase's individual words do not contribute to the figurative meaning implied when those words are put together in an idiomatic phrase. Nunberg, Sag, and Wasow (1994) later provided a linguistic analysis demonstrating that idioms

are, in fact, compositional; subsequent psycholinguistic analyses provided further support for this argument (e.g., Gibbs, 1985, 1993). Gibbs and Nayak (1989) demonstrated that idioms can be divided into three different groups, or types, based on how decomposable – or *analyzable* they are. According to this view, there is a relatively transparent relationship between the surface structure of some idioms and their figurative meanings (e.g., *lose your temper*). For these "normally decomposable" phrases, the literal and figurative meanings are semantically similar enough that very little analysis needs to take place for the figurative meaning to be understood based only on the literal form of the phrase. The relationship is a bit more distant between the literal and figurative meanings of other idioms (e.g., *flip your lid*). According to Gibbs and Nayak, these "abnormally decomposable" phrases require more analysis, since the metaphorical links between the literal meanings and figurative concepts need to be mapped out in order for the phrases to make sense. Still other idioms (e.g., *kick the bucket*) are those whose surface structure has little relation to the intended figurative meaning. The figurative meanings of these "nondecomposable" phrases are often based on some historic occurrence or culturally instantiated linguistic usage that has long since been forgotten, while the phrases themselves have become fixed in the language as single units.

In support of this analysis, Gibbs, Nayak, and Cutting (1989) found that processing times are different for the different classes of idioms. Notably, they found that people judged abnormally decomposable idioms to be meaningful *faster* than they did normally decomposable idioms. Though this difference may seem counterintuitive, the researchers argued that it is attributable to the fact that certain experiences or events (e.g., getting angry) are conceptualized in a manner consistent with the entailments of specific conceptual metaphors (e.g., anger = "a heated substance under pressure") and that processing the surface forms of these metaphors (e.g., *flip your lid*) is speeded relative to more literally stated representations of the events (e.g., *lose your temper*). Gibbs and his colleagues (1989) observed that people judged both types of decomposable idioms faster than they did nondecomposable idioms. Based on these findings, they concluded that the analyzability of an idiom is a matter of degree.

Undoubtedly, cultural knowledge plays a role in understanding how figuratively intended concepts are lexicalized in any given language (Kecskes & Papp, 2000). For example, where English speakers say *spill the beans* when they mean that information has been revealed, Latvian speakers say *spill [the] water*. Interestingly, McGlone, Glucksberg, and Cacciari (1994) used this idiom as an example of why speakers would not use a mass noun to represent something conceptualized as a count noun. But clearly this is not the case for speakers of Latvian, highlighting how experience with a specific language can bias one's views about idiom analyzability in general. In fact, Latvian speakers argue that the use of *spill the beans* in English is strange (to them, at least) precisely because beans can be picked up – thus reversing the outcome of being spilled – where the action of spilling water is impossible to reverse, just as one cannot reverse the outcome of revealing information (Bortfeld, 1998, 2002).

As this example illustrates, whether and to what degree an idiom is analyzable is a fairly subjective measure and is open to considerable debate. Native speakers of any language will have pre-established biases stemming from the phrases they have analyzed (or been unable to analyze) from their own language. They also know what many of the idioms from that language mean, regardless of whether they understand why this is so. While some researchers pursue the important work of differentiating characteristics of idioms within a single language (e.g., Blasko, 1999; Blasko & Connine, 1993; Titone & Connine, 1994), examining how people make sense of idioms from *other* languages (e.g., unfamiliar idioms) will help clarify the degree to which idioms are variably analyzable and how processing is influenced by familiarity with a phrase's surface form. This may also clarify why some data are consistent with an on-line account of figurative language comprehension, while other data support an account focused on pre-existing structures.

Cross-Linguistic Idiom Comprehension

Given that idioms are used in virtually every language and because surface representations of similar concepts differ across languages, these phrases pose a particularly difficult problem for language learners. Even native speakers can be convinced that the figurative meanings of certain idiomatic phrases in their language mean the opposite of what they actually mean (e.g., Keysar & Bly, 1995), indicating that analysis would not lead them to think otherwise. Nonetheless, fluency in a language is often inferred based on how flexibly a speaker understands and produces figurative speech (Bortfeld & Brennan, 1997). Just as idioms from one's own language are variably analyzable, varying degrees of analyzability must be a characteristic of idioms from other languages as well. Following this logic, normally analyzable (e.g., *to lift into the light*) and abnormally analyzable (e.g., *to release into the wind*) idioms should make sense

across languages despite the fact that the phrases may be unfamiliar units from language to language. Unanalyzable idioms (e.g., *to scatter ducklings*) should present more of a challenge. Given the variability in the analyzability of their surface structure, other languages' idioms can serve as a tool to address a confound common in much idiom processing research: that any given language's idioms are familiar to native speakers of that language and hence, will already have been analyzed to the degree that they can be analyzed by those speakers. By introducing uniformly unfamiliar phrases that are not uniformly analyzable, any differences in the processing required to make sense of the phrases can be measured.

Although examining knowledge to make sense of figurative speech that varies in familiarity and analyzability, cross-linguistic idiom comprehension is a phenomenon that has received little attention in the traditional processing literature (but see Danesi, 1993; Irujo, 1986, 1993; Kecskes, 2001, 2002; Kövecses & Szabó, 1996 for insightful work on idiom learning, and see Matlock & Heredia, 2002 for work on the acquisition of other fixed phrases). The experiments reported here examine whether and to what degree people can identify the figurative meanings underlying a variety of idioms. Difference between how people process idioms from their own and from two other languages will indicate whether idioms are analyzable to *varying degrees* across languages, as well as how processing might differ for the different phrases *given this variability.*

Experiment 1

In an extension of Gibbs et al.'s (1989) findings, participants in Experiment 1 were presented with idiomatic phrases that were either normally analyzable, abnormally analyzable, or unanalyzable. Their task was to read each visually presented idiom and decide which meaning category it belonged to given its figurative interpretation. The phrases belonged in equal parts to the meaning categories: revelation, secrecy, insanity, anger, and control. Phrases had been previously selected based on the analyzability ratings of independent judges, such that a third of the phrases' figurative and literal meanings had a transparent relationship (normally analyzable idioms); another third had a metaphorical relationship (abnormally analyzable idioms); the final third had an opaque relationship (unanalyzable idioms). Based on Gibbs, Nayak, and Cutting's (1989) earlier findings, it is expected that participants in this experiment should correctly categorize abnormally analyzable idioms from their native language the fastest, followed by normally analyzable idioms, while they should correctly categorize unanalyzable idioms the slowest. This result would confirm that the analyzability of an idiom 1) is a matter of degree and 2) influences the type of processing used to make sense of it in a nonlinear manner.

The three idiom types Gibbs, Nayak, and Cutting (1989) used in their processing studies were culled from Gibbs and Nayak (1989), in which idioms were grouped as one of three types (normally decomposable, abnormally decomposable, or nondecomposable). Based on the idea that idiom analyzability ranges along a continuum, with normally analyzable idioms anchoring one end, unanalyzable idioms anchoring the other, and abnormally analyzable idioms occupying the middle, it should be possible to identify idioms that are located on either end of the analyzability continuum, as well as in the middle. These three locations on the analyzability continuum should mirror the three idiom types proposed by Gibbs and Nayak (1989). To capture these types for the current study, an idiom rating technique was developed that identified those idioms clearly at either end of the analyzability continuum or directly in the middle. This technique served two purposes: it allowed idioms to be grouped as very clearly normally analyzable, abnormally analyzable, or unanalyzable and it underscored the fact that idioms fall along a continuum of analyzability. This idiom rating technique (outlined below) guided stimuli selection for each of the three experiments reported here.

Method

Participants

A group of 34 undergraduate students (15 males and 19 females) received research credit for their participation. All were native speakers of American English and none reported that they were functionally fluent in another language.

Material

Using idiom dictionaries (e.g., Boatner, Gates, & Makkai, 1975; Long & Summers, 1979; Makkai, Boatner, & Gates, 1995), a corpus of 150 idiomatic English phrases was compiled representing five concept groups (30 idioms per group). Each idiom was then paired with its figurative paraphrase, as defined by English idiom dictionaries (e.g., Long & Summers, 1979; Makkai, Boatner, & Gates, 1995).

Two native English speakers, both graduate students in English literature, rated these 150 idioms. Each was paid 50 dollars. Both raters were naive to

the experimental hypothesis being tested. Raters were presented with a booklet containing a written set of instructions, a set of 25 practice idiom-paraphrase pairs and the 150 randomly ordered idiom-paraphrase pairs relevant to the current study. Every idiom-paraphrase pair was accompanied by a 5-point Likert scale anchored with 1 = "transparent" and 5 = "opaque." Raters were instructed to locate each idiom along its accompanying 5-point scale, being sure to use all 5 points. Every idiom-paraphrase pair was followed by a question regarding the familiarity of that idiom ("Is this idiom familiar or unfamiliar to you?"). Raters were instructed to answer the familiarity question for each idiom based on whether they had previously used or witnessed someone else using the idiom in either conversation or writing. After reading the instructions, the raters rated the 25 practice idioms. Following the practice session, raters discussed any confusions or clarification questions they had about the rating process with the experimenter. When there were no more clarification questions, they rated the 150 idioms.

Transparent–opaque distinction

Instructions to raters (see Appendix A) paraphrased and extended the original instructions of Gibbs and Nayak (1989) to their participants. Raters were instructed to base an idiom's rating on how many words in the phrase they considered to have transparent (or seemingly literal) relations to their figurative referents ("1" on the scale) or, conversely, opaque (or arbitrary) relations to their figurative referents ("5" on the scale). Idioms whose individual components had transparent relations to their idiomatic referents were to be rated at the low end of the scale (toward 1), those whose components had opaque relations to their idiomatic referents were at the high end (toward 5), and those whose components had metaphorical relations to their idiomatic referents were in the middle (on or near 3).

Further differentiation was made by examining combinations of the individual words. If only one word in a phrase was metaphorically related to its figurative referent, while all other words were literally related to their figurative referents, then this idiom was given a rating along the lower half of the scale (2). If an idiom contained both a metaphorically related word and a seemingly arbitrary word, this phrase was rated further up the scale (4).

Raters worked independently, rating each of the 150 idioms and indicating whether or not each phrase was familiar. A reliability measure indicated satisfactory agreement between raters (88.6% agreement, Cohen's Kappa = .87). Ratings were averaged and distributions of the means were compiled for each concept group in order to determine which 15 idioms from each concept category were consistently rated such that five fell on one end of the scale (on or near 1), five fell on the other end (on or near 5), and five in the middle (on or near 3). The five idioms consistently rated as transparent for the five different concept groups composed the normally analyzable idioms for use in the subsequent categorization study. The five idioms consistently rated as metaphorical for the five different concept groups composed the abnormally analyzable idioms. The five idioms consistently rated as opaque for the five different concept groups composed the unanalyzable idioms. Idioms roughly matched for length across the three types, such that normally analyzable idioms averaged 92 characters (SD = 7.4), abnormally analyzable idioms averaged 93 (SD = 2.5) characters, and unanalyzable idioms averaged 94.4 (SD = 9.2) characters. Only idioms familiar to both raters were included in the final set of stimuli (see Appendix B for the full set of idioms for the concept *insanity*).

Procedure

Each person completed the categorization experiment individually. Idiomatic phrases were presented on a computer screen one at a time in a different random order for each participant. The computer's keyboard was covered such that only 5 keys were exposed. These keys were labeled with capital letters (R, I, C, A, or S) representing one of each of the five categories (revelation, insanity, control, anger, or secretiveness). The full name of each category was located above the appropriate key with the first letter of each category name so that the match between each key and its corresponding category was clear.

Participants were asked to press a key for each idiom according to which concept group they felt best represented that idiom's figurative meaning. Prior to the experiment, they read instructions regarding the procedure and completed a brief practice session that included one idiom from each of the five concept categories. This session was intended to familiarize participants with the task and the data were not included in the final analysis. Participants took approximately 15 minutes to complete the experiment. Response latencies were measured from the onset of the word string to the time at which participants pressed one of the five buttons.

Analyses

For all experiments reported here, two analyses were conducted for each result; F_1 is the analysis by-sub-

Table 1. Native English Speakers' Response Times (in Seconds) for Correct Conceptual Categorization of Three Idiom Types for English (Experiment 1), Latvian (Experiment 2), and Mandarin (Experiment 3).

Idiom Type	Source Language for Idiomatic Stimuli		
	English	Latvian	Mandarin
Normally Analyzable	3.684 (1.2) (89%)	3.498 (.81) (79%)	4.018 (.86) (82%)
Abnormally Analyzable	3.301 (.88) (76%)	3.852 (1.1) (69%)	5.123 (1.1) (59%)
Unanalyzable	3.939 (1.3) (68%)	4.850 (.82) (44%)	5.788 (1.4) (50%)

Note. Standard deviations are listed next to the corresponding response time. Overall percentage correct is listed below the corresponding response time.

jects (with participants as the random factor), and F_2 is the analysis by-items. Where appropriate, planned comparisons of the effects of the three conditions were conducted, comparing the three idiom types (that is, the normally analyzable versus the abnormally analyzable idiom groups, and both of these versus the unanalyzable idiom group). All analyses are based on correct response averaged across the five concept groups.

Results and Discussion

The response times for conceptual categorization are presented in Table 1, along with the proportion of idiom types (e.g., normally, abnormally, or unanalyzable) that were correctly categorized (that is, categorized in a manner consistent with the figurative definitions provided in standard American English idiom dictionaries). Response times from incorrectly categorized phrases were eliminated from the analysis, as were response times greater than 3 *SD* beyond each individual's mean response time.

First, the percentage of correctly categorized idioms decreased across the three idiom types, such that normally analyzable idioms were most often categorized correctly (89%), followed by abnormally analyzable idioms (76%), and unanalyzable idioms (68%), linear trend, $F_1(1, 99) = 56.09$, $p < .001$; $F_2(1, 72) = 13.75$, $p < .001$.

Differences in the analyzability of the idiom types were reflected in participants' response times as well. An analysis of variance was performed with idiom type (normally vs. abnormally vs. unanalyzable) as a within-subject factor, revealing a significant effect, $F_1(2, 66) = 17.34$, $p < .001$; $F_2(2, 73) = 3.07$, $p = .05$. Participants took a longer time to correctly categorize the unanalyzable idioms based on their conceptual meanings than they did for either the normally or abnormally analyzable idioms. Participants correctly categorized the figurative meanings of both normally and abnormally analyzable idioms (3.685 s and 3.301 s, respectively) faster than they did for unanalyzable idioms (3.939 s), planned comparison, $F_1(1, 66) = 22.84$, $p < .001$; $F_2(1, 72) = 8.88$, $p < .01$. Finally, participants correctly categorized the meanings of abnormally analyzable idioms (3.301 s) faster than they did for normally analyzable idioms (3.685 s), planned comparison, $F_1(1, 66) = 12.10$, $p < .001$; $F_2(1, 72) = 3.65$, $p = .06$, although the analysis by items reached only marginal significance. These results are consistent with the earlier findings of Gibbs, Nayak, and Cutting's (1989).

The differences in response times for the three groups of idioms indicate that there are differences in how people process the different phrases depending on how analyzable they are. Furthermore, at least when the idioms are familiar (e.g., from one's own language), this difference goes in the counterintuitive direction observed in previous research (e.g., Gibbs, Nayak, & Cutting, 1989; Nunberg, 1978). These data indicate that the speed does, indeed, depend on the degree to which the phrases are analyzable. But to what degree did participants' pre-established knowledge of the figurative meanings of the idiomatic phrases influence their ability to perform this task? Experiment 2 was designed to address this question.

Experiment 2

Experiment 2 measured the speed and accuracy with which native English speakers were able to categorize idioms translated from another language. Toward this end, participants were asked to complete the same conceptual categorization task used in Experiment 1, but this time they categorized 75 literally translated Latvian idioms.

Method

Materials

As in Experiment 1, the first half of Experiment 2 established which Latvian idioms would be included as stimuli in the subsequent categorization study. A series of Latvian idioms were submitted to the same rating task as were the English idioms in Experiment 1. Raters for Experiment 2 were two philology graduate students at the University of Latvia. Each was paid the equivalent of fifty dollars in exchange for completing the rating task. Each was naive to the experimental hypothesis being tested. The same guidelines were followed to train the Latvian speaking raters that had been used to train the English speaking raters. Although both were able to speak English to some degree, neither had been exposed to English before their teenage years and both had started studying English relatively late (during secondary school).

A list of 150 Latvian idioms was compiled with 30 idioms for each of the five concepts, as described in Experiment 1. Raters were presented with a booklet containing a written set of instructions, 25 practice idiom-paraphrase pairs, and the 150 Latvian idioms paired with paraphrases of their figurative meanings as defined by Latvian idiom dictionaries. The English instructions (see Appendix A) were translated into Latvian by a bilingual (Latvian-English) speaker who also was a doctoral candidate in linguistics. The two raters independently rated all of the idioms and indicated whether or not they were familiar with each phrase. A reliability measure indicated satisfactory agreement between them (86.7% agreement, Cohen's Kappa = .85). Their ratings were averaged and a distribution of those means was compiled.

Once it was determined which 75 idioms were to be used in the categorization study in the manner followed in Experiment 1, another language sensitive Latvian-English bilingual translated each idiom literally (word for word) into English (Appendix C). These were reviewed by the first Latvian-English bilingual, and any ambiguous or problematic translations were discussed until both speakers agreed upon acceptable translations. Idioms roughly matched for length across the three types, such that normally analyzable idioms averaged 93.2 characters (*SD* = 13.6); abnormally analyzable idioms averaged 92.6 (*SD* = 11.6); unanalyzable idioms averaged 93 (*SD* = 5.1).

Participants

A different group of 34 undergraduate students (16 female and 18 males) received research credit for their participation in the experiment. All were native speakers of American English and none reported that they were functionally fluent in another language.

Procedure

The procedure was identical to that followed in Experiment 1, except that participants were told that the idioms they would be categorizing were going to be unfamiliar to them and that their job was to try to determine the figurative meaning of each phrase. Prior to the experiment, participants read instructions regarding the conceptual categorization task that included relevant examples of idioms that were literally translated from Latvian. As in Experiment 1, participants completed a brief practice session that included five literally translated Latvian idioms, with one from each of the five concept categories. This session familiarized participants with the task and the data were not included in the final analysis. Participants took approximately 20 minutes to complete the experiment.

Results and Discussion

Response times for conceptual categorization are presented in Table 1, along with the proportion of idiom types (normally, abnormally, or unanalyzable) that were correctly categorized. Response times from incorrectly categorized phrases were eliminated from the analysis, as were response times greater than 3 *SD* beyond each individual's mean response time.

Error rates for native English speakers' conceptual categorization of Latvian idioms directly translated into English show a pattern similar to that seen for English idioms in Experiment 1, although there was an overall reduction in the number of idioms categorized correctly. Normally analyzable idioms again were most often correctly categorized (79%), followed by abnormally analyzable idioms (69%), and unanalyzable idioms (44%), linear trend, $F_1(1, 99) = 238.69$, $p < .001$; $F_2(1, 72) = 24.71$, $p < .001$. The most noticeable drop in English speakers' correct conceptual categorization of Latvian idioms is in the unanalyzable group. This is consistent with the characterization of unanalyzable idioms as having extreme historical/cultural bases that seem opaque to individuals not privy to the language in which these idioms are used. In fact, it is notable that native English speakers were able to correctly categorize over 40% of the so-called unanalyzable idioms, particularly since there were five possibilities from which to choose, meaning that chance performance would be a correct categorization rate of 20%.

However, where the accuracy with which native English speakers were able to correctly conceptually categorize Latvian idioms followed a pattern consistent with that seen for English idioms, their processing times followed a different pattern. An analysis of variance was performed with idiom type (normally vs. abnormally vs. unanalyzable) as a within-subject factor. These analyses revealed a significant effect of idiom type, $F_1(2, 66) = 39.93$, $p < .001$; $F_2(2, 72) = 12.46$, $p < .001$. Further analysis of the individual means indicated that participants were able to correctly categorize normally analyzable idioms significantly faster (3.498 s) than abnormally analyzable idioms (3.852 s), planned comparison, $F_1(2, 66) = 11.92$, $p < .001$; $F_2(2, 74) = 6.80$, $p < .01$. This is in contrast to the response times seen for English speakers' categorization of English idioms in Experiment 1 and follows a more intuitive pattern than that seen for English idioms, given that abnormally analyzable idioms are further down the analyzability continuum than are normally analyzable idioms.

This difference in response time patterns for English speakers categorizing English idioms and for English speakers categorizing literally translated Latvian idioms highlights the different processing used to comprehend abnormally analyzable idioms that are highly familiar (e.g., from one's native language) and those seen or heard for the first time (e.g., from another language). It appears that since people have already done the conceptual mappings for abnormally analyzable idioms from their native language, they are fast to process these phrases. These conceptual mappings may be accessed as single units, such that the processing of the phrases is faster even than for the more literal (normally analyzable) phrases, which still must be processed word by word. That is, the conceptual mappings for familiar abnormally analyzable idioms are readily accessed. In contrast, although native English speakers may find Latvian abnormally analyzable idioms apt (e.g., can analyze them for their figurative meaning), they take measurable time to do so. Results from Experiment 2 indicate that native English speakers had to – for the first time – figure out the conceptual mappings for these novel phrases. Their extended response times for categorizing such phrases reflect this mapping process. Finally, when participants were able to correctly categorize unanalyzable Latvian idioms that had been literally translated into English, they took significantly longer to do so than for either of the other two idiom types (4.850 vs. 3.675 s, planned comparison, $F_1(2, 44) = 11.92$, $p < .001$; $F_2(2, 74) = 6.80$, $p < .01$). This finding is consistent with results for unanalyzable English idioms, since the phrases are not readily mappable to any underlying conceptual structure. It appears that, although these phrases may come to be accessed as single units (e.g. "frozen" phrases) over time, no corresponding conceptual structure is available to speed this process. The processing of such phrases therefore remains slow relative to more analyzable phrases.

Experiment 3

Findings from Experiment 2 indicate that native English speakers are able to comprehend many literally translated Latvian idioms. However, this cross-linguistic analyzability may have emerged in part from historical similarities existing between English and Latvian, since both are Indo-European languages (though they are only distantly related within the Indo-European family). There may likewise be cultural factors at play that have shaped the idioms used in the two languages. Experiment 3 was designed to address this possibility. Using the same paradigm described previously, native English speakers in Experiment 3 conceptually categorized literally translated Mandarin Chinese idioms. Since Mandarin is a Sino-Tibetan language, this change in language addressed any confound that the genetic similarities between Latvian and English may have introduced to influence whether native English speakers were able to determine the figurative meanings of idioms directly translated from Latvian.

As in the previous experiments, Experiment 3 measured the speed and accuracy with which native English speakers could conceptually categorize 75 Mandarin idioms based on their intuitions about each idiom's figurative meaning.

Method

Participants

A group of 34 undergraduate students (14 males and 20 females) received research credit for their participation in the experiment. All were native speakers of American English and none reported that they were functionally fluent in another language.

Materials

Similar to the technique described for selection of Latvian idioms in Experiment 2, a list of 150 Mandarin idioms was compiled by native Mandarin speaking research assistants. Again, the 150 idioms contained 30 phrases for each of the five concepts outlined in Experiments 1 and 2. These 150 idioms, along with their accompanying figurative paraphrases, were included in a rating questionnaire iden-

tical to those used for the English and Latvian idioms. The rating questionnaire was written in Mandarin characters using a commercially available software package.

The two native Mandarin speaking raters were both graduate students in Linguistics who had received their primary and secondary education in Mainland China and who had arrived in the United States to begin their graduate studies within the year that this study was conducted. Although both were able to speak English to varying degrees, neither had been raised in an English speaking home and both had begun studying English relatively late (during secondary school) in their academic careers. Both were naive to the experimental hypothesis being tested. Each was paid fifty dollars. The same rating guidelines were followed as those followed to train raters in the previous experiments. Instructions for the rating task were translated by a Mandarin-English bilingual who was a doctoral candidate in psychology.

In the manner outlined in the previous two experiments, Mandarin-speaking raters independently rated all of the idioms and indicated whether or not they were familiar with each phrase. A reliability measure indicated satisfactory agreement between them (82.1% agreement, Cohen's Kappa = .80). Their ratings were averaged and a distribution of those means was compiled. Only idioms that were familiar to both raters were included in the final set of 75 (Appendix D).

Once it was determined which 75 idioms were to be used in the categorization study, another Mandarin-English bilingual who was naïve to the experimental hypothesis being tested translated each idiom into English. These translations were checked for accuracy by the Mandarin-English bilingual who had translated the rating instructions, and any ambiguous or problematic translations were discussed until acceptable translations had been determined. Idioms roughly matched for length across the three types such that normally analyzable idioms averaged 147.8 characters (*SD* = 10.1), abnormally analyzable idioms averaged 147.4 (*SD* = 29.2) characters, and unanalyzable idioms averaged 147.6 (*SD* = 21) characters.

Procedure

The procedure was identical to that followed in Experiments 1 and 2. Prior to beginning the experiment, participants read instructions regarding the conceptual categorization task that included relevant examples of idioms that were literally translated from Mandarin. Participants completed a brief practice session that included five literally translated Mandarin idioms that were not included in the experiment, with one from each of the five concept categories. This session familiarized the participants with the task and the data were not included in the final analysis. Participants took approximately 20 minutes to complete the experiment.

Results and Discussion

Response times for conceptual categorization are presented in Table 1, along with the proportion of idiom types (e.g., normally, abnormally, or unanalyzable) that were correctly categorized. Response times from incorrectly categorized phrases were eliminated from the analysis, as were response times greater than 3 *SD* beyond each individual's mean response time.

Native English speakers were able to correctly categorize literally translated Mandarin idioms according to concept in a pattern similar to that seen for native English speakers conceptually categorizing English and Latvian idioms, although there was an overall reduction in the number of idioms categorized correctly. Normally analyzable idioms were most often correctly categorized (82%), followed by abnormally analyzable idioms (59%), and unanalyzable idioms (50%), linear trend, $F_1(1, 99) = 90.98$, $p < .001$; $F_2(1, 72) = 24.13$, $p < .001$. The most noticeable drop in English speakers' correct conceptual categorization of Mandarin idioms relative to Latvian and English idioms was between the normally and abnormally analyzable group. However, the difference between the number of abnormally analyzable and unanalyzable idioms that were correctly categorized was consistent with the characterization of unanalyzable idioms as having historical/cultural bases that are opaque to individuals unfamiliar with the language in which the idioms are used.

Native English speakers' response times mirrored results for Latvian idioms in Experiment 2 and diverged from response times for English idioms in Experiment 1. Response times for Mandarin idioms were, on average, longer than for either English or Latvian idioms. This can be attributed to the length of Mandarin idioms when translated into English, since the phrases were quite a bit longer than either the English idioms from Experiment 1 or the translated Latvian idioms from Experiment 2.

An analysis of variance revealed a significant effect of idiom type, $F_1(2, 66) = 78.81$, $p < .001$; $F_2(2, 74) = 25.07$, $p < .001$. Further analysis of the individual means indicated that participants were able to correctly categorize normally analyzable idioms significantly faster (4.018 s) than abnormally analyzable idioms (5.123 s), planned comparison, $F_1(1, 66) = 74.21$, $p < .001$; $F_2(1, 72) = 7.62$, $p < .01$.

Furthermore, when participants did correctly categorize unanalyzable idioms, it took them significantly longer to do so than for either of the other idiom types (5.788 s vs. 4.571 s), planned comparison, $F_1(1, 66) = 83.10$, $p < .001$; $F_2(2, 74) = 6.54$, $p < .025$.

Just as they do when categorizing idioms from their own language, native English speakers correctly categorized many Mandarin idioms and did so at different rates depending on whether the phrases were normally, abnormally, or unanalyzable. As for the Latvian idioms, response times were fastest for normally analyzable idioms, followed by abnormally analyzable idioms, while unanalyzable idioms took participants the longest to categorize correctly. The response time differences for native English speakers' conceptual categorization of Mandarin idioms adds further support to the notion that idioms can be differentiated along an analyzability continuum.

Much of the debate regarding idiom processing has focused on whether we have pre-established mappings between the phrases and underlying conceptual structures or whether we create those mappings as we hear and comprehend a phrase (e.g., Gibbs, Bogdanovich, Sykes, & Barr, 1997; Glucksberg et al., 1992, 1993). The studies presented here establish a technique for examining speakers' processing of familiar and unfamiliar idioms and clearly demonstrate a difference between the time-course for accessing familiar mappings for making new mappings from source to target domains. Since it is difficult to control for the breadth of speakers' idiomatic knowledge in their native language, examining their processing of other languages' idioms provides another way to differentiate old from new conceptual mappings.

General Discussion

By examining the analyzability of idiomatic phrases from multiple languages and what people do when confronted with these phrases, the studies reported here demonstrate that the surface forms of many idioms are, in fact, analyzable for their underlying figurative meaning. Despite cross-linguistic differences in lexicalization, the conceptual underpinnings of many phrases from a variety of languages are identifiable.

In a forced choice task, native English speakers categorized idioms from both their own and two other languages according to the phrases' figurative meanings, where the other languages' phrases had been literally translated into English. People correctly categorized normally analyzable idioms from the two unfamiliar languages significantly more often and faster than they were able to correctly categorize abnormally analyzable idioms from those languages. In turn, they correctly categorized those languages' abnormally analyzable idioms more often and faster than they did the languages' unanalyzable idioms. These results may at first seem intuitive. However, when categorizing idioms from their own language, participants' response times were reversed for normally and abnormally analyzable idioms, such that they correctly categorized abnormally analyzable idioms faster than they did normally analyzable idioms. This finding is consistent with earlier work by Gibbs and his colleagues (1989) and contrasts with the pattern of response times observed for idioms from two other languages. Finally, people took the longest to correctly categorize unanalyzable idioms, even when the phrases were from their native language.

Differences in the speed and accuracy with which people are able to categorize idioms from their own and other languages reflect the degree to which figurative language can be variably analyzed. The differences in response times within each language suggest different forms of processing at various points along a continuum of analyzability. The divergence in response times for abnormally analyzable idioms from one's own and other languages supports arguments for the on-line mapping of idioms' surface forms to underlying structures, as well as arguments for pre-established structures, although the exact nature of those structures remains unclear. Finally, that native English speakers experienced the most difficulty making sense of unanalyzable idioms from their own and other languages supports the view that, while many idioms are uniformly analyzable, there are many idioms that are culturally- and / or historically-based. People are slow to make sense of such idioms, whether they are native speakers of a particular language or not. Perhaps unanalyzable idioms from one's own language never get mapped to any conceptual structure, thereby keeping processing slow. Likewise, unanalyzable idioms from unfamiliar languages are not readily mappable to any conceptual structure. One essentially must memorize such a phrase or (perhaps) create a new structure for the phrase to make any sense. The data reported here indicate that people are able to determine the figurative meanings of many such phrases, but this is most likely attributable to deductive processes they go through while focusing on very specific words within the phrases. The extremely long response times for this class of idiom support such an interpretation. For example, in Latvian *to go to butter* means "to go crazy". The words *go to* may have prompted native English speakers to identify a transformation (e.g., of mental state) and decide that this phrase indicated insanity, despite the fact that not much more in the phrase would have led them to conclude that.

This finding is consistent with other work showing that it is common for verbs of motion (such as "go") to take on new meanings indicating change or transformation (e.g., Sweetser, 1990). Regardless, the mapping took measurable time for people to achieve, as reflected by the response times for unanalyzable Latvian idioms.

Generally, results from the present study highlight the utility of conceptualizing analyzability as ranging along a continuum, with transparency anchoring one end and opacity anchoring the other. Analyzability is not absolute. The idea of three idiom "types" is an artificial distinctionto be sure, but it is one that has allowed experimental isolation of specific processing differences. The degree of analysis that needs to take place to make sense of any given idiomatic phrase can be located anywhere along the length of this analyzability continuum. The mapping between surface form and underlying structure along such a continuum might be better described as anchored at less metaphorical (and more literal) on one end, becoming increasingly metaphorical as one moves away from that end, and finally becoming relatively arbitrary, or culturally and/or historically based at the opposite end of the continuum. Initial-encounter processing times would likewise be located along this continuum, as indicated by the pattern of response times reported here. The variability in idiom analyzability observed in the present studies may account for many of the differences among experimental findings on idiom processing reported in the literature. That people determined the figurative meanings of idiomatic phrases that were simply translated word-for-word from another language into their own indicates that people can and do create conceptual mappings on-line. On the other hand, that people categorized abnormally analyzable idioms from their own language faster than normally analyzable idioms supports the view that certain conceptual mappings are well established and, thus, readily accessible.

Figurative language is of interest precisely because it is so prevalent and in such diverse forms across languages. It is important to understand how people who speak languages that are, quite literally, a world apart can relatively easily map figurative turns of phrase from one language onto corresponding concepts established in another. Although becoming fluent in a second language undoubtedly requires a more complex representation of the concepts captured by these phrases, at least initially this mapping between new words and established concepts must guide comprehension (Gonzales, 1999; Kroll & Stewart, 1994; Paradis, 1997). Our ability to appreciate literature that has been translated from languages around the world can also guide us in our pursuit of the answer to this question. Of course, translators generally do not translate idioms literally from the source language to the target language. They instead try to find an idiom in the target language that best represents the conceptual sense intended by the idiom used in the source language. But the art of translation depends on the translator's ability to recognize the concept or group of concepts being tapped by a given idiom in the source language and find the closest lexicalization to it.

The present study examines how analyzability plays a role in the processing of idiomatic expressions. Research that further characterizes the structure of figurative speech in general, and idiomatic phrases in particular, should include categorization studies where concept groups that are not represented by test idioms (e.g., happiness, disgust) are added as possible responses, eliminating the degree to which participants' responses are constrained. Additional "cross-linguistic" examinations are needed that include, for example, the same language as used by different communities of speakers (e.g., American vs. British vs. Australian English), eliminating some of the problems inherent to translation in general (e.g., Au, 1983, 1984). Such pursuits will help flesh out the picture that has emerged from the studies reported here regarding the cross-linguistic nature of certain forms of figurative language.

More generally, these results provide support for proposals that language should not be classified dichotomously, as either literal or figurative (e.g., Gibbs, 1994). Because people have the same or at least similar embodied experiences (e.g., how they perceive things, how they move, how they interact with things) in language group after language group and in culture after culture, they must develop the same schematic representations (for example, that containers contain things, that motion proceeds along a path, and so forth). These schematic representations should motivate and structure how specific phrases – such as idioms – evolve in any single language and account for how some of these phrases can be understood across languages. Of course, each language will have its own unique set of terms as well. But classifying these language-specific phrases as identical to those that can easily cross linguistic boundaries would be a mistake, parallel to the mistake researchers make when they collapse the process of speeded and functional (e.g., on-line) language comprehension with that of complex (e.g., time intensive) language appreciation (see Gerrig & Healy, 1983 regarding this distinction). Results from the present study highlight the difference between these two processes. The dissociation demonstrated here between response times for abnormally analyzable idioms from one's own and from another language captures the two processes in action. One happens quickly, guided by years of native-speaking experience and the body of native-speaker knowledge

that has established, and the other happens gradually, guided by semantic and conceptual knowledge about the world and one's place in it.

References

Au, T. (1983). Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited. *Cognition, 15,* 155–187.

Au, T. (1984). Counterfactuals: In reply to Alfred Bloom. *Cognition, 17,* 289–302.

Blasko, D. (1999). Only the tip of the iceberg: Who understands what about metaphor? *Journal of Pragmatics, 31,* 1675–1683.

Blasko, D., & Connine, C. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 295–308.

Boatner, M., Gates, J., & Makkai, A. (1975). *A dictionary of American idioms*. Woodbury, NY: Barron's Educational Series, Inc.

Bortfeld, H. (1998). *A cross-linguistic analysis of idiom comprehension by native and non-native speakers*. Unpublished doctoral dissertation, State University of New York, Stony Brook, NY.

Bortfeld, H. (2002). What native and non-native speakers' images for idioms tell us about figurative language. In R. Heredia & J. Altarriba (Eds.), Advances in psychology: Bilingual sentence processing (pp. 275–295). North Holland: Elsevier Press.

Bortfeld, H., & Brennan, S. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes, 23,* 119–147.

Bortfeld, H., & McGlone, M. (2001). The continuum of metaphor processing. *Metaphor and Symbol, 16,* 75–86.

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27,* 668–683.

Danesi, M. (1993). Metaphorical competence in second language acquisition and second language teaching: The neglected dimension. In J. E. Alatis (Ed.) *Georgetown University Round Table on languages and linguistics* (pp. 489–500). Washington, DC: Georgetown University Press.

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language, 6, 22–42.*

Gentner, D., & Wolff, P. (1997). Argument in the processing of metaphor. *Journal of Memory and Language, 37,* 331–355.

Gerrig, R., & Bortfeld, H. (1999). Sense creation in and out of discourse contexts. *Journal of Memory and Language, 41,* 457–468.

Gerrig, R., & Healy, A. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 667–675.*

Gibbs, R. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition, 8, 149–156.*

Gibbs, R. (1985). On the process of understanding idioms. *Journal of Psycholinguistic Research, 14,* 465–472.

Gibbs, R. (1992). What do idioms really mean? *Journal of Memory and Language, 31,* 485–506.

Gibbs, R. (1993). Why idioms are not dead metaphors. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure and interpretation* (pp. 57–78). Hillsdale, NJ: Erlbaum and Associates.

Gibbs, R. (1994). *The poetics of mind*. New York: Cambridge University Press.

Gibbs, R., Bogdanovich, J., Sykes, J., & Barr, D. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language, 37,* 141–154.

Gibbs, R., & Gonzales, G. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition, 20,* 243–259.

Gibbs, R., & Nayak, N. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology, 21,* 100–138.

Gibbs, R., Nayak, N., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language, 28,* 576–593.

Giora, R. (1997). Understanding figurative and literal languages: The Graded Salience Hypothesis. *Cognitive Linguistics, 8, 183–206.*

Giora, R. (2002). *On our mind: Salience, context, and figurative language*. Oxford, UK: Oxford University Press.

Glucksberg, S., Brown, M., & McGlone, M. (1993). Conceptual metaphors are not accessed during idiom comprehension. *Memory & Cognition, 21,* 711–719.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97,* 3–18.

Glucksberg, S., Keysar, B., & McGlone, M. (1992). Metaphor understanding and accessing conceptual schema: Reply to Gibbs. *Psychological Review, 99,* 578–581.

Gonzalez, V. (1999). *Language and cognitive development in second language learning*. Boston, MA: Allyn and Bacon.

Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quaterly, 2, 287–304.*

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics, 31,* 205–219.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R., & Aaron, D. (1991). Review of Lakoff & Turner: More than cool reason: A field guide to poetic metaphor. *Language, 67,* 320–228.

Katz, J., & Postal, P. (1963). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Kecskes, I. (2001). The graded salience hypothesis in second language acquisition. *Applied Cognitive Linguistics, 1, 249–271.*

Kecskes, I. (2002). *Situation-bound utterances in L1 and L2*. New York: Mouton de Gruyter.

Kecskes, I., & Papp, T. (2000). *Foreign language and mother tongue*. Mahwah, NJ: Lawrence Erlbaum.

Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language, 34,* 89–109.

Kövecses, Z., & Szabó, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics, 17,* 326–355.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics, 1, 39–74.*
Long, T., & Summers, D. (1979). *Longman dictionary of English idioms*. London: Longman.
Makkai, A., Boatner, M., & Gates, J. (1995). *A dictionary of American idioms*. Hauppauge, NY: Barron's Educational Series, Inc.
Matlock, T., & Heredia, R. R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 251–274). North Holland: Elsevier Press.
McGlone, M., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes, 17,* 167–190.
Murphy, G. (1996). On metaphoric representation. *Cognition, 60,* 173–203.
Murphy, G. (1997). Reasons to doubt the present evidence for metaphoric representation. *Cognition, 62,* 99–108.
Nunberg, G. (1978). *The pragmatics of reference*. Indiana University Linguistics Club.
Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language, 70,* 491–538.
Paradis, M. (1997). The cognitive neuropsychology of bilingualism. In A. De Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspective* (pp. 331–354). Hillsdale, NJ: Lawrence Erlbaum.
Sweetser, E. (1990). *From etymology to pragmatics: Cultural and metaphorical aspects of semantic structure*. New York: Cambridge University Press.
Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18,* 523–534.
Titone, D., & Connine, C. (1994). The comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1126–1138.
Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (Ed.), *Substance and structure of language*. Berkeley, CA: University of California.

Received December 12, 2002
Revision received April 1, 2003
Accepted April 2, 2003

Heather Bortfeld

Department of Psychology
Texas A&M University
College Station, TX 77843–4235
USA
Tel.: +1 979 845–2746
Fax: +1 979 845–2747
E-mail: Bortfeld@psyc.tamu.edu

Appendix A:

Summary of Rating Instructions

Some idioms have words whose meanings directly relate to their figurative interpretations. For example, in the phrase *pop the question*, the word *pop* is transparently related to the idea of "suddenly asking" or "suddenly proposing" while the word *question* refers to a particular kind of question, namely a "marriage proposal." Idiomatic phrases like this should be rated 1 on a continuum that runs between 1 (= transparent) and 5 (= opaque). On the other hand, there are idioms that are not analyzable. The words in the phrase do not clearly relate to its figurative interpretation. For example, in the phrase *break a leg*, the words *break* and *leg* are only opaquely related to the idea of "performing extremely well in some event." Idiomatic phrases like this should be rated as on or near 5 along the transparent-opaque continuum mentioned above. Finally, there are idioms that are analyzable to varying degrees, such that certain or all of their individual words have some relation to their figurative meanings. For some of these idioms, the relation between the phrases' words and their figurative referents are not entirely transparent, nor are they entirely opaque. Rather, the words can be understood as being metaphorically related to their figurative referents. For example, the phrase *line one's pocket* means something like "to embezzle funds." Both the words *line* and *pocket* have metaphorical relations to their figurative referents "embezzle" and "funds." Idioms like this, whose words are metaphorically related to their figurative referents, should be rated on or around a 3 along the transparent-opaque continuum.

Appendix B:

English Idioms for Insanity (Experiment 1)

INSANITY

Normally Analyzable:
to boggle one's mind
to be not all there
to be out of one's mind
to be all mixed up
to take leave of one's senses

Abnormally Analyzable:
to go off the deep end
to have bats in one's belfry
to come apart at the seams
to go to pieces
to lose one's marbles

Unanalyzable:
to go ape over something
to be off the wall
to be a basket case
to drive one bananas with
to be mad as a hatter

Appendix C:

Latvian Idioms for Insanity (Experiment 2)

INSANITY

Normally Analyzable:
for one to go out of order
to leave one's reason
to lose one's brain
when all five senses are not home
to darken one's mind

Abnormally Analyzable:
to become unyoked
to come off one's hinges
to be short of film
to let go of the reins
to walk on blowing sand

Unanalyzable:
to have a squeezed head
to go to butter
to turn into a Swede
to go in all heaven's directions
to be caught like a fish

Appendix D:

Mandarin Idioms for Insanity (Experiment 3)

INSANITY

Normally Analyzable:
for someone to act with no cause and no reason
the state of the mind is not at ease
to be inconsistent in joy and anger
to have unclear spirit and intelligence
to be frightened and out of control

Abnormally Analyzable:
the spirit is out of the house
lost soul and fallen spirit
for one's nerves to misfire
to be beyond rescue
to run without legs

Unanalyzable:
the man of Qi worries about the sky falling
seven up and eight down
to fish the moon out of the water
chickens fly and dogs jump
Szechwan dogs bark at the sun