Of black sheep and white crows: Extending the bilingual dual coding theory to memory for idioms

Lena K. Pritchett1, Jyotsna Vaid1* and Sumeyra Tosun2

Abstract: Are idioms stored in memory in ways that preserve their surface form or language or are they represented amodally? We examined this question using an incidental cued recall paradigm in which two word idiomatic expressions were presented to adult bilinguals proficient in Russian and English. Stimuli included phrases with idiomatic equivalents in both languages (e.g. “empty words/пустые слова”) or in one language only (English—e.g. “empty suit/пустой костюм” or Russian—e.g. “empty sound/пустой звук”), or in neither language (e.g. “empty rain/пустой дождь”). If idioms are stored in a language-specific format, then phrases with idiomatic equivalents in both languages would have dual representation, and should therefore be more easily recalled than phrases with idiomatic meaning in only one language. This result was obtained. As such, the findings support the dual-coding theory of memory and are also compatible with models of the bilingual lexicon that include language tags or nodes.

Subjects: Bilingualism - Second Language; Bilingualism & Multilingualism; Bilingualism / ESL; Memory

Keywords: bilingual memory; idioms; multiword units; dual-coding theory; Russian; formulaic language; collocations

ABOUT THE AUTHORS

Lena K. Pritchett received her undergraduate degree in Psychology and Russian at Texas A&M University in 2011. Her interests lie in how bilingual speakers understand figurative language.

Jyotsna Vaid is a professor of Psychology at Texas A&M University and editor of the journal, Writing Systems Research. Her work examines cognitive and neurocognitive aspects of multiple language experience, the processing of creative language (jokes, proverbs, idioms, and metaphors), and the impact of writing system properties on cognition.

Sumeyra Tosun is an assistant professor of psychology at Süleyman Şah University in Istanbul. She received her doctorate in cognitive psychology from Texas A&M University. Her research examines evidentiality in relation to memory and discourse, cognitive processes underlying humor production, and directional biases related to handedness and reading/writing direction. This research was conducted in the Language and Cognition Lab directed by Jyotsna Vaid, which examines the processing and memory repercussions of knowing two or more languages.

PUBLIC INTEREST STATEMENT

Everyday language uses many formulaic, pre-fabricated expressions such as idioms. Most previous investigations of idiom processing have been based on single language users. The present research examined bilinguals’ memory for idioms. Specifically, the study asked whether idioms that have equivalents in both languages of bilinguals are more salient, and thus more easily recalled, than those that do not have an equivalent in both languages. Our study showed that bilinguals were better at recalling idiomatic expressions for which an equivalent existed in the other language than idioms that did not have an equivalent in the other language. This finding is consistent with the predictions of the dual-coding theory of memory and suggests that when representing an idiom in memory, we preserve its form as well as its underlying meaning.
1. Introduction

Figurative language refers to expressions ranging from metaphors and idioms to jokes, proverbs, or multiword formulaic utterances, that is expressions in which the intended meaning of a phrase is not fully recoverable by considering the literal meaning of its constituent parts (see Wray, 2012, for a review). There is an extensive body of work on figurative language processing in single language users (e.g. Gibbs, 1994; Giora, 2003; Glucksberg, 1991). However, as noted by Cieślicka (2006, p. 119), “The abundance of L1 idiom processing studies has been accompanied by a regrettable lack of comparable research into the representation and processing of idiomatic expressions by second language learners” or by bilinguals, we might add.

Motivated by the need for more studies on idiom processing in users of multiple languages, the present research sought to contribute to emerging scholarship in this field (Carrol & Conklin, 2015; Heredia & Cieslicka, 2015; Lopez, 2015; Vaid, 2000; Vaid, López, & Martínez, 2015). Our study examined memory for two word idiomatic expressions. The particular question of interest in our study was how collocations with fixed, idiomatic meanings in both languages are represented relative to those with idiomatic meaning in only one of the bilingual’s languages.

Our research question built on prior work conducted within the framework of a bilingual extension of the dual-coding theory of memory (Paivio, 1990, 2014). According to this theory, items that have a dual representation will be remembered more easily than those that have a single representation in memory. Whereas previous tests of this theory have relied primarily on memory at the level of single words or pictures (e.g. Jared, Poh, & Paivio, 2013), our study examined memory for two word idiomatic phrases. Based on other research on the processing of multiword units with fixed meanings, we expected two word idiomatic expressions to behave essentially like single words in the sense that their meaning is likely stored and retrieved holistically, rather than being computed. Our study exploited the fact that some idioms only have an idiomatic meaning in one language, whereas others have idiomatic counterparts in both languages of a bilingual. We hypothesized that idioms with a shared idiomatic meaning across the two languages of bilinguals would have a heightened salience and accessibility than those with an idiomatic meaning in only one language.

More broadly, our study also speaks to the issue of the form in which linguistic knowledge is represented; that is, is the representation of an expression tied to the form and language in which the expression is encountered or is knowledge represented in a propositional, language-independent, amodal form? This issue has been studied previously in the context of memory for notation of numerical information; for example, the number “3” can be represented in digit form or in word form (“three”) (see Frenck-Mestre & Vaid, 1992; Vaid & Frenc-Mestre, 1991). Similarly, a given word can be represented in two different scripts (Park & Vaid, 1995). If notation is not preserved, we would expect memory for form to be poor. However, the findings suggest otherwise. The fact that there is better than chance recall of notation would suggest that bilinguals would be good at remembering the form (language) in which a particular word or figurative expression was encountered. At least one previous study on this issue found that this was the case for memory for the language in which a proverb was stated (Vaid & Martinez, 2001). The present study indirectly contributes to the issue of memory for format but it does so in a somewhat different way than in previous studies. The main assumption of the present study is as follows: if bilinguals pay attention to the form of an utterance, then if an idiomatic expression has a counterpart in another language of a bilingual, memory should be better for that expression than if it exists only in one language. This prediction follows directly from the dual-coding theory of memory (Paivio, 2014).

Before turning to our study, we give a brief overview of relevant psycholinguistic research on figurative language processing in single and multiple language users.

1.1. Approaches to figurative language processing

Several studies have sought to evaluate the contribution of the two cerebral hemispheres to the comprehension of conventional vs. novel figurative expressions. Using a variety of forms of
figurative language and a range of tasks, and conducted with unilaterally brain-damaged patients (e.g. Brownell, Simpson, Bihrl, Potter, & Gardner, 1990), or with optimally functioning individuals using lateralized stimulus presentation (e.g. Faust & Mashal, 2007; Mashal, Faust, Hendler, & Jung-Beeman, 2007), many of these studies argue that the right cerebral hemisphere is specialized for understanding and producing figurative meaning (e.g. Brownell et al., 1990; Klepousniotou & Baum, 2005).

However, other studies suggest no hemispheric differences in metaphor comprehension or even a left hemisphere superiority (Faust & Weisper, 2000; Lee & Dapretto, 2006; Oliveri, Romero, & Papagno, 2004). A study by Mashal and Faust (2009) found a right hemisphere advantage only when the expression was novel (and thus its meaning had to be created, rather than retrieved); when encountered a second time, the originally novel expression was processed differently and no longer showed a right hemisphere advantage. Taken together, these studies suggest that the right hemisphere may indeed have a special role in understanding and producing metaphorical language, particularly for novel metaphorical expressions. Yet, more research still needs to be done to pinpoint and confirm the actual mechanisms underlying right hemisphere involvement.

Another set of studies has used online and offline behavioral methods to test claims of different models of how figurative meanings are computed and accessed. The models vary in terms of whether they consider figurative meaning to be activated in parallel with literal meaning (temporal primacy debate) and in terms of whether different forms of figurative expressions may activate literal meaning to different degrees (compositionality debate). For example, the traditional view of language processing, known as the Standard Pragmatic View (Grice, 1975; Searle, 1975), holds that, in order to understand a figurative phrase, one initially must comprehend its literal meaning, and if it does not make sense, only then does one decode the figurative meaning. This model implies that literal meaning is activated before figurative meaning.

Other models of figurative processing propose that both literal and non-literal meanings are activated when comprehending idiomatic phrases. These models vary in terms of whether they prioritize literal or figurative meaning. One such alternative model is the Idiom Decomposition Model (Gibbs & Nayak, 1989; Gibbs, Nayak, & Cutting, 1989), which suggests that comprehending idiomatic phrases depends on the degree to which individual meanings of every word contribute to the overall understanding of the phrase. The Idiom Decomposition Model was further developed and tested by Gibbs and Nayak (1989) in their study of the syntactic behavior of idioms. They hypothesized that because some idioms can be syntactically altered and still hold their figurative meanings, (e.g. “John laid down the law” can be passivized as “The law was laid down by John”), while others tend to lose their figurative meaning when altered, e.g. “John kicked the bucket” cannot be passivized into “The bucket was kicked by John,” and the time required to process these two categories of idioms will vary. Gibbs and Nayak’s hypothesis was supported: people found it challenging to assign independent meanings to individual constituents of non-decomposable idioms. In short, these phrases required more time to process.

Another compositional model of figurative language processing emphasizes the role of literal meaning in constructing the meaning of a figurative expression (Cacciari & Glucksberg, 1991; Cacciari & Tobossi, 1988). This model suggests that a metaphorical phrase is initially processed literally, but it may eventually be replaced by a figurative meaning. Later studies also support the idea that idiom recognition is necessary prior to idiom meaning activation (Cacciari, Padovani, & Corradini, 2007).

Yet, another model takes a hybrid position. According to a study by Titone and Connine (1999), idiomatic phrases are processed as non-compositional and compositional word sequences simultaneously. More specifically, these authors argued for parallel representation of the idiom’s meaning as a whole unit along with the individual representation of its constituent parts.
These models of figurative language processing differ in terms of whether literal or figurative meanings are prioritized in processing; yet, not all models of the cognitive processing of figurative language emphasize this issue. Giora’s graded salience hypothesis (1999, 2002, 2003) proposes that salience rather than degree of figurativeness is the critical factor in determining primacy of processing. Giora defines salient meanings as the ones that “enjoy prominence due to their conventionality, frequency, familiarity, or prototypicality” (2002, p. 490). Thus, for Giora, salient meanings (whether literal or figurative) are activated initially.

1.2. Figurative and literal meanings in L2
Models introduced above demonstrate the various existing views on processing idioms in a person’s first language. What might be the case for the second language, or for individuals who acquired two languages simultaneously?

One model, developed as an extension of Giora’s salience view to the case of second language learners, was proposed by Cieślicka (2006). Termed the Literal Salience Model (Cieślicka, 2006, 2010; Liontas, 2002), this model argues that literal meaning is more salient in L2 users even if a phrase is presented in figurative context. Cieślicka (2006) employed a cross-modal lexical priming paradigm to test this idea. Her participants (Polish–English bilinguals from Poland) were auditorily presented with sentences that contained familiar idioms. While listening to each phrase, participants were visually presented with a word that either related to figurative or to literal meaning of the idiom; participants had to perform a lexical decision task on that word. Cieślicka (2006) found that priming effects obtained by targets that were related to the literal meaning were greater than priming elicited by targets related to the idiomatic meaning. Thus, literal meanings were initially accessed much faster than figurative meanings in L2 idiom processing, supporting the Literal Salience Model.

Another model, the Dual Idiom Representation model (Abel, 2003), extends to the L2 the findings of Titone and Connine (1999) in their study of figurative L1 language. Titone and Connine discovered that metaphorical phrases were simultaneously processed as non-compositional and compositional word sequences. More specifically, they argued for parallel representation of the idiom’s meaning as a whole unit along with the individual representation of its constituent parts. Similarly, the Dual Idiom Representation model in regard to figurative L2 processing postulates that decomposability determines representation of the idiom. Non-decomposable idioms require a separate lexical entry while decomposable idioms do not.

In addition to decomposability, frequency was also found to play an important role in the development of an idiom’s entry in the bilingual’s mind (Abel, 2003). In this study, a group of native German speakers and another group of non-native German speakers judged idioms’ decomposability and frequency. Native and non-native group showed the same tendency in their judgments of familiarity of idioms and its relationship to decomposability (such as non-decomposable idioms require an idiom entry, while decomposable idioms are represented by their constituent parts). Yet, in Abel’s study, frequency also played a role in constructing an idiom’s entry (such as, the more frequently a phrase is used in its metaphorical sense, the more likely it will have its own lexical entry).

Using a somewhat different approach, Martinez (2003) and Vaid et al. (2015) examined whether metaphor expressions are automatically activated in both languages. Using a bilingual adaptation of the metaphor interference task first used by Glucksberg (Gildea & Glucksberg, 1983), Martinez presented Spanish–English speakers with sentences in both their languages on which they were to make speeded true/false judgments on the basis of whether the sentences were literally true or literally false. Inserted among the sentences were metaphorically true sentences that were, nevertheless, literally false, e.g. “some cats are detectives” and “lawyers are sharks.” Martinez hypothesized that if figurative meanings are automatically activated, participants should take longer to reject such sentences as literally false, resulting in the so-called “metaphor interference effect.” The findings indicated that language proficiency plays a crucial role in determining whether metaphorical phrase is accessed or not and therefore whether metaphor interference effect takes place.
1.3. Bilingual memory research
A dominant issue underlying research in bilingualism and memory from its earliest days (e.g. Ervin & Osgood, 1954) has been to examine, through a range of experimental methods, whether word meanings in the bilinguals’ two languages are organized in a single, shared system or in separate systems (De Groot, 2002; Durgunoğlu & Roediger, 1987; Kroll & De Groot, 1997; see Heredia, 2008 for a review of bilingual memory models). The shared system view is also known in the literature as the “interdependence hypothesis” and the separate systems view is known as the “independence hypothesis.” Moreover, differences in the context of language acquisition by bilinguals were thought to favor the development of one or the other form of lexical organization; that is, an interdependent or shared system was thought to be more likely among bilinguals who acquired their two languages simultaneously and/or in similar contexts (so-called “compound” bilinguals), whereas an independent form of organization was thought to characterize bilinguals who acquired their two languages in separate contexts, typically, with the second language acquired much later than the first one, so-called “coordinate” bilinguals (see Ervin & Osgood, 1954). A number of studies have been conducted to test these hypotheses and empirical support has been obtained for each.

In an attempt to reconcile the findings, some researchers have proposed that the evidence that supports a single store view or a separate store view of memory representation may depend on the processing demands of the retrieval tasks used. That is, conceptually driven tasks such as free recall and recognition tasks, it was proposed, would more likely yield support for a shared store view, whereas data-driven tasks such as lexical decision, word fragment completion, and naming were thought to more likely support a separate store view (Durgunoğlu & Roediger, 1987).

The debate about bilingual lexical organization and the effect of particular circumstances of a bilingual’s language acquisition on lexical organization has, in recent years, given way to questions about whether words in the bilingual’s two languages are selectively or non-selectively activated. This shift in focus has arisen as online measures have increasingly come to be used in psycholinguistic research (see Kroll & De Groot, 1997). Nevertheless, the basic questions remain.

1.4. Figurative language and bilingual memory
Only a few studies to date have examined bilingual figurative language memory. Harris, Tebbe, Leka, Garcia, and Erramouspe (1999) used a cued recall memory task to assess memory for sentential metaphors (“Playful monkeys are clowns”) and similes (“Playful monkeys are like clowns”) by bilingual English–Spanish speakers in both languages. The results showed that concrete metaphors were remembered better than abstract ones and Spanish metaphors were recalled more as similes.

Of relevance to the present study is a study by Vaid and Martinez (2001), who examined Spanish–English bilinguals’ incidental recognition memory for the language of proverbs presented in a mixed language list. The aim of the study was to determine whether the wording of proverbs is retained or if proverb meaning is stored conceptually. Memory of language of presentation was tested for familiar and less familiar proverbs in each language under different encoding conditions. The results showed that bilinguals were good at recognizing the language in which the proverb had been presented, suggesting that they retained the wording of the proverbs. If proverbs’ meanings are stored in a conceptual mode, participants should have been poor at detecting the initial language in which the proverb had been presented.

1.5. The present study
Given the ubiquity of figurative expressions in everyday language, there is an urgent need for more studies of how multiword units, formulaic expressions, and other forms of figurative language are comprehended and organized in memory in users of more than one language. The present research was conducted with this in mind.
1.6. Dual-coding model

Our starting point was a bilingual extension of the dual-coding model of memory developed by Paivio and Desrochers (1980; see also Paivio & Lambert, 1981; Paivio, 1990). The original version of Paivio’s dual-coding model argued that lexical entries have two interconnected mental representations: a symbolic representation and an imaginal representation. Several studies support the claim of the model that memory for pictorially encoded stimuli should be superior to that for verbally encoded stimuli. The model has also led to a veritable cottage industry of research on the advantage in recall for concrete over abstract words, as concrete words presumably tap into both the symbolic and the imaginal representations. This “concreteness effect” is a robust finding in the bilingual memory literature as well (see De Groot, 2002).

The bilingual adaptation of the dual-coding model, initially proposed by Paivio and Desrochers (1980), argued for a language-free imaginal representation and two symbolic representations, one corresponding to each language (see Figure 1). The two symbolic (or verbal) systems are separate but linked by connections. As Heredia (2008, p. 51) notes in his review of bilingual memory models, the bilingual dual-coding model, unlike previous models, “is formulated well enough so as to generate specific predictions about bilingual memory.”

The model proposes that connections between entries across the two verbal systems are stronger than those within each system. As such, the model predicts that memory should be better for translation equivalents than for words that are synonyms within a language. Studies by Paivio and Lambert (1981), Paivio, Clark, and Lambert (1988) tested this model using an incidental memory procedure and found empirical support for the view that retrieval is better for words that were pictorially encoded than for words that were verbally encoded (consistent with the general dual-coding principle of superior retrieval for imaginally represented mental representations). Moreover, it was discovered that words that had been translated in the acquisition phase showed better recall than words that had been copied or paraphrased in the same language (Vaid, 1988). Thus, retrieval was better when the task required activation of entries in different languages than when it required activation of entries in a single language.
1.7. The present study

The present study examined whether memory for two word idiomatic expressions, such as “blue moon” (meaning, “a rare occurrence”) is represented in a language-specific format or in an amodal, conceptual form. To test this question, an incidental cued recall test was administered. Stimuli were adjective–noun, non-decomposable, idiomatic phrases of the following types: idioms with idiomatic equivalents in both languages of the bilinguals (Russian and English), idioms that had an idiomatic meaning in only one language (English or Russian), and novel two word phrases that had no prior idiomatic meaning in either language. According to the bilingual dual-coding model (Paivio & Desrochers, 1980), memory should be better for items that are represented twice in the lexicon. Thus, it was hypothesized that idiomatic phrases that have a shared idiomatic meaning in both languages of bilinguals will show a higher level of recall than phrases that have an idiomatic meaning in only one of the languages or in neither language.

An additional question we examined was whether retrieval of phrase meaning would be greater when there was a match between the language of the retrieval cue (which was the first word of the two word phrase) and the language in which the phrase was initially presented in the encoding phase. Based on the encoding specificity principle (Tulving & Thomson, 1973), we expected this to be the case.

2. Method

The procedures in this study were approved by the human subjects protection committee of the Institutional Review Board of the university where it was conducted and followed the ethical guidelines of the Helsinki Declaration of 1975, as revised in 1983.

2.1. Participants

Adult speakers of English and Russian were recruited for the study from a large university and surrounding community in the southwestern region of the USA. Participants ranged in age from 17 to 30 with a mean age of 27. They were administered a 12-item language background questionnaire developed for this study which contained items about the age at which they had acquired each language, their current pattern of use of each language, and their self-ratings of proficiency in each language. To be eligible to participate in the study, participants had to rate themselves as at least 4 out of 7 on speaking, reading, writing, and comprehending each language, where 1 = very little knowledge and 7 = like a native speaker of the language. Of the 25 participants recruited for the study, 3 did not meet this criterion and were excluded from the analyses.

Of the remaining 22 participants, all were late bilinguals, having learned their second language after the age of 6. They included 15 participants whose first language was Russian (11 women, mean age of 29) and 7 whose first language was English (6 women, mean age of 23). The native English speakers were undergraduate students majoring in the Russian language at the university where the study was conducted; native Russian speakers were members of the Russian immigrant community settled in the vicinity of the university, and had lived in the USA for an average of eight years.

2.2. Materials and procedure

Stimuli consisted of 96 adjective + noun collocations in English and their Russian translation equivalents. The stimuli were constructed by combining each of the 24 adjectives with 4 different nouns, such that the resulting phrases had a commonly known figurative meaning in English and in Russian translation (Fig-Both), a figurative meaning only in English (Fig-English), a figurative meaning only in Russian (Fig-Russian), and a novel figurative meaning that did not exist in either language (Fig-Neither). For example, “blue blood/голубая кровь” has a figurative meaning in both languages, “blue moon/голубая луна” only has a figurative meaning in English, “blue distances/голубые дали” only has a figurative meaning in Russian, and “blue smell” has no known idiomatic meaning in either language. A list of the stimuli is available on request.
Phrases were pre-tested with native speakers of each language to ensure that the intended figurative meaning was recognizable in the respective languages. A given noun was paired only once with an adjective (e.g. the noun “blood/кровь” was only used in combination with “blue/голубой” and no other adjective). Thus, each of the 24 adjectives (e.g. “blue”) was paired with 4 different nouns (blood, moon, distance, and smell), resulting in a total of 96 phrases in each language.

Participants were tested individually or in small groups of two to three people. The study was conducted in two phases. In the acquisition phase, participants were shown a list of the 96 phrases presented in English or Russian in a fixed random order. Separate lists were prepared to counterbalance the language of presentation of a given phrase across participants. Participants were not told that they would be tested on their memory of the phrases; instead, they were told the study involved their judgments of the pleasantness of the phrases. On reading each phrase, participants were asked to rate it in degree of pleasantness using a five-point scale, with 1 being “very unpleasant” and 5 being “very pleasant.” For example, “dirty joke” implies an unpleasant meaning and could be rated as 1, while “warm greeting (or тёплое приветствие) usually has a positive connotation and could be rated as 5. Participants were informed that some phrases might not make sense to them (e.g. “blue smell,” or “голубой запах,” “rich parachute/богатый парашют,” or “dirty cough/грязный кашель”) and were advised to rate those phrases to the best of their ability.

Upon completion of this task (the acquisition phase), a language background questionnaire was administered. Aside from its role in classifying the bilinguals, the questionnaire served a double purpose of introducing some delay before the test phase was administered as it took approximately 5–10 min to complete.

The test phase was then administered. This consisted of a cued recall task. It came as a surprise to participants. In this task, 24 adjectives that had previously been shown in the acquisition phase were again presented, and for each adjective (now termed “cue”), participants were to write down from memory the four nouns that had been presented with it earlier. Importantly, half of the adjective cues in the test phase appeared in the same language as at original presentation, whereas the remainder appeared in translation (i.e. a phrase that had previously been presented in English was now presented in Russian translation, and one that had initially been presented in Russian was now presented in English translation). This was counterbalanced across participants.

Participants were to write down the noun in the language in which they thought it had been presented earlier. In addition, they were asked to rate their confidence on a five-point scale, with 1 being “not at all confident” and 5 being “very confident” of their response. Participants were required to take the language of recall into consideration when doing the confidence rating; that is, their confidence was to reflect both a particular word and the language in which the word had been presented initially.

2.3. Design
The design was a 4x2x2x2 mixed factorial model, with the within-subjects variables being Phrase Type (Figurative meaning in both languages, Figurative in English only, Figurative in Russian only, or Figurative in neither), and the between-subjects variables being Language at Initial Presentation (English or Russian), Language of the Cue in the Test Phase (Same or Different), and Participant’s Native Language (English or Russian).

3. Results
A four-way analysis of variance was conducted on three response measures: mean, accuracy of recall considered in two ways, and mean confidence ratings. The accuracy data were analyzed in the following ways. The first analysis considered all responses generated by participants without regard to whether they were in the correct language (i.e. the same language in which the phrase had been presented at initial presentation). In this analysis, if a participant saw “blue moon” in the acquisition phase but recalled it as “лунa” (Russian word for “moon”), it was still considered a correct answer.
The second accuracy analysis looked only at responses that were generated in the correct language (i.e. the language in which the phrase had initially appeared).

Mean percent accuracy of recall per phrase type is summarized in Table 1.

### 3.1. Accuracy of word recall without regard to language accuracy

For this accuracy analysis, there was a significant main effect of Phrase Type, $F (1, 20) = 27.77$, $p < .001$, which indicated that Fig-Both phrases were remembered significantly more accurately than any other category of phrases. Comparison of the means revealed that Fig-Both phrases were remembered significantly more accurately than Fig-English phrases, $t (21) = 2.93, p < .01$, Fig-Russian phrases, $t (21) = 3.43, p < .01$, and Fig-Neither, $t (21) = 5.59, p < .001$. Additionally, Fig-English phrases were recalled significantly better than Fig-Neither, $t (21) = 3.23, p < .01$, and Fig-Russian phrases were remembered significantly better than Fig-Neither, $t (21) = 2.24, p < .05$. Thus, overall, nonsense phrases seemed to be the hardest to retrieve; phrases that had figurative meanings in both languages were more easily retrieved that those that only had a figurative meaning in one language (see Figure 2).

| Table 1. Mean confidence ratings (out of 5) and mean number of items recalled (and standard deviations) by phrase type |
|-------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|
| Accuracy regardless of language at presentation | Accuracy with regard to language at presentation | Confidence ratings | Accuracy regardless of language accuracy |
| Figurative in both languages | Figurative in English only | Figurative in Russian only | Nonsense |
| Accuracy regardless of language at presentation | 5.77(3.32) | 3.91(2.45) | 3.50(1.99) | 2.41(2.46) |
| Accuracy with regard to language at presentation | 4.95(2.98) | 3.45(2.34) | 2.95(2.08) | 2.18(2.32) |
| Confidence ratings | 3.10(1.22) | 2.74(1.31) | 2.36(1.12) | 1.70(1.30) |

*Out of a maximum of 24 items per phrase type.

Figure 2. This figure demonstrates the mean overall recall accuracy as a function of phrase type in the condition without regard to language of initial presentation.
3.2. Accuracy of word recall and language recall

For this accuracy analysis, only phrases that were remembered in the language of acquisition were considered accurate. Thus, for example, if a participant saw “blue moon” in the acquisition task, but recalled it as “лунa,” it was not considered a correct answer. The analysis revealed a Phrase Type main effect, $F(1, 20) = 27.06, p < .001$. As before, Fig-Both phrases were more likely to be recalled than the other three types (see Figure 3).

A post hoc analysis showed similar relationships between the different categories of phrases as noted above: Fig-Both phrases were recalled significantly more accurately than Fig-English phrases, $t(21) = 2.54, p < .05$, Fig-Russian phrases, $t(21) = 3.51, p < .01$, and also more accurately than Fig-Neither phrases, $t(21) = 5.27, p < .001$. Furthermore, Fig-English phrases were remembered significantly better than Fig-Neither, $t(21) = 2.70, p < .05$. But unlike in the previous analysis, Fig-Russian phrases were not remembered significantly better than Fig-Neither, $t(21) = 1.86, p = .077$. Overall, phrases that were novel in both languages were recalled the worst while phrases with figurative meanings in both languages enjoyed an advantage in recall.

In addition to the main effect of Phrase Type, there was a near significant interaction between Language at Acquisition and Cue Language, $F(1, 20) = 4.07, p = .057$ (see Figure 4). Post hoc analyses showed that participants tended to recall more phrases when they originally saw them in English and were presented with an English cue at recall time ($M = 4.05, SD = 2.26$) than when they initially saw them in Russian and were presented with a Russian cue ($M = 2.77, SD = 2.94$), $t(21) = 1.84, p = .08$. Additionally, participants tended to recall more phrases they initially saw in Russian when presented with an English cue at recall time ($M = 3.86, SD = 3.47$) than when presented with a Russian cue at recall time ($M = 2.77, SD = 2.94$); $t(21) = 1.929, p = .067$.

3.3. Analysis of confidence ratings

A Phrase Type main effect was obtained, $F(1, 24) = 25.823, p < .001$, indicating that participants were significantly more confident when remembering phrases that shared a figurative meaning in both languages than phrases with no recognizable figurative meaning in either language (see Table 1).
Participants were significantly more confident when recalling Fig-Both phrases ($M = 3.10$, $SD = 1.22$) than Fig-Neither phrases ($M = 1.70$, $SD = 1.30$), $t(21) = 4.28$, $p < .001$, or Fig-Russian phrases ($M = 2.36$, $SD = 1.12$), $t(21) = 2.27$, $p < .05$. Additionally, participants were significantly more confident when remembering Fig-English phrases ($M = 2.74$, $SD = 1.31$) than Fig-Neither phrases ($M = 1.70$, $SD = 1.30$), $t(21) = 4.01$, $p < .001$. When recalling Fig-Russian phrases ($M = 2.36$, $SD = 1.30$), participants felt more confident than when recalling Fig-Neither phrases ($M = 1.70$, $SD = 1.30$), $t(21) = 2.07$, $p = .051$. Overall, participants felt more confident when recalling Fig-Both phrases and least confident when remembering Fig-Neither phrases (see Figure 5). There was no interaction effect.

To rule out some alternative explanations for the discovered tendencies, we collected data about the phrases’ perceived frequency (see Table 2) and imageability (see Table 3) from monolingual speakers of each language (these were individuals who had not been tested in previous tasks and know only one language—English or Russian).

A $2 \times 4$ repeated measures ANOVA was conducted on the frequency ratings reported by monolinguals (i.e. how often they have encountered each of the 96 phrases). A Phrase Type main effect was found, $F(1, 10) = 245.26$, $p < .001$ as was an interaction of Phrase Type and Native Language, $F(1, 10) = 10.16$, $p < .01$ (see Figure 6). The interaction effect indicated that for native Russian speakers, there was no difference in perceived frequency of phrases with figurative meanings in both languages and phrases with figurative meanings only in Russian. Similarly, for English monolinguals, there was no difference in perceived frequency of phrases with figurative meanings in both languages and phrases with figurative meanings only in English.

Analysis of imageability of the phrases (i.e. how easy it is to visualize the meaning of each of the 96 phrases) showed a main effect of Phrase Type, $F(1, 10) = 38.9$, $p < .001$, and a Phrase Type by Native Language interaction effect, $F(3, 30) = 11.87$, $p < .001$ (see Figure 7). The interaction effect indicated that for native Russian speakers, there was no difference in perceived imageability of phrases with figurative meanings in both languages and phrases with figurative meanings only in...
Russian. Similarly, for English monolinguals, there was no difference in perceived imageability of phrases with figurative meanings in both languages and phrases with figurative meanings only in English.

These results allow us to conclude that stimuli belonging to the Figurative-Both condition were not intrinsically more familiar or more imageable than stimuli belonging to the Figurative in the native language conditions, as judged by native speakers of each language. Therefore, the differences in retrievability observed in the present study are not due to any greater familiarity or imageability of phrases in the Figurative-Both condition.

Table 2. Mean ratings of phrase subjective frequency by phrase type (7 pt. scale)

<table>
<thead>
<tr>
<th></th>
<th>English monolinguals (n = 7)</th>
<th>Russian monolinguals (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figurative in both languages</td>
<td>4.44</td>
<td>4.43</td>
</tr>
<tr>
<td>Figurative in English</td>
<td>4.42</td>
<td>2.39</td>
</tr>
<tr>
<td>Figurative in Russian</td>
<td>2.35</td>
<td>4.71</td>
</tr>
<tr>
<td>Nonsense</td>
<td>1.04</td>
<td>.95</td>
</tr>
</tbody>
</table>

Table 3. Mean ratings of phrase imageability by phrase type (7 pt. scale)

<table>
<thead>
<tr>
<th></th>
<th>English monolinguals (n = 7)</th>
<th>Russian monolinguals (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figurative in both languages</td>
<td>4.49</td>
<td>4.98</td>
</tr>
<tr>
<td>Figurative in English</td>
<td>4.54</td>
<td>2.91</td>
</tr>
<tr>
<td>Figurative in Russian</td>
<td>3.86</td>
<td>4.62</td>
</tr>
<tr>
<td>Figurative in neither (Nonsense)</td>
<td>2.69</td>
<td>1.25</td>
</tr>
</tbody>
</table>
4. Discussion

In this study, we examined the recall of idiomatic expressions that had a relatively fixed structure and established figurative meanings in either one or both languages of bilinguals. To the extent that
phrases with a shared figurative meaning in both languages may be considered to have dual entries in the mental lexicon compared to those in which the figurative meaning existed in only one of the languages, we hypothesized that recall would be superior for the former than for the latter phrase type. This hypothesis was motivated by the notion of bilingual memory developed in the framework of a dual-coding approach, according to which words that have shared meanings in both languages of bilinguals will be better retrieved than those that only have meanings in one language (Paivio, 1990, 2014).

Despite the fact that the recall task was very demanding and, therefore, that overall percent recall was around 30%, we still observed a consistent advantage in recall of phrases that had a shared figurative meaning in both of the bilinguals’ languages. As such, our hypothesis was supported.

4.1. Accuracy of recall independent of language of stimulus in acquisition phase
Bilinguals were significantly better at recalling phrases with figurative meaning in both languages than phrases in any other condition (i.e. figurative in one of their languages or figurative in neither language). Furthermore, participants were significantly better at remembering phrases with figurative meaning in only one language than nonsense ones. This trend was also predicted because nonsense metaphors are novel and do not have an entry in the mental lexicon. These findings are consistent with Paivio and Desrochers’s (1980) dual-coding theory which proposed that lexical entries with dual representations are more likely to be remembered than those that are only represented once (see also Vaid, 1988). Likewise, nonsense phrases are the hardest ones to retrieve since they presumably do not map onto any existing representation in the lexicon.

4.2. Accuracy of recall with regard to the language of the stimulus in the acquisition phase
Similarly, when we only considered phrases recalled in the language in which they were first presented, recall was highest for phrases with figurative meanings in both languages. Even though the overall level of recall was much lower in this way of analyzing the data than in the one reported in the previous section, recall of phrases figurative in one language was still higher than recall of nonsense ones.

Additionally, there was one near significant interaction between Input language and Cue language (p = .057). The interaction suggests that the condition yielding the highest recall was when the phrase language was English and the cue language was also English (on average, 4.05 items recalled); the next highest condition was when the phrase language was Russian and the cue language was English (on average, 3.86 items recalled). In general, recall was poorer when the language of the cue at recall was Russian. Thus, our expectation that recall would be higher when language at the acquisition phase and cue phase was the same was only partially supported; English language cues (for phrases that appeared initially in either English or Russian) seemed to have a beneficial effect in recall. It would appear that participants were more comfortable with the English language than with Russian, even though more than half of them had acquired English as a second language. A possible explanation of this phenomenon could be the fact that participants had lived in the USA for some time and were therefore more used to operating in English. To test this interpretation of the findings, a follow-up study should be conducted with bilinguals living in Russia to determine if the dominant language of the environment influences people’s recall accuracy.

Our findings also showed that participants’ level of confidence was highest for phrases that were of the Figurative-Both type, relative to Figurative-Russian and Figurative-Neither phrases. Participants were least confident about nonsense phrases and significantly more confident about Figurative-English and Figurative-Russian phrases than about nonsense ones. Thus, not only did participants show more accurate recall of phrases that had figurative meanings in both languages, they were also more confident about encountering them previously.
4.3. Language acquisition background
Participants were drawn from two different backgrounds: one in which people acquired English as their second language in their teen years, the other one in which people learned Russian as their second language in college. The variable representing their native language was controlled for and was not found to be statistically significant. That is, participants’ language acquisition background did not affect their accuracy of recall of the different phrase types. It is possible that with a larger sample differences may have emerged.

4.4. Phrase type frequency and imageability
To examine if the findings could be attributable to other factors such as differences in familiarity or imageability of the phrases, an analysis of frequency and imageability ratings by a sample of monolingual Russian and monolingual English speakers was conducted but showed no evidence for this alternative potential explanation. The results for both dimensions showed no difference in ratings for the figurative-in-both items and the figurative-in-their-native-language items, thereby ruling out possible differences in perceived frequency and/or imageability of the different phrase types as an alternative explanation of the observed difference in recall.

One possible reason for why recall was so low (particularly for phrases in Russian) may be that participants were less familiar with the idiomatic meanings of some of the Russian phrases. We conducted a post-test check with a subset of participants and found that participants generally were familiar with the phrases in both languages. A more likely reason for the low overall level of recall was that the task was made very difficult by the fact that there were simply too many items to be recalled. Use of a recognition procedure rather than a recall procedure would probably have resulted in better performance. Nevertheless, despite the low level of overall recall, our findings showed a significant difference in relative recall by phrase type, in support of our prediction.

5. Conclusion
Taken together, and consistent with Paivio and Desrochers’s (1980) dual-coding theory of bilingual memory, the present findings suggest that items that have a dual representation in memory (as is presumably the case for the phrases that have a figurative meaning in both languages) yield better retrieval than items that have a single representation in memory (as is presumably the case for phrases that have a figurative meaning in only one of the languages).

In terms of their relevance to the debate on whether bilinguals’ linguistic systems are separate or integrated (the so-called independence vs. interdependence issue), the findings do not rule out either view. They are clearly compatible with an independence view, which holds that there are two separate representations of words or idiomatic phrases, one for each language; in this view, the idioms that have meaning in both languages would be dually represented. However, the findings may also support an interdependence or common store view in which language information is stored as a tag attached to the word (or idiomatic phrase) and is retrieved along with it. An example of such a model is the BIMOLA model which posts a language node which preactivates all the entries sharing a particular language tag (Lewy & Grosjean, 2008). Another model that also posits language nodes is the Bilingual Interactive Activation model (see Dijkstra & van Heuven, 1998). Both these models would be able to accommodate the present findings as well.

Given that there are so few existing studies on memory for figurative language in bilinguals, converging evidence from bilingual adaptations of other experimental approaches, such as studies by Schweigert (2009) of idiom comprehension as a function of multiple presentations on the ratings and memorability of figurative phrases, and priming studies (e.g. Cieślicka, 2006) will be needed to add support to our findings. It is hoped that the findings from the present research will lead to more investigations into figurative language comprehension and representation in bilinguals.


Hillsdale, N.J: Lawrence Erlbaum.


http://dx.doi.org/10.1093/acprof:oso/9780195136166.001.0001


http://dx.doi.org/10.1093/acprof:oso/9780195136166.001.0001


Hillsdale, N.J: Lawrence Erlbaum.


http://dx.doi.org/10.1093/acprof:oso/9780195136166.001.0001


