

12 What Native and Non-Native Speakers' Images for Idioms Tell Us About Figurative Language

Heather Bortfeld
Brown University

Abstract

In a series of experiments, native and non-native English speakers were asked to form mental images of familiar American English idioms. Even when given explicit instructions to base their images on the phrases' literal meanings, both groups of speakers reported images that reflected fusions of the phrases' literal and figurative meanings. In a subsequent rating task, non-native speakers were judged to have produced more natural sounding sentences using the target idioms after the imaging task than they did before the task. In a second experiment, non-native English speakers were asked to form images of unfamiliar American English idioms both before and after being told the phrases' figurative meanings. Non-native speakers' images did not noticeably change between the first and the second imaging session, indicating that their images were not dependent on explicit knowledge of the phrases' figurative meanings. Emergent properties of metaphorical language are discussed.

Introduction

How is it that we understand not only literal language (when a phrase means what is explicitly stated), but also figurative language (when a phrase means something different from what is explicitly stated)? Various accounts have been offered to explain how people understand figurative language. Some researchers have argued that conceptual metaphors guide people's comprehension of figurative language (e.g., Gibbs, 1994). Others propose that conceptual schemas are constructed on-line and made available during comprehension (Glucksberg, Brown, & McGlone, 1993).

These proposals are based largely on data collected from speakers comprehending idioms from their native language. Yet even as native speakers, people have variable knowledge of their own language's idioms. Some phrases are highly overlearned while others have never been heard before. Some studies have examined the effects of differences in idiom familiarity (e.g., Blasko & Connine, 1993), but determining which processes are at work, given different knowledge states, is difficult. Few studies have examined how people process completely unfamiliar idioms (however see Keysar & Bly, 1995), but this is exactly the situation one finds oneself in when learning another language. Such a situation

UNIVERSITY OF NORTH TEXAS



ILLiad TN: 217296

Borrower: TXA

IN PROCESS: 20060830

Lending String: *INT,OKU,KKU,RCE,TLIS

Patron: Bortfeld, Heather

Journal Title: Bilingual sentence processing /

Volume: Issue:

Month/Year: 2002**Pages:** pp. 275-295

Article Author:

Article Title: Bortfeld, Heather; What native and non-native speakers' images for idioms tell us about figurative language.

Imprint: Amsterdam ; New York ; North-Holland/Els

ILL Number: 22673320



Call #: P115.4 .B55 2002

Location: WILLIS LIBRARY-3rd
FLOOR AVAILABLE

ARIEL

Charge

Maxcost: \$50IFM

Shipping Address:

Sterling C. Evans Library (41 HOU)
Texas A&M University - Inter-Library Ser
TAMU 5000, EVANS Library
College Station, TX 77843-5000

Fax:

Aril: 128.194.84.50

renders the adult learner almost child-like in his or her ignorance of the figurative meanings intended by these novel turns of phrase. Anyone who has spent time learning an additional language as an adult is familiar with this phenomenon. Yet, unlike children, adult learners are fully developed language comprehenders. They understand that their job is to determine what the novel combinations of words mean when used figuratively in the target language.

Why Study Idioms?

Idioms are remarkably frequent in everyday language and, indeed, entire dictionaries have been written about these phrases. For example, Boatner, Gates, and Makkai's original (1975) *Dictionary of American Idioms* presented over 4000 expressions that occur in English alone. The latest edition of this same dictionary (Makkai, Boatner, & Gates, 1995) contains over 8000. Increasingly, language scholars are acknowledging the wide range of idiomatic phrases that people use daily. Of course, language learners become aware of this as soon as they try to join a colloquial conversation being held in their new tongue. Since idioms are abundant in spoken language (any spoken language), they must be accounted for by any theory of language comprehension.

Idiom Processing

Gibbs and his colleagues have actively focused on the "decomposition" of idiomatic phrases into smaller meaningful units. Based on this work, they have come to characterize idiomatic phrases as variably decomposable (Gibbs & Nayak, 1989). This characterization is consistent with more recent linguistic analyses of idiomaticity (e.g., Nunberg, Sag, & Wasow, 1994) that criticize the traditional view (e.g., that idioms are noncompositional). Gibbs' analysis of idiom structure has led him to examine the relationship between metaphors and the phrases' surface content. He proposes that idiomatic meaning is motivated to varying degrees by speakers' implicit knowledge of the metaphors the phrases refer to. Gibbs argues that these conceptual structures arise from the connections we make everyday between domains of meaning. Other researchers have argued that these conceptual structures are instead constructed anew and made accessible in working memory during idiom comprehension (e.g., Glucksberg & Keysar, 1990; Glucksberg, Brown, & McGlone, 1993; Glucksberg, Keysar, & McGlone, 1992).

Cultural knowledge in idiom processing: But cultural knowledge and experience play an important role in the lexicalization of concepts in a given language as well. Cultural differences are often what make idiomatic phrases appear arbitrary to cultural outsiders. Of course, even members of a particular language group may find historically-based phrases in their language opaque in origin. Yet as native speakers of the language, these individuals come to access the intended meanings of the seemingly arbitrary phrases automatically, such that they quickly and easily understand them. But of interest here are those culturally-colored phrases that require analysis to be understood yet that are not completely opaque (e.g., based on some elusive historical event or fact). These variably "analyzable" phrases, which

differ from language to language, will help us understand more about how idiom processing works.

Perhaps the kernels of meaning upon which many idioms are based are general across languages and it is the stretching and warping that takes place when those concepts are lexicalized that is culturally motivated (Kittay, 1997). If so, can people apply conceptual knowledge established in their first language to make sense of figurative speech of this sort in another? Where recognizing a familiar idiom's figurative meaning won't necessarily require that we access the phrase's underlying metaphorical reference, trying to understand a novel idiom (e.g., one from another language) should prompt us to do the necessary mapping between the phrase's surface content and the conceptual structure that motivates its figurative meaning.

Mentally Imaging Idioms

Linguists and other scholars have documented the historical origins of many idioms. Native speakers tend not to have access to this historical information (e.g., why *kick the bucket* means 'to die'). However, the same speakers demonstrate tacit knowledge of the metaphorical basis for many other idioms. When asked to form a mental image for a phrase like *spill the beans*, people can do so without difficulty and can answer a variety of questions about that image (Gibbs & O'Brien, 1990). Although their reports are somewhat varied, native speakers generally report that the container for the beans is something like a large jar or bowl about the size of a person's head, that the beans were supposed to remain in the container but were accidentally spilled, that the beans go all over the place once spilled, and that they are not easy to get back into the container. How does this knowledge come about? There is no linguistic explanation for this tacit knowledge about certain idioms. The studies reported here attempt to determine whether this kind of knowledge crosses linguistic boundaries to help non-native speakers make sense of novel idioms from other languages.

Gibbs and O'Brien (1990) found that native English speakers' images for subsets of idiomatic phrases with similar figurative meanings were highly consistent, despite large differences in their surface forms (e.g., *keep it under your hat* and *hold your tongue*). Responses to follow-up questions about the causes and effects of different events within their images were also highly consistent. These researchers concluded that the image schemas they observed across speakers' images reflect knowledge that is the product of conceptual metaphors (e.g., THE MIND IS A CONTAINER and IDEAS ARE ENTITIES) that motivate the phrases' figurative meanings.

However, in trying to replicate Gibbs and O'Brien's (1990) finding with a group of native Italian speakers, Cacciari and Glucksberg (1995) found that the majority of the images these speakers generated for Italian idioms reflected only the literal actions and events denoted by an idiom's surface structure. Cacciari and Glucksberg concluded that such literal interpretations demonstrate that people's images do not reflect figurative meanings, "let alone the conceptual metaphors that presumably underlie those meanings" (p. 290).

Given these contradictory findings, perhaps there is some confusion over what aspects of people's mental images reflect what aspects of the phrases' meanings.

Indeed, in other work (e.g., Gibbs, 1994) Gibbs has argued that the distinction between literal and figurative is not at all categorical and should instead be conceptualized as falling along a continuum. It is not possible in either of these two imaging studies to determine whether participants construed experimental instructions to be asking for an image of a phrase's literal or figurative meaning. If no further instruction was given and no clarification was sought, then perhaps the difference between the two studies' results is due to different interpretations of the task by the different participant populations.

Given that figurative interpretations of familiar idioms are made just as quickly as literal ones (Gibbs, 1980, 1986; McElree & Nordlie, 1999), even if native speakers were explicitly asked to produce an image based on an idiom's literal meaning, they would likely experience interference from the phrase's figurative meaning while forming that image. The findings reported by Gibbs and O'Brien (1990) may reflect difficulties that native speakers have in separating familiar idioms' literal and figurative meanings. Non-native speakers who are unfamiliar with their new language's idioms, on the other hand, shouldn't have automatic access to any pre-stored figurative meanings for the phrases and thus should find it easier to produce strictly literal images for them. If non-native speakers instead produce the kind of hybrid images that the native speakers in Gibbs and O'Brien's study produced, this would indicate that they engage in a process of mapping the phrases' surface forms to conceptual structures similar to those held by the native speakers. While the on-line aspect of this mapping process is consistent with the view that Glucksberg and his colleagues hold on idiom processing, the similarity between the outcomes of native and non-native speakers' mapping would likewise support the view held by Gibbs and his colleagues. Experiment 1 was designed to test this hypothesis.

Experiment 1: Native and Non-Native English Speakers' Mental Images for American English Idioms

Method

Native English and native Latvian speakers learning English as a foreign language were asked to form mental images based on the literal meanings of a group of American English idiomatic phrases.

Participants

Native speakers: Twenty-five undergraduates (17 women and 8 men) at the State University of New York at Stony Brook received course credit in exchange for participation in this experiment. All were native speakers of English and none reported the ability to speak another language.

Non-native speakers: Twenty-five undergraduate students (20 women and 5 men) at the University of Latvia received course credit in exchange for participation in this experiment. All were in their first or second year of the English philology

concentration. All were native speakers of Latvian, spoke Russian¹ to varying degrees, and had been studying English for five years or less.

Materials

Stimuli were the 25 idioms used by Gibbs and O'Brien (1990). The phrases represent 5 concepts (e.g., anger, revelation, secretiveness, insanity, and control), with 5 different idioms expressing each specific concept (e.g., anger = *blow your stack*, *flip your lid*, *lose your cool*, *hit the ceiling*, and *foam at the mouth*). All idioms were "abnormally analyzable," as defined by Gibbs and Nayak (1989) in their taxonomy of idiom types. These five groups of idioms are listed in the Appendix.

The six probe questions from Gibbs and O'Brien's (1990) original study were also used. As noted by those researchers, the probe questions were designed to reveal participants' beliefs about various aspects of their mental images. Probe questions were of two types: 1) questions requiring yes/no responses (e.g., Intentionality--"Was the action done intentionally?"; Reversibility--"Is it difficult to reverse the action?"), and 2) questions eliciting open-ended responses (e.g., Causation--"What caused the action to happen?"; Manner--"How was the action performed?"; Consequence--"What happens as a result of the action?"; Negative Consequence--"What would happen if the action didn't happen?"). During the imaging task, each probe question was altered to match the specific mental image that participants described for each idiom. For example, if a subject reported that her image for *blow your stack* was someone with fire and smoke shooting out of the top of her head, they would be asked "What causes the fire and smoke to shoot out of the top of this person's head?" for the Causation question. All probe questions were adjusted in this way according to participants' images.

Procedure

Native speakers: Native speaking participants were first given a sheet of paper with a list of the 25 idioms on it and were asked to generate a sentence for each, using the phrase figuratively. They were then asked to form mental images based on the literal meanings of the 25 different idiomatic phrases. Idioms were presented to each participant sequentially in a random order. Once a mental image was generated, native speakers were asked the series of detailed probe questions (outlined above) in an invariant order (e.g., Causation, Intentionality, Manner, Consequence, Negative Consequence, and Reversibility). Each imaging session was audiotaped and later transcribed verbatim for subsequent coding. Following the imaging task, native speakers were given a second sheet of paper listing the 25 idioms and again were asked to generate sentences using the phrases figuratively.

Idioms for which native speakers were unable to provide a mental image were skipped and re-presented at the end of the interview. If a native speaker was still unable to provide a mental image for that phrase, the phrase was skipped entirely. The interview was conducted according to a script, and very little discussion of the images took place, aside from simple clarification questions (e.g., "What do you see coming out of the person's head?" when someone described "stuff" exploding out of

the top of a person's head for *blow your stack*). Each interview took less than one hour to complete.

Non-native speakers: The protocol for the non-native speakers was similar to that followed for the native speakers. Upon scheduling an imaging session, they were given a list of the 25 idioms (in English) accompanied by the phrases' figurative paraphrases (also in English). Paraphrases for each idiom were based on Boatner, Gates, and Makkai's (1995) definition for the phrase. Non-native speakers were told to "become familiar with" the figurative meanings of the idiomatic phrases, as they would be asked several questions about those phrases during the imaging session. Sessions always took place two weeks after scheduling.

As was the case with native speakers, non-native speakers were asked to generate written sentences using the idioms figuratively both prior to and following the imaging session. At the beginning of the session, they were told that they would be presented with the idiomatic phrases they had been given two weeks earlier and that they were to form a mental image based on the literal meaning of that phrase. Non-native speakers were asked to describe the first image that came to mind, and to answer the probe questions in as detailed and accurate a manner as possible. Each interview took less than one hour to complete.

Coding

The interviews were transcribed and imported into Sequence[®], a software package designed specifically for coding linguistic corpora. Two research assistants naive to the experimental hypothesis being tested were trained to code the images and answers to probe questions following the procedure outlined by Gibbs and O'Brien (1990). The native speakers' images and answers to probe questions were coded first, in order to establish a baseline for comparing the non-native speakers' images and answers.

Images were first identified as reflecting either the literal or figurative meaning of an idiom, or some hybrid of the two meanings. In order to be considered entirely literal, an image had to describe *only* the literal contents of the phrase, with no reference to the underlying emotion or desire described by the figurative meaning on the phrase. Conversely, those images that were considered entirely figurative had to describe *only* the emotion or desire that the figurative meaning of the phrase entailed with no reference to the surface structure of the phrase. Hybrid images had to include references to both the concrete actions described by a literal reading of the phrase, in addition to references to an underlying (figuratively accurate) emotion or desire that motivated those actions. Examples of the three types of images for the idiom *go off your rocker* from the insanity concept group are presented in Table 1.

For example, when a native speaker reported that his image for *flip your lid* was of the top of a person's head being forcefully blown up, this was scored as a hybrid image in which "some force causes the top of a container to release pressure violently." Reliability was coded with satisfactory agreement between coders (96.5% agreement, *Cohen's Kappa* = .95). Coders discussed any discrepancies until agreement was reached.

Table 1. Sample Responses by Native Speakers to Questions Regarding Their Image for the Phrase "Go Off your Rocker."

Literal	
General Image:	Getting off of my rocking chair to get food
Causation:	Hunger
Intentionality:	Yes, it was intentional
Manner:	Casually standing up
Consequence:	I get food
Negative Consequence:	More hunger
Reversibility:	It's not difficult to reverse, just sit back down
Figurative	
General Image:	The old man loses his mind
Causation:	He's getting too old
Intentionality:	No
Manner:	Absent minded (effect of old age)
Consequence:	He gets put away
Negative Consequence:	He'd be normal
Reversibility:	Difficult too, because it happens sooner or later
Hybrid	
General Image:	A crazy woman rocks and rocks in her rocking chair until she loses her grip on the chair and flies off
Causation:	Loss of balance due to her agitation
Intentionality:	No
Manner:	Surprisingly
Consequence:	Breaks her hip
Negative Consequence:	Healthy hip
Reversibility:	No, it's hard to stop once it's happening

Collapsing across concept groups, a total of 70% of the native English speakers' images qualified as hybrids of phrases' figurative and literal meanings, 28% were based entirely on a literal reading of the phrases, and 2% clearly referred to only the phrases' figurative meanings.

Since the majority of images for phrases represented a hybrid of the two meanings, the hybrid images in each concept group were then coded for their general characteristics. In order to do this, each coder read through all of the participants' hybrid images and responses to the corresponding probe questions and independently developed five schemas intended to capture the most common answers provided for each conceptual group. Once coders had established

independent schemas for the images' general characteristics, these were discussed until a single representative schema for each concept was agreed upon. Table 2 outlines the image schema derived from native English speakers' most frequent images and responses to probe questions for anger idioms. A schema was established for each of the five concept groups.

Table 2. Schema Derived from Native English Speakers' Most Frequent Responses for Anger Idioms

Anger	
General Image:	Pressure forces some container to release its contents violently upwards and outward Stress, anger, and/or frustration
Causation:	Action is not intentional
Intentionality:	Forcefully, violently, uncontrollably
Manner:	Pressure is released or relieved
Consequence:	Container remains intact
Negative Consequence:	Action is difficult if not impossible to reverse
Reversibility:	

Once schemas were established, coders independently scored the native and non-native speakers' image descriptions and probe question responses for all images (including those that had been classified as entirely literal or figurative) according to whether or not they matched the schemas. The coding program allowed images and responses to be grouped according to concept with the variable labels hidden, so that coders would not be biased by such information (e.g., whether the image had been considered literal, figurative, or a hybrid of the two; which participant provided that image; which idiom was being described). Coders were able to view the image schema key (developed earlier) corresponding to a particular concept while coding any given image and probe question answers. Images and probe responses were coded independently of one another. This prevented coders from "reading into" or projecting a response onto a probe question based on supplementary information available in responses to other questions for the same idiom.

Images and responses were coded as either matching or not matching the corresponding schema. In order to be considered a match, a general image had to include both the literal and figurative aspects of the image schema that had been identified for that image's concept group. For example, if a participant reported that she saw "a lion tamer cracking a whip to discipline a lion" for the idiom *crack the whip*, this was coded as a match with the general schema for control. That is, this general image includes a concrete action (cracking the whip) for a purpose consistent with the figurative meaning of the phrase (that order is imposed in some way by someone else). A general image was considered a non-match if only a literal

action was included (e.g., "a man jumping too high and hitting his head on the ceiling" for *hit the ceiling*) with no reference to the underlying figurative meaning of the phrase. Conversely, if the general image included only a reference to the figurative meaning (e.g., "to sneak around with a secret" for *behind one's back*) with no reference to the literal contents of the phrase, this was also considered a non-match.

In order for responses to the open-ended probe questions to be coded as matches, they had to include at least one of the characteristics listed in the corresponding image schema. For example, if a person answered the Causation question for the phrase *blow your stack* by saying "something upset them," this was counted as a match since the image schema for that concept's Causation probe includes "stress, anger, and/or frustration." Responses were coded as a non-match if they did not include at least one aspect of that probe's listing for the corresponding concept. For example, when a participant reported that "someone wanted to roll a ball to someone else" for the phrase *keep the ball rolling*, this was coded as a non-match because the Causation probe for control had to include "exertion of authority or control over others."

A test of reliability showed satisfactory agreement between coders for native speakers' images (92.4% agreement, *Cohen's Kappa* = .91) and for non-native speakers' images (94% agreement, *Cohen's Kappa* = .92). Again, all coding discrepancies were discussed until agreement was reached. Compiling codes across participants and across idioms within each of the five conceptual groups provided a baseline summary of the most frequent image characteristics for each conceptual group for the two groups of speakers.

Results

Where relevant, two analyses are reported; t_1 is the analysis by-subjects, and t_2 is the analysis by-items. The criterion for significance was set at $p < .05$. Overall, both groups of speakers produced images and responses to probe questions that matched the hybrid image schemas developed for each conceptual group of idioms. Averaging across the different idiom groups reveals that 74% of native speakers' mental images matched the schemas for hybrid images (ranging from 61% to 89% for the five concepts). This proportion is significantly different from chance ($t_1(24) = 1.93, p < .05$; $t_2(24) = 1.72, p < .05$) even with a conservative null hypothesis of .50 that assumes participants could only give two possible responses. An average of 72% of non-native speakers' images matched the native speakers' most common schemas for hybrid images (ranging from 58% to 87%), also significantly different from chance ($t_1(24) = 1.73, p < .05$; $t_2(24) = 1.72, p < .05$).

In addition to their general image schemas, participants' answers to the probe questions illustrate participants' understanding of the details characterizing their images. Data from probe questions were of two types: Some questions required yes/no responses and others elicited open-ended responses that were coded in the same manner as participants' images. Across the five idiom groups, native-speakers gave responses that matched the schemas 80% of the time to the Causation questions, 86% to the Intentionality questions, 79% to the Manner questions, 82% to the Consequence questions, 76% to the Negative Consequence questions, and

75% of the time to the Reversibility questions. The proportions for all the probes were greater than chance ($t(24) = 2.03, 2.78, 1.98, 2.23, 1.81, \text{ and } 1.79$, respectively, $p < .05$ for each comparison). Non-native speakers gave responses that matched the schemas 73% of the time to the Causation questions, 82% to the Intentionality questions, 78% to the Manner questions, 81% to the Consequence questions, 75% to the Negative Consequence questions, and 72% of the time to the Reversibility questions. The proportions for all but the reversibility probes were greater than chance ($t(24) = 1.74, 2.19, 1.88, 2.25, \text{ and } 1.78$ respectively, $p < .05$ for each comparison). Since each percentage represents a single item from the imaging questionnaire, no item analyses were carried out for this set of data.

Discussion

Native and non-native speakers both formed images for idioms that reflect a fusing of the phrases' figurative and literal meanings. For example, the idioms *go off your rocker* and *lose your marbles* both mean "to go crazy" when used figuratively. Both native and non-native speakers described images for these idioms that involved a complete loss of stability (most often physical) on the part of the individuals imagined, yet involved rocking chairs in the former instance and marbles in the second. If they were interpreting the phrases literally, as they had been instructed to do, nothing in the lexical makeup of these phrases should have constrained their image production to such a degree. Instead, both groups appeared to have difficulty forming images separating the phrases' literal from their figurative meanings.

These results support and extend Gibbs and O'Brien's (1990) findings regarding native speakers' images for idioms. First, native speakers more often produced images for idioms that hybridized the phrases' figurative and literal meanings than reflected just their literal or figurative meanings. Non-native speakers did the same. In fact, in certain instances, non-native speakers' images were *more* consistent with the established schemas than a subset of the native speakers' images. This was because a subset of the native English speakers interpreted certain idioms in terms of more contemporary, slang meanings of the terms (e.g., *blow your stack* to mean 'spend all your money' rather than 'to get extremely angry'). The non-native speakers, on the other hand, had only just learned the phrases' canonical, idiomatic meanings and did not have enough cultural exposure to English to be aware of the phrases' more recently developed colloquial meanings. Despite this, the overall similarity between both groups' images and their responses to subsequent probe questions is noteworthy. More surprising is the fact that both groups consistently hybridized the phrases' literal and figurative meanings while imaging them.

Experiment 2: Naturalness of Native and Non-Native English Speakers' Use of Idioms Before and After Imaging

If native speakers' images of idioms reflect mappings between surface structures and underlying conceptual structures, then non-native speakers stand to gain

linguistic knowledge in their new language by doing this mapping themselves. Of course, the rich linguistic knowledge that native speakers have to characterize each of these domains is what distinguishes them as native speakers. Nevertheless, adult learners who successfully acquire another language do so in large part because they adapt to the way speakers of the new language lexicalize familiar concepts. One way to promote this is for adult learners to analyze phrases from the new language in terms of those underlying structures. This is a process that the mental imaging task appears to promote as well.

If this is the case, then the sentences that the native and non-native speakers produced both prior to and following the mental imaging task in Experiment 1 provide an ideal opportunity to examine whether the task had any influence on the non-native speakers' linguistic abilities. Experiment 2 tests the prediction that non-native speakers will use idiomatic phrases from their new language more naturally following participation in a task that prompts them to analyze how the phrases' surface content maps to corresponding conceptual structures.

Method

Native English speakers rated the sentences that were collected before and after the imaging task (completed in Experiment 1) for naturalness.

Participants

Thirty undergraduate psychology students (16 women and 14 men) at the State University of New York at Stony Brook received course credit for participating in the study. All identified themselves as native speakers of English and none reported the ability to speak another language.

Materials

The sentences generated by both native and non-native speakers before and after the imaging task were compiled into groups reflecting the five concepts described in Experiment 1. One idiom was randomly selected from each set in the five concept groups. Before-and-after sentences produced by each of the 50 participants in Experiment 1 (25 native speakers and the 25 non-native speakers) were then transferred into a booklet, resulting in 250 pairs of sentences. The sentences were blocked by concept. Within each block, sentences were randomized such that pairs of sentences produced by native and non-native speakers were mixed. Within each pair, sentences produced prior to and following the mental imaging task were also randomized. The following instructions preceded each pair of sentences: (1) *Circle the sentence that sounds more natural to you* and (2) *Rate how natural that sentence sounds*. The second question was followed by a 7-point scale with 1 anchored as "not very natural" and 7 anchored as "very natural."

Procedure

Participants were instructed to compare each pair of sentences, circle the more natural sounding one, and then rate that sentence for how natural it sounded. Participants completed the questionnaire individually.

Results

Before and after choices: The naturalness of native speakers' expressions was not affected by their participation in the mental imaging task; the sentences they produced following the mental imaging task were judged to sound more natural than those produced before the task 52.9% of the time, a difference that is not greater than chance ($t_1(29) = 1.57, p > .05$; $t_2(123) = .52, p > .05$). In contrast, the sentences that non-native speakers produced after participating in the mental imaging task were judged to sound more natural than those they produced before the task 62.8% of the time, a difference that is greater than chance ($t_1(29) = 6.44, p < .001$; $t_2(123) = 3.21, p < .003$).

Naturalness ratings: Although the non-native speakers used English idioms in sentences more naturally following the mental imaging task, the sentences they produced were still rated as less natural than those produced by native speakers (4.01 vs. 5.47; $t_1(29) = 6.50, p < .001$; $t_2(248) = 7.76, p < .001$).

Discussion

Findings from Experiment 2 indicate that analysis of an idiom's surface form (as is promoted by the mental imaging task) can help people recognize an idiom's conceptual underpinnings. Despite the fact that non-native speakers were aware of the figurative meanings of the idioms prior to participating in the mental imagery task, the sentences they produced based on that knowledge were not as natural sounding as the sentences they produced following the imaging task. Although the increase in naturalness in post-imaging sentences could be a product of the increased time non-native speakers were actually aware of the canonical definitions of the phrases, it seems likely that analyzing the phrases helped non-native speakers better understand the mapping between surface forms and conceptual structures.

Experiment 3: Influence of Prior Knowledge on Native and Non-Native Speakers' Mental Images for Idioms

It is still unclear whether non-native speakers' knowledge of the canonical definitions of target idioms prior to participation in the imaging task resulted in their production of hybrid images or whether their hybrid images emerged from the analysis promoted by the task. Three issues that need to be resolved include, 1) whether or not one's prior knowledge of the figurative meaning of an idiomatic phrase influences the mental image one forms for that phrase's literal meaning; 2) whether the similarities seen between native and non-native speakers' images for idioms in Experiment 1 disappear when non-native speakers' native language is not

from the same language family as that of the target language; and 3) whether analyzing idioms' surface structure helps non-native speakers determine their figurative meanings. Experiment 3 was designed to address these issues.

Method

As in Experiment 1, Experiment 3 compares the mental images generated by native and non-native speakers for the same set of 25 English idioms. In the present study, however, participants performed the imaging task twice, both before and after being told the target phrases' canonical figurative meanings. To eliminate any confounds that may have been introduced due to genetic similarities between English and Latvian (both Indo-European languages), non-native speakers in the present study were native speakers of Mandarin (a Sino-Tibetan language). As in Experiment 1, participants generated sample sentences using the target idioms both prior to and following the first imaging session. This served the dual purpose of measuring the participants' familiarity with the idioms and whether engaging in the imaging task helped participants determine the figurative meanings of the phrases. Finally, a questionnaire was substituted for the interview format used in Experiment 1 to control for any influence the interactive nature of the interview may have had on participants' responses.

Participants

Native speakers: Twenty-five undergraduate students (12 women and 13 men) of the State University of New York at Stony Brook received monetary compensation for their participation in the experiment. All identified themselves as native speakers of English and none reported the ability to speak another language.

Non-native speakers: Participants were twenty-five graduate students (14 women and 11 men) at the State University of New York at Stony Brook who received monetary compensation for their participation in the experiment. All were native Mandarin Chinese speakers from Mainland China.

Although it is common for people from China to speak any number of Sino-Tibetan dialects² depending on what region they are from, formal education takes place in Mandarin and all forms of media in the country are transmitted in Mandarin. As graduate students, the non-native speakers in Experiment 3 all completed several years of higher education in China and thus were all fluent in Mandarin. Given that any of the Chinese dialects the students might speak in addition to Mandarin are also unrelated to Indo-European languages, their multilingualism did not present itself as a confound in the present study. Participants were recruited from English as a Second Language (ESL) courses offered by Stony Brook's Department of Applied Linguistics to new non-English speaking graduate students. All had Test of English as a Foreign Language (TOEFL) scores between 500 and 600 (scores can range from 200 to 800). TOEFL is a comprehension and writing test that does not involve spoken English. TOEFL scores were used as a rough indicator of English knowledge because the test is used as a standard measure of fluency among foreign students. Graduate students with scores under 600 are typically asked to attend an ESL course in order to improve their English skills prior

to serving as teaching assistants. However, even a score above 600 does not guarantee fluency in spoken English nor familiarity with idiomatic English.

Materials

Stimuli were the same 25 idiomatic phrases used in Experiment 1 (see Appendix). Five versions of the questionnaire were prepared, each with the target idioms presented in a different random order. Questionnaires were identical for native and non-native speakers, except that the non-native speakers' questionnaires consisted of the English idiomatic phrases and their literal translations in Mandarin. The literal translations were provided to discourage participants from using Mandarin-English dictionaries to look up individual words. Written instructions and probe questions were also translated.

Translations: Although the difficulties inherent in translating experimental stimuli from one language to another are well-documented (see Au, 1983, 1984), literal translations of the idioms were necessary, given the non-native speakers' relatively limited English vocabularies. Every effort was made to ensure that the translations from English to Mandarin were accurate and reflected the appropriate meanings of both the instructions and experimental stimuli. Participants were encouraged to write in Mandarin if they preferred to do so, and all did.

Translations were completed by two fully bilingual Mandarin-English speakers and were judged for accuracy by 5 native Mandarin speakers who had resided in the United States for 10 years or more. Any ambiguities were corrected, based on these 5 speakers' recommendations. Of the test idioms, only two out of the 25 literal translations needed to be discussed and the ambiguities in those translations were noted by only 2 of the 5 reviewers.

Procedure

The experimental protocol for each of the two sessions included a pre-test, the mental image questionnaire, and a post-test identical to the pre-test. Participants were tested individually. Data from the native speakers' pre-tests established a baseline for how familiar they were with the test idioms and data from the non-native speakers provided information about how much idiomatic English they knew prior to participating in the study. For the pre-test, participants were told to provide a paraphrase of the figurative meaning for each phrase they identified. Participants then completed the mental image questionnaire. Following completion of the questionnaire, participants were given a post-test and again were told to simply provide figurative definitions for as many of the phrases as they could.

Following completion of the first session, all participants were given a sheet of paper with the canonical figurative definitions of the 25 idiomatic phrases. They were told to check their understanding of the idioms' meanings against the canonical figurative definitions provided on the sheet of paper prior to returning for the second session. At this time, a second session was scheduled to take place approximately two weeks after the first session. The second session followed the same format as the first, except that no post-test was administered.

Coding

Native English speakers' responses to the first and second mental image questionnaires were transcribed and imported into Sequence[®] (in the same manner described in Experiment 1). Native Mandarin speakers' responses were translated by two bilingual Mandarin-English speakers who were naïve to the experimental hypotheses being tested. No significant semantic differences were noted between the two sets of translations, ensuring that participants' responses were accurately represented by the translations. These responses were then transcribed and imported into the coding program.

Coding was completed by the two research assistants who coded participant responses for Experiment 1. The image schemas established in that experiment were used as a comparison measure. As in the first experiment, images were coded for whether they were based on an idiom's figurative or literal meaning, or instead represented a hybrid of the two. The definitions that each participant provided preceding and following the first mental imaging session were coded for whether participants knew the canonical figurative meanings of each phrase or not. A test of reliability indicated satisfactory agreement between coders across the different material (91.6% agreement, *Cohen's Kappa* = .90).

Results

The percentage of images reported by native and non-native speakers that were coded as literal, figurative, and hybridized are shown in Table 3. Again, even when explicitly instructed to base their images on the literal meanings of the phrases, both groups of speakers reported hybrid images the majority of the time. Collapsing across the different idiom groups for the first imaging session, an average of 69% of native English speakers gave hybrid images for the different English idioms (ranging from 56.1% to 76.1%). For the second session this average was 67% (ranging from 58.7% to 74.8%). Collapsed across the different idiom groups, an average of 47% of the native Mandarin speakers' images during the first session were hybrids (ranging from 27.7% to 57.4%). For the second session this average was 54% (ranging from 27.1% to 73.4%). These data indicate that it is more difficult for native speakers to separate the literal from the figurative meanings of idiomatic phrases from their own language than it is for non-native speakers. However, the data also indicate that non-native speakers fused the literal and figurative meanings of the idioms almost 50% of the time both prior to and following being made explicitly aware of the phrases' canonical meanings. This was the case even when non-native speakers had no prior knowledge of the majority of the phrases' figurative meanings. Data from the pre-tests indicate that non-native speakers were familiar with just 15% of the English idioms prior to the imaging task.

Focusing on results from the hybrid images, the data from the first and second imaging sessions indicate that both groups of participants gave responses that were consistent with the concept schemas established during Experiment 1. Collapsing across the different idiom groups, an average of 76% of native English speakers gave responses that matched the schemas for the different English idioms (ranging

from 69.5% to 80.1%) during the first imaging session, a proportion that is significantly different from chance ($t(30) = 1.81, p < .05$). An average of 73% of the non-native speakers' responses during the first session matched the schemas of the different English idioms (ranging from 66.5% to 78.5%). This result is marginally different from chance ($t(31) = 1.62, p < .10$). These results indicate that non-native speakers' analyses of the phrases during the imaging task gave them an indication of the phrases' figurative meanings, at least to some degree.

Table 3. *Image Types Reported by Participants During First and Second Mental Imaging Sessions*

English Background:	Concept Group	Session: Image Type	Native			Non-Native		
			I	II	III	I	II	III
Anger		Hybrid:	56.1	65.2	54.2	65.8		
		Literat:	40.0	29.0	32.9	12.3		
		Figurative:	3.9	5.8	12.9	21.9		
Control		Hybrid:	76.1	74.8	57.4	73.4		
		Literat:	22.6	21.9	37.4	18.8		
		Figurative:	1.3	3.2	5.2	7.8		
Secrecy		Hybrid:	75.9	74.2	56.8	60.0		
		Literat:	23.2	21.3	26.5	14.2		
		Figurative:	0.6	4.5	16.8	25.8		
Insanity		Hybrid:	72.3	64.5	37.0	41.3		
		Literat:	25.8	30.3	51.9	36.1		
		Figurative:	1.9	5.2	11.0	22.6		
Revelation		Hybrid:	65.8	58.7	27.7	27.1		
		Literat:	32.3	36.1	66.5	44.5		
		Figurative:	1.9	5.2	5.8	28.4		

Collapsing across idiom groups for the second session reveals that an average of 77% of native English speakers' hybrid images for the English idioms matched the baseline schemas (ranging from 69.5% to 80.8%), a proportion that is significantly different from chance ($t(31) = 1.95, p < .05$). An average of 76% of non-native English speakers' hybrid images for the different English idioms collapsed across the different idiom groups described similar general images (ranging from 73% to 78.5%), also significantly different from chance ($t(31) = 1.93, p < .05$). These data indicate that the Mandarin speakers were even more likely to fuse literal and

figurative meanings of the phrases once they were made aware of the canonical figurative definitions of the phrases.

Overall, both groups of participants reported highly consistent responses to the different probe questions about their images for the English idioms (see Table 3). This was particularly so for images produced during the second imaging session. The probe question data indicate that when participants' images were hybrids, their intuitions about the characteristics of their images tended to match the schemas developed based on native speakers' most common responses. Finally, data from pre- and post-tests conducted during the first imaging session support the hypothesis that analysis of idioms' surface structure helps non-native speakers deduce their figurative meanings. These data are presented in Table 4.

Table 4. *Types of Definitions Provided at Pre- and Post-Test During Imaging Session I*

Concept Group:	Test:	Natives			Non-natives		
		Pre-	Post-	Post-	Pre-	Post-	Post-
Anger	Correct:	74.8	91.0	9.7	9.7	31.0	
	Incorrect:	7.1	6.5	10.3	10.3	58.1	
	Slang:	11.0	2.6	--	--	--	
Control	Correct:	75.5	87.7	5.2	5.2	49.0	
	Incorrect:	5.2	1.9	22.6	22.6	36.8	
	Slang:	16.1	10.3	--	--	--	
Secrecy	Correct:	91.6	97.4	45.2	45.2	82.6	
	Incorrect:	2.6	0.6	12.9	12.9	13.5	
	Slang:	3.2	1.9	--	--	--	
Insanity	Correct:	76.8	90.3	5.8	5.8	41.0	
	Incorrect:	5.2	1.9	11.7	11.7	51.3	
	Slang:	15.5	7.7	--	--	--	
Revelation	Correct:	81.3	96.1	12.3	12.3	45.2	
	Incorrect:	6.5	1.9	16.1	16.1	45.2	
	Slang:	8.4	1.9	--	--	--	

Where non-native speakers gave correct responses only 15% of the time on the pre-tests, they correctly defined the test idioms 49.7% following completion of the imaging task. This result makes a compelling argument for the analyzability of at least a subset of idiomatic phrases. It is notable that at no time prior to the post-test were these non-native speakers given information about the phrases' figurative meanings. The increase in correct figurative definitions at post-test can only have been a product of analysis of the phrases' surface structure and how this mapped to conceptual knowledge (an analysis promoted by the imaging task). In contrast,

changes between native speakers' pre- and post-test responses were almost entirely due to a reduction in slang interpretations of certain phrases at post-test.

Discussion

Results from Experiment 3 are consistent with the initial findings regarding speakers' mental images for idioms. Non-native speakers produced lots of hybrid images and these hybrids matched schemas that were developed based on native speakers' images, as did their answers to a standard set of probe questions asked regarding those images. Following a two week period, during which non-native speakers had access to the test idioms' canonical figurative meanings, even more of their responses matched native speaker-based image schemas.

Pre- and post-tests from the first imaging session indicate that speakers are able to infer figurative meaning based on an analysis of the surface structure of idiomatic phrases, without additional supporting context. That these phrases can be understood by non-native speakers with no instruction regarding their canonical meanings is compelling evidence that at least a subset of idiomatic phrases are analyzable.

Although the images that non-native speakers produced prior to being told the phrases' figurative meanings were more often based on literal interpretations of the phrases than those produced by native speakers, the increase in accurate definitions of the idioms' figurative meanings at post-test indicates that the non-native speakers were engaged in this process of analysis. The similarities between the native and non-native speakers' images in both Experiment 1 and 3 indicate that this process of analysis guides and constrains idiom comprehension more than knowledge of the phrases' canonical figurative meanings itself.

Again, there is no explanation for such regularity in people's knowledge about figurative phrases in traditional linguistic theories of idiomaticity. Data from the probe questions indicate that both native and non-native speakers are able to access very specific information about the causes and effects of actions in their images for idioms. Given the striking similarities between native and non-native speakers' images and their answers to probe questions about those images, these findings indicate that specific conceptual structures underlie at least some subset of languages' idiomatic phrases.

Another issue that these data address is that of the universality of these conceptual structures. Figurative language is fascinating precisely because it is so prevalent and in such diverse forms across languages. It is important to ask 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 turns of phrase in another. The fact that figurative phrases originating in one language can be understood and appreciated by speakers of another attests to the importance of this question. The conceptual structure or structures tapped by a given idiomatic phrase must be shared by the two languages in order for it to be analyzed. This is indicated by the cross-linguistic data reported here.

Our ability to appreciate literature that has been translated from other languages addresses the universality question as well. Translators do not translate idioms

word-for-word from the source to the target language. Instead, they try to locate an idiom in the target language that best represents the sense intended by the idiom used in the source language. In doing so, good translators maintain the source language's original mapping from words to conceptual structure. 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 in the target language. Truly talented translators work to preserve original structural representation (e.g., turns of phrase) so that people can at least in part experience different manners of lexicalizing a familiar concept. It is this opportunity--the opportunity to experience new lexicalizations of old concepts--that makes the experience of reading well-translated foreign literature distinct from the experience of reading literature that originated in one's own language.

Conclusion

These studies provide evidence that our use and understanding of idioms is guided by conceptual structures that are not created or constrained by our *a priori* knowledge of the phrases' canonical figurative meanings (cf., Keysar & Bly, 1995, 1999). Analyzability itself can be conceptualized as ranging along a continuum (Bortfeld & McGlone, 2001). The data presented here suggest that the continuum of analyzability may well reflect a more fundamental continuum ranging between universal conceptual knowledge and culture-specific knowledge. The idioms used in the current studies can all be located roughly in the middle of the analyzability continuum; they are all somewhat metaphorical. Further work needs to be done examining how both native and non-native speakers process idioms from the rest of the analyzability continuum.

Ultimately, it seems that the aptness of the conceptual mapping rather than simply one's familiarity with the phrase is what determines an idiom's analyzability. After all, language--both one's first and any subsequently acquired--is initially encountered without the benefit of familiarity. For figurative speech in our own language, we must initially map words to conceptual structures in whatever way is meaningful. Over time, that mapping becomes well established. The sense that an unfamiliar figurative phrase is apt is something that most people who have learned a second language can appreciate, whether that learning took place during childhood. Certain phrases in a new language seem appropriate and, quite simply, make sense. It is not the case that the phrases are familiar; rather it is that they are apt, regardless of the particular surface structure the source language has given them. By looking more closely at what motivates this apparent aptness, we stand to learn much about the structures guiding our comprehension of nonliteral language.

Thus far, this discussion has been framed in terms of second language learners dealing with figurative speech in their new language. However, the research reported here can likewise be evaluated in terms of the benefits of multilingualism. Perhaps what distinguishes good language learners is how willing they are to consider an unfamiliar turn of phrase as representative of a familiar concept. All the non-native speakers who took part in these experiments were able to form images for various idiomatic phrases. But some people seemed more comfortable or adept

at mapping words to meaning than others. Perhaps individuals who have been raised in a bilingual environment are particularly flexible in their approach to this mapping process. Just as bilingual speakers have less difficulty than monolinguals in appreciating that two different words can both be symbols for the same conceptual object, they may likewise more readily appreciate that distinct turns of phrase map to the same conceptual structure. Perhaps bilingual speakers simply map words to underlying meaning more flexibly than monolingual speakers. Such a finding would add further support to claims about the universality of conceptual structures. Our understanding of language processing, in general, and idiom comprehension, in particular, stands to gain from the extension of the questions addressed here to bilingual populations. Such research would also benefit adult second language learners by enriching our understanding of the source of at least some of their difficulties.

Idioms are fascinating precisely because they are so prevalent and in such diverse forms across languages. Language processing research must continue examining how idioms can or cannot cross linguistic boundaries. Only then will we understand how it is that people who speak languages that are, quite literally, a world apart can relatively easily map the surface forms of novel phrases to their appropriate underlying concepts.

Notes

¹ Russian is a Slavic language, relatively closely related to Latvian in that both stem from the Balto-Slavic branch of Indo-European languages.

² These dialects tend to be entirely different languages, although for political and cultural reasons they are typically not referred to as such.

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., & 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., & McGlone, M. (2001). The continuum of metaphor processing. *Metaphor and Symbol*, 16, 75-86.
- Cacciari, C., & Glucksberg, S. (1995). Understanding idioms: Do visual images reflect figurative meanings? *European Journal of Cognitive Psychology*, 7, 283-305.
- Gibbs, R. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8, 449-456.

Gibbs, R. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, 9, 17-30.

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

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

Gibbs, R., & O'Brien, J. (1990). Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition*, 36, 35-38.

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.

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.

Keysar, B., & Bly, B. (1999). Swimmers against the current: Do idioms reflect conceptual structure? *Journal of Pragmatics*, 31, 1559-1578.

Kittay, E. (1997). Of "men" and metaphors: Shakespeare, embodiment, and filing cabinets. In T. Ward, S. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 375-402). Washington, DC: American Philosophical Association.

Makkai, A., Boatner, M., & Gates, J. (1995). *A dictionary of American idioms*. Hauppauge, NY: Barron's Educational Series, Inc.

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. *Psychonomic Bulletin and Review*, 6, 486-494.

Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language*, 70, 491-538.

Appendix

Idioms used for Mental Imaging Task (adapted from Gibbs & O'Brien, 1990)

Anger	Control	Secretiveness
blow your stack	crack the whip	keep it under your hat
hit the ceiling	lay down the law	button your lips
lose your cool	call the shots	hold your tongue
foam at the mouth	wear the pants	behind one's back
flip your lid	keep the ball rolling	keep in the dark
Insanity	Revelation	
go off your rocker	spill the beans	
lose your marbles	let the cat out of the bag	
go to pieces	blow the whistle	
lose your grip	blow the lid off	
bounce off the walls	loose lips	