Published in *Metaphor and Symbol*, 1999, 14, 1-16.

Running Head: Memory for Similes and Metaphors

Monolingual and Bilingual Memory for English and Spanish Metaphors and Similes

Richard Jackson Harris

Kansas State University

Michael R. Tebbe

University of Kansas

Gary E. Leka

University of Texas-Pan American

Reina Coral Garcia and Raquel Erramouspe

Universidad Católica del Uruguay

Address correspondence to:

Richard J. Harris

Department of Psychology, Bluemont Hall

Kansas State University

1100 Mid-Campus Drive

Manhattan KS 66506-5302 USA

ph (785) 532-0610 fax (785) 532-7004

e-mail: rjharris@ksu.edu

Abstract

Four experiments tested cued-recall memory for simple sentential metaphors (Playful monkeys are clowns) and similes (Playful monkeys are like clowns) in Spanish and English. Overall, concrete metaphors and similes were remembered much better than abstract ones. In the all-English Experiment 1, both metaphors and similes were usually recalled as metaphors, while in the all-Spanish Experiment 2 both metaphors and similes were usually recalled as similes. Experiments 3 and 4 used mixed-language lists (Spanish and English) and bilingual participants fluent in both languages. In these studies the trend of recalling both English figurative types as metaphors essentially went away, although there was still some tendency to recall Spanish metaphors as similes. Results were interpreted in terms of possible language-specific production biases, in the context of Glucksbergs property-attribution model and Grosjeans bilingual-monolingual mode distinction.

Memory for Concrete and Abstract Metaphors and Similes in English and Spanish

Two of the most common figures of speech are similes and metaphors, which differ only in the presence of the word **like** in the simile**.** Although there has been considerable psycholinguistic research on metaphor in recent years (see Gibbs, 1994, for a thorough review), there has been much less on similes. Although similes are usually contained in taxonomies of figurative language, they may in some sense be considered literal comparisons as well. Stating that Juliet is like the sun could be perfectly true literally, depending on the basis for the comparison. It is not literally anomalous or false like Her eyes are pearlsor The ivy cuddled up to the window*.* However it is classified, the simile provides an interesting contrast to the metaphor because it is so similar structurally and semantically.

Many, if not most, models of metaphor in psychology and elsewhere consider metaphors, like similes, as statements of comparison of the topic and the vehicle. Many years ago, Aristotle noted in hisPoeticsand Rhetoric that a good metaphor reflects the speaker or writers perception of similarity in otherwise dissimilar things. In this view metaphors are thus basically comparisons, and they are more powerful than similes because they are more condensed (Gibbs, 1994). Many classic models in philosophy and later psychology posited some sort of substitution or comparison model (e.g., Black, 1962; Richards, 1936), sometimes involving the transfer of a semantic feature or an imbalance of salience of properties of the topic and vehicle.

Although not all models of figurative language processing make comparative predictions about metaphors and similes, most that do (e.g., Miller, 1979) consider similes a more basic, direct, and explicit comparison than metaphors, because, after all, the comparison in a simile is unambiguously signaled by the presence of the word **like.** Metaphors are thus abbreviated similes which must be recoded to similes as part of the comprehension process..

Other models, notably Glucksbergs (1991; Glucksberg & Keysar, 1990) attributive categorization (class-inclusion) model, say that metaphors are the more direct, psychologically simpler statement. Glucksberg argues that a metaphorical statement asserts a class inclusion. For example, saying the job is a jailasserts that the job belongs to a category of unpleasant entities denoted by jail*.* However, a simile actually requires an additional processing step because the comparison is qualified by **like**and does not so directly assert a class-inclusion statement *.* Metaphors are not implicitly transformed into similes to be understood but rather are understood as class-inclusion statements (Glucksberg, McGlone, & Manfredi, 1997; McGlone, 1996).

Most models of metaphor have been developed in and for English. However, Johnson (1995; see also Johnson, 1996) conducted two studies, identical except that one study was conducted in English with English speakers and the other in Spanish with Spanish speakers. A computer presented participants with a list of sentences, which contained either metaphors (Cigarettes are time bombs*)* or similes *(*Cigarettes are like time bombs*)*. The time was measured from stimulus onset until the participant pressed the space bar to indicate they had understood the sentence. There then followed a task judging if a second sentence logically followed from the first. At the end of the whole list participants were given a surprise cued-recall memory test, where they were presented with the sentential subject and linking verb as a stem and had to fill in the rest of the sentence.

Johnson found that metaphors were comprehended faster than similes in both English and Spanish, and that concrete sentences were comprehended faster than abstract sentences. Due to results from some literal control sentences of the same length, Johnson concluded that this difference was not merely due to the longer sentence length of similes. Thus, comprehension time data supported the idea that metaphors are more primary, more basic psychologically, a finding consistent with Glucksbergs attributive categorization model, whereby metaphors state a class-inclusion relationship, but similes must be translated into metaphors and thus take longer.

However, in terms of memory, the two languages showed a strong and curious asymmetry in Johnsons studies. In English, there was a strong tendency to recall similes as metaphors, while the reverse seldom occurred. In Spanish, however, the opposite was true. Metaphors were most often recalled as similes but similes very seldom recalled as metaphors.

The present research contained two pairs of studies that further examined this strong language asymmetry in memory identified by Johnson. Experiments 1 and 2 were designed as extensions of Johnson (1995, 1996) as memory-only studies. They also more systematically controlled the numbers of concrete and abstract sentences. Experiments 1 and 2 were identical in procedure and design except that Experiment 1 was conducted entirely in English in the United States and Experiment 2 was conducted entirely in Spanish in Uruguay. Their data were analyzed together considering language as a between-subjects factor. They are presented as Experiment 1" and Experiment 2" below because they were conducted by entirely different experimenters using different pairs of stimulus tapes. Experiments 3 and 4 used the same procedure but with mixed-language (Spanish-English) lists of metaphors and similes. The participants in these studies were bilinguals fluent in English and Spanish. Experiments 3 and 4 differed only in the length of the stimulus lists.

Experiment 1: Method

Participants. The participants were 110 native-English-speaking introductory psychology students from a large Midwestern U.S. university with open admissions. They received course credit for participation and were tested in small groups.

Materials. Two lists of sentences were constructed, similar to those of Johnson (1996). Each list contained 32 simple sentences of the form (The) (Adj) N is/are (like) (Adj)N*.* Half of the sentences contained metaphors and half contained similes, with the two types differing only in the presence of the word **like** in the simile sentences. The two lists were identical except that any sentence which was a metaphor in one list was a simile in the other list and vice versa (e.g., Playful monkeys are clowns, Playful monkeys are like clowns*)*. Half of the metaphors and half of the similes in each list were concrete, e.g.,Playful monkeys are (like) clowns, and half contained abstract subjects, e.g., Short tempers are (like) smoking volcanoes. All vehicles were concrete.

Procedure. The participants were told to listen to a list of figurative sentences, because they would be tested on their memory for those sentences later. The experimenter then started the tape player. The sentences on the tape were read by a male voice speaking slowly and carefully and allowing 3-4 seconds between sentences. After the tape was over, the answer sheet was handed out. For each sentence in order the subject and linking verb was given (e.g., Playful monkeys are *\_\_\_\_\_\_\_)*, with instructions to fill in the rest of the sentence as best you can remember. Participants were given as much time as needed.

Experiment 2: Method

Experiment 2 was an exact replication of Experiment 1 but with all the materials in Spanish. The sample consisted of 78 native-Spanish-speaking students from a private university in Montevideo, Uruguay. The materials from Experiment 1 were translated into Spanish and confirmed by a native speaker of Uruguayan Spanish. Some minor changes had to be made in the content of a few sentences to make them reasonable in Spanish in that particular culture. The stimulus tapes were then re-recorded by a native speaker of the local dialect of Spanish. Almost all of the students had studied another language (most often English), a situation typical of university students in most countries outside the U.S.

Experiments 1 and 2: Results

Scoring

The responses were scored blind to figurative language condition in two ways. It was noted whether the recall was verbatim (exact words of input sentence), gist (paraphrase or incomplete report of words in input sentence but retaining the basic semantic content), or error (substantially changing semantic content). Secondly, it was noted whether the recalled sentence was in the form of a simile (i.e., with the presence of **like***)* or a metaphor (no **like***).* Because of the presence of the sentential subject and to be verb as the stem on the answer sheet, all responses (other than omitted ones) were constrained to be in the form of metaphor or simile. A final category of response were omitted responses, in which case the line was left blank or filled in with a question mark. For purposes of data presentation and analyses below, the verbatim and gist responses were combined into a single category. The distribution of these two types of responses did not systematically differ across the figurative language or concreteness categories.

Data Analyses

Per cent of recalls in each of the four response categories (metaphor, simile, error, and omit) for Experiments 1 and 2 appear in Table 1. Only effects with p<.01 are considered. In general, results were very systematic and clear. In English, both similes and metaphors tended to be recalled as metaphors, while in Spanish both tended to be recalled as similes. There were very few other content errors, although there were also large numbers of omitted responses, especially on abstract sentences. While analyses of variance were performed on results (described below), a greater reliance was placed on measures of effect-size and confidence-intervals in interpreting results, considering the current concern about the misuse of null hypothesis statistical testing (e.g., see Hunter, 1997; Abelson, 1997; Estes, 1997).

Omits and errors. Before discussing the two major categories, let us briefly discuss the error and omit data. For the errors, the number of such responses was always very small (1-2%); thus participants generally recalled the content more or less correctly if they recalled it at all. The very large number of omits suggests the task was quite difficult, especially for the abstract sentences, which were forgotten 60-70% of the time. However, the remaining data produced such strong effects that that this great amount of missing data was not a major measurement problem. A repeated-measures analysis of variance on the number of omitted responses showed main effects of all three variables, but by far the largest effect size was for concreteness, F(1,186)=748.37, MSe=1.89, partial eta squared=.80, with abstract sentences being forgotten much more often than concrete ones. There were also weaker effects showing that Spanish sentences were forgotten more often than English ones, F(1,186)=8.69, MSe=1.37, partial eta squared=.04, and similes forgotten more often than metaphors, F(1,196)=74.27, MSe=4.89, partial eta squared=.08. Finally, there were two double interactions with language factor, which reflected the fact that the omitted responses occurred more often for Spanish similes than metaphors (but not so for English) and that English concrete sentences were forgotten less than Spanish concrete sentences, but this language difference did not exist for the abstract sentences.

Semantically correct recalls. Two three-way mixed analyses of variance were performed on the number of sentences recalled as (a) metaphors and (b) similes, combining responses recalled either in verbatim or gist form. The figurative language type and concreteness factors were both repeated measures, while language was a between-subjects factor. Confidence intervals around the means for each cell were calculated. Both analyses showed main effects of all three factors and double interactions of concreteness with each of the other factors. However, the bulk of the variance was accounted for by the language and concreteness main effects, with very large partial eta squared values for these Fs (see Table 2). Many more English sentences were recalled as metaphors, and many more Spanish sentences were recalled as similes. Concrete sentences produced more of both types of recalls than did abstract sentences.

Experiments 1 and 2: Discussion

The results from Experiments 1 and 2 were completely consistent with the memory results of Johnson (1995) but were obtained in a procedure without a prior sentence-judgment reaction-time task. The fact that the size of the metaphor-simile asymmetry was at least as large as that in Johnson (1995) is important, in that Johnsons Spanish-speaking sample were all Spanish-English bilingual students living in the U.S., while those in the present Experiment 2 were tested in their own Spanish-speaking culture in their strongly dominant or only language. Apparently those diverging factors were not the critical ones.

Although memory for concrete material has long been known to be better than memory for abstract material, the magnitude of the difference in the present study is still very striking. The present task apparently somehow enhanced this natural advantage for concrete material, all the more striking given that the abstract sentences were not totally abstract; the vehicle was always concrete. In the abstract sentences that were remembered, however, the results from the concrete sentences were replicated. Although there is a remote possibility that these results could be due to a greater predictability of the concrete sentences, the very small number of error responses suggests that few participants were using a guessing strategy.

The strong language asymmetry in figurative type recall was initially unexpected and was not immediately interpretable. On the one hand, there may be differential stylistic production constraints in the two languages, with the metaphorical form somehow being more natural or frequent or preferred in English, while the simile is in Spanish. Several Spanish speakers and language teachers have commented that they believe similes are more common in Spanish than in English, but there is no known frequency data to support or refute those impressions.

To further explore this curious language asymmetry in figurative language memory, the next two studies varied language as a within-subjects factor, using mixed-language lists and bilingual participants. The method for Experiments 3 and 4 was identical except for Experiment 4 using a shorter task (16 sentences instead of 32).

Experiments 3 and 4

Method

Participants. The participants in Experiment 3 were 100 Spanish-English and English-Spanish bilinguals (69% women, 31% men) and in Experiment 4 were 50 such bilinguals (59% women, 30% men, 11% unreported). All were from undergraduate psychology classes at the University of Texas-Pan American, a 90% Mexican-American university in the Rio Grande Valley in far southern Texas. In this region both Spanish and English are used on a daily basis by most people. Of the combined sample, 80% were born in the US and 17% in Mexico. Spanish was the native language of 54%, while English was the native language of 41% (11% other or unreported). On a 7-point scale of self-rated general language fluency, the mean ratings were 5.98 in English and 4.49 in Spanish; thus the sample rated themselves as somewhat more fluent in English but reasonably fluent in both.

Materials and Design. The lists were the same as used in Experiments 1 and 2, except that half of the sentences in each list were in Spanish and half in English. Because of the addition of the language variable, there were four lists instead of two. These lists were completely counterbalanced for language and figurative type (metaphor or simile). Thus each list had 16 metaphors and 16 similes. Half of each of these were in English and half in Spanish. Half of each of those eight were concrete and other half abstract, e.g., Playful monkeys are like clowns, Las mentiras improvisadas son bumerangs, Los submarinos son como ballenas, Shameless ignorance is soft clay, etc. There were 4 blocks of eight sentences each, one sentence in each language by figurative type by concreteness combination. Within each block the order of the sentences was random but constant across participants.

Procedure. The procedure was identical to that of Experiments 1 and 2, except that all instructions were in both languages. The answer sheets presented each stem in both languages, written one on top of the other and counterbalanced for order of language. The instructions appeared in both languages and said to fill in the blank in English or Spanish, according to the form of the original sentence. This format allowed examination of language switches in recall.

At the end of the study, participants were asked to rate their overall fluency in Spanish and English on two 7-point scales and also were asked their country of birth and first language acquired. Because of the nature of the study, participants would have obviously been put into a bilingual mode, in the sense of Grosjean (1997), in which both of their languages could be assumed to be in a reasonably high level of activation, unlike the situation of the participants in Experiments 1 and 2. They knew that both English and Spanish and their own bilingual skills were relevant to the experiment.

Results and Discussion

There were two primary questions of interest. First of all, what was the frequency and pattern of switching from one language to another (e.g., sentence heard in Spanish and recalled in English)? Secondly, what was the pattern of switching from one figurative form to another (e.g., heard as metaphor and recalled as simile) and how was this similar or different from the comparable data in Experiments 1 and 2? Although this latter question is the primary focus, we will examine the language switching data first.

Language switches in recall. Table 3 presents the per cent of responses in Experiments 3 and 4 recalled as metaphors or similes but switched to the other language. Although there were more switches in concrete sentences (12-16% overall) than abstract (4-8% overall), this was due to the very large number of abstract sentences which were totally forgotten (see Table 4 and discussion below). There were no significant differences (by chi-square tests) as a function of either language or figurative type (metaphor vs. simile). Thus we can probably rule out some overpowering production bias that drives recall for material, regardless of the input language, to be recalled systematically in either English or Spanish.

Participants were in what Grosjean (1997) called the bilingual mode, meaning that they were well aware that their bilingual skills were relevant to the experiment and both languages were in some degree of activation during the study. Much as two bilinguals talking with each other frequently code-switch to facilitate communication, it may be that participants in this experiment were doing a sort of code-switching in responding to the memory task, i.e., responding in whichever language came to mind first. However, that there was virtually no case of a mixed-language response to a single sentence (e.g., Abusive control is like a cracking latigo)*.*

Patterns of recall by figurative type. Table 4 presents the per cent recall by response category for Experiment 3. Before examining the trends of recall by figurative type, it is important to note the large number of omitted responses, especially for the abstract sentences. This bilingual task was apparently quite difficult, and a very large per cent of the responses were left blank. This greater difficulty of the abstract sentences is consistent with various conceptual models of bilingual memory (see Kroll & de Groot, 1997).

After seeing the large number of omitted responses on Experiment 3, the study was replicated in briefer form with 50 new participants from the same pool. Except for halving the number of sentences from 32 to 16, i.e., using only the first half of each tape, Experiment 4 was identical to like Experiment 3 in procedure. Surprisingly, however, cutting in half the number of sentences to remember made the task only very slightly easier. Still, 23-35% of the concrete sentences and 53-72% of the abstract sentences were totally forgotten. Otherwise the pattern of results was virtually identical to that of Experiment 3. Therefore, only the results from Experiment 3 are presented from this point on, given that the sample and the number of data points per participant were each twice as large as those numbers in Experiment 4.

An examination of Table 4 shows that overall the huge language asymmetry observed in Experiments 1 and 2 was greatly attenuated or totally eliminated in the bilingual study. For the English input sentences, concrete metaphors and similes were most often recalled as their own type, although there are also large numbers of switches to the other type. Unlike in Experiment 1, however, there was no large bias of recalling English similes as metaphors.

With the Spanish input sentences, however, there is more evidence of the retention of the bias to recall both metaphors and similes as similes, as occurred in the monolingual Experiment 2. Both concrete metaphors and similes were three times as likely to be recalled as similes as they were as metaphors.

The data from Experiment 3 were analyzed by four three-way (language, concreteness, figurative type) analyses of variance, with repeated measures on all three factors. Separate analyses were done for the recall categories of English metaphors, English similes, Spanish metaphors, and Spanish similes, with confidence intervals computed around each cell mean. Effect sizes for each significant (p<.01) effect are presented in Table 5. All four analyses showed main effects of concreteness, with concrete sentences more often recalled than abstract ones. For the analyses of sentences responded to as English metaphors, several effects were due primarily to the fact that there were many more English concrete metaphor input sentences recalled in this category than any other (no overlap of 95% confidence interval with any other response category). For the English simile recalls, English concrete metaphors and English concrete similes were recalled as English similes more than was any other category of input sentences. The Spanish metaphor recall category showed only a main effect of concreteness, largely due to a very low number of responses in this category overall, given that very few English sentences were recalled in Spanish and very few Spanish sentences were recalled as metaphors. Finally, the Spanish simile recalls showed significantly more recalls of Spanish concrete metaphors and Spanish concrete simile input sentences than from any other category.

The role of fluency. One possible variable to consider in any bilingual task is the relative fluency of participants in the two languages. Although it is not clear what would be the most appropriate fluency measure in this case, given that the task involved both listeningandwriting, participants did provide a self-rating of general overall fluency in both Spanish and English. Because the overall rated fluency was higher in English (in spite of Spanish being the native language of a clear majority), neither a median split nor matching groups on fluency was feasible. As a reasonable estimate of results from these more precise but unattainable procedures, the data presented in Table 4 were recalculated only for the subset of participants who rated their English as better than their Spanish, who comprised two-thirds of the original sample (N=66). If there were large effects of relative fluency of the two languages, removing those who rated their Spanish as better or equal to their English should homogenize the data and remove some of the variance. Table 6 presents exactly the same data as Table 4 but only for the subset of the sample who rated their English better than their Spanish. The data are virtually indistinguishable from those in Table 4, so no further analyses are reported.

Why did the repartitioning of the data by relative fluency not produce different results? It may be that the present sample was sufficiently balanced that the existing modest differences in fluency did not matter in this task. It is also possible that a global self-rating of general fluency cannot be made reliably enough to affect this particular task, which involves both oral comprehension and writing. Many participants commented to the experimenter that they felt they were far more proficient in listening and conversing in Spanish than in reading or writing it. Perhaps separate fluency rating scales for oral and written Spanish would have revealed greater discrepancy in the fluency ratings. Also, perhaps a chance to respond orally to the memory task would have made the task easier.

Another possibility is that the differential fluency ratings may be an artifact of the experimental setting. English fluency ratings may have been inflated by virtue of being done in the academic setting, where participants are used to dealing more in English than in Spanish. Perhaps if they had been collected in the home environment, when most of the participants spoke much more Spanish, skill in the two languages would have been rated more similarly. It may also be that individuals who are in a bilingual mode are more attentive to the literal form of the input, and thus the surface form of whether it is a metaphor or simile has a relatively higher salience than in Experiment 1 and 2. The data from Experiments 3 and 4 support this view but only for the English sentences, which may further reflect that English is the language of school and Spanish is the language of the home. Finally, it may be that the rather modest differences in self-rated fluency in the present participants are simply not important in the present task.

It is also not clear if the difficulties with this task are due more to comprehension problems in hearing the sentences or production problems in writing the completions of the sentence stems on the answer sheet. The fact that results from Experiment 4 replicated Experiment 3 suggests that the difficulty is not due solely, or even primarily, to the task being too long.

General Discussion

The large memory differences of metaphors and similes in Spanish and English are awkward for most models of figurative language, none of which make language-specific predictions. Although Johnsons (1996) reaction-time data showing that metaphors were comprehended more quickly than similes in both languages are entirely consistent with Glucksbergs attributive categorization model, neither this model nor any other predict language differences in recall. The fact that the present research so strongly supported Johnson (1995) confirms that this large difference was not a statistical or methodological fluke. The fact that the Spanish simile bias, but not the English metaphor bias, was maintained in the bilingual studies suggests that it may be the stronger of the two effects.

One likely explanation of the language difference is some sort of production output constraint, perhaps metaphors being more common or stylistically preferred in English, whereas similes are in Spanish. Even if this is true, it is not clear how such a bias could explain the magnitude of the obtained difference. In this task participants heard many sentences of two clearly different types; one might expect the task demands, if anything, to bias the results toward a more balanced number of metaphor and simile recalls, because the participant would be aware that there were two types and probably could guess that their numbers might be fairly even.

In terms of Grosjeans (1997) call for a greater attention to whether a participant is in a monolingual or bilingual mode during an experiment, the present research determined the mode through the task itself. In Experiments 1 and 2 there was no mention of any other language but the language of the experiment (also the participants best and usually only language). In Experiments 3 and 4, however, not only were participants recruited with a qualification of being bilingual, but also all experimental instructions and stimulus materials were in both languages. This procedure necessarily activated both languages. If such activation could have attenuated the effects obtained in Experiments 1 and 2, it does not explain why the Spanish simile bias was retained more than the English metaphor bias. Also, the fact that Experiment 2 replicated Johnsons (1995) data with English-Spanish bilinguals performing a monolingual Spanish task suggests that the monolingual/bilingual mode was not the critical factor in this case.

Future research would do well to try the present task with a greater variety of memory measures, such as recognition memory, free recall, or even some sort of implicit memory measure. Presenting sentences in a more meaningful context of discourse would also be useful. It may be, for example, that abstract sentences would be easier to remember in a meaningful context. Finally, studies examining additional different types of figurative and literal language or alternative types of metaphors or similes would be worthwhile. Although present participants tended to recall both Cigarettes are time bombs and Cigarettes are like time bombs as the former, how might they recall variations like Cigarettes kill or Cigarettes kill like time bombs?

Results such as these remind us again that there is a great need to replicate and extend psycholinguistic studies to languages other than English. Sometimes the results replicate completely, as in the reaction-time studies of Johnson (1995), but sometimes they do not, as in the present studies. To study only one language is to risk mistaking a language-specific effect for a linguistic or cognitive universal. Replication of the present studies with other languages may shed further light on the obtained results.

Although there is certainly a need to understand the cognitive processes involved in the comprehension and memory for figurative language for its own sake, there may also be application beyond the domain of metaphor and similes. For example, Wisniewski (1997) argues that the comprehension of metaphors is but one type of a more general type of language comprehension of conceptual combinations like seafood sausage, butterfly clip, or mushroom cloud*.* In any event, figurative language is so prevalent that it must be accounted for by any comprehensive language processing models.

Reference Note

Author Note: Michael Tebbe is now at the University of Kansas School of Law. Correspondence about the paper may be addressed to R.J. Harris, Department of Psychology, Kansas State University, 492 Bluemont Hall, 1100 Mid-Campus Drive, Manhattan KS 66506-5302 (e-mail: rjharris@ksu.edu).

References

Abelson, R.P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test. Psychological Science, **8**, 12-15.

Black, M. (1962). Models and metaphors. Ithaca NY : Cornell University Press.

Estes, W.K. (1997). Significance testing in psychological research: Some persisting issues. Psychological Science, **8**, 18-20.

Gibbs, R.W., Jr. (1994). The poetics of mind. Cambridge UK: Cambridge U. Press.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusions. Psychological Science, **3**, 146-152.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphoric comparisons: Beyond similarity. PsychologicalReview*,* **97,** 3-18.

Glucksberg, S., McGlone, M.S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. Journal of Memory and Language *,* **36***,* 50-67.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A.M.B. de Groot and J.F. Kroll (Eds.), Tutorials in bilingualism: Psycholinguistic perspectives. (Pp. 225-254). Mahwah NJ: Lawrence Erlbaum Associates.

Hunter, J.E. (1997). Needed: A ban on the significance test. Psychological Science, **8**, 3-7.

Johnson, A.T. (1995). Comprehension of metaphors and similes: Reaction time and memory studies. Unpublished doctoral dissertation. Kansas State University.

Johnson, A.T. (1996). Comprehension of metaphors and similes: A reaction time study. Metaphor and Symbolic Activity, **11**, 145-159.

Kroll, J.F., & de Groot, A.M.B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A.M.B. de Groot and J.F. Kroll (Eds.), Tutorials in bilingualism: Psycholinguistic perspectives. (Pp. 169-199). Mahwah NJ: Lawrence Erlbaum.

McGlone, M.S. (1996). Conceptual metaphors and figurative language interpretation: Food for thought? Journal of Memory and Language *,* **35**, 544-565.

Miller, G.A. (1979). Images and models, similes and metaphors. In A. Ortony (Ed.), Metaphor and thought*.* (pp. 202-250). Cambridge UK: Cambridge University Press.

Richards, I.A. (1936) The philosophy of rhetoric. London: Oxford University Press.

Wisniewski, E.J. (1997). When concepts combine. Psychonomic Bulletin and Review, **4,** 167-183.

Table 1

Per Cent Recall Responses by Category in All-English Study (Experiment 1) and All-Spanish Study (Experiment 2)

Recalled as

Input Sentence Metaphor Simile Error Omitted

English Metaphors

Concrete 69 9 1 20

Abstract 26 5 2 68

English Similes

Concrete 58 18 1 24

Abstract 26 8 2 64

Spanish Metaphors

Concrete 21 40 2 36

Abstract 10 26 1 61

Spanish Similes

Concrete 12 44 2 41

Abstract 6 25 1 69

Notes: Sentences recalled as metaphors and similes include both verbatim and gist recall. Rows sum to 100%, minus effects of rounding.

Table 2

Effect Sizes (Partial Eta Squared) for Significant (p<.01) Effects from ANOVAs in Experiments 1 and 2

Recall Category

Effect Metaphor Simile

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Language .46 .44

Fig. Type .17 .06

Concreteness .59 .35

Lang X Concr. .35 .09

Fig. Type X Concr. .08 .05

Note: Neither the Lang X Fig. Type nor the three-way interaction was significant (p<.01) in either analysis.

Table 3

Per Cent Recall Responses in Other Language (pooled across Experiments 3 and 4)

Concreteness Type

Input Form Concrete Abstract

English Metaphors 12 4

English Similes 13 4

Spanish Metaphors 16 7

Spanish Similes 14 8

.

Table 4

Bilingual Mixed-Language Studies: % Recall Responses by Category (Experiment 3, N=101)

Recalled in English Recalled in Spanish Omit Error

Input Type Met Sim Met Sim

English

Met-Conc. 31 23 6 6 30 4

Met-Abst. 12 14 2 2 68 2

Sim-Conc. 16 32 4 11 34 3

Sim-Abst. 8 14 2 2 68 6

Spanish

Met-Conc. 7 6 6 18 47 16

Met-Abst. 4 3 5 4 80 4

Sim-Conc. 5 10 6 21 48 10

Sim-Abst. 3 3 4 5 80 5

Notes: Metaphor and Similes responses include both verbatim and gist recalls. Rows sum to 100%, minus rounding errors.

Table 5

Effect Sizes (Partial Eta Squared) for Significant (p<.01) Effects from ANOVAs (Experiment 3)

Recall Category

Effect English Metaphor English Simile Spanish Metaphor Spanish Simile

Language .41 .42 ns .26

Fig. Type .24 .07 ns .06

Concreteness .30 .31 .07 .46

Lang X Fig. Type .16 ns ns ns

Lang X Concr. .22 .13 ns .14

Fig. Type X Concr. .09 .07 ns ns

Note: The three-way interaction was not significant (p<.01) in any analysis.

Table 6

Bilingual Mixed-Language Studies: % Recall Responses by Category (Experiment 3) where English Fluency > Spanish Fluency (N=66)

Recalled in English Recalled in Spanish Omit Error

Input Type Met Sim Met Sim

English

Met-Conc. 37 23 4 4 28 4

Met-Abst. 13 13 3 1 68 2

Sim-Conc. 20 34 4 7 30 5

Sim-Abst. 8 16 2 1 65 8

Spanish

Met-Conc. 8 6 7 15 52 12

Met-Abst. 2 3 4 4 80 7

Sim-Conc. 6 8 6 15 56 9

Sim-Abst. 4 2 3 4 80 7

Notes: Metaphor and similes responses include both verbatim and gist recall. Rows sum to 100%, minus effects of rounding.

Data analysis. The per cent responses in each response category by input sentence type are presented in Table 1.

Two-way repeated-measures analyses of variance were performed on the correct form verbatim, correct-form-gist, reversed form-verbatim, reversed form-gist, and omitted responses, with p=.01. Results for the correct verbatim responses showed main effects of figurative type, F(1,109)=162.76, MSe=2.30, and concreteness, F(1,109)=443.03, MSe=2.30, and their interaction, F(1,109)=89.92, MSe=2.30. The correct gist analysis showed only a main effect of figurative type, F(1,109)=42.66, MSe=.43. These results reflected many more of both types of correct responses to metaphors than to similes. Also, there were many more correct verbatim responses to concrete metaphors than to any of the other three types.

Analyses for the reversed-form responses (i.e., metaphors recalled as similes and similes recalled as metaphors) showed main effects of figurative type, F(1,109)=138.74, MSe=2.12, and concreteness, F(1,109)=246.31, MSe=2.12, and their interaction, F(1,109)=85.37, MSe=2.12, for reversed-verbatim responses. The reversed-gist category also produced significant effects of figurative type, F(1,109)=50.93, MSe=.46, and concreteness, F(1,109)=18.16, MSe=.46, and their interaction, F(1,109)=6.76, MSe=.46. These results reflect many more both verbatim and gist reversed-type responses for similes than metaphors. There were more reversed-type verbatim responses for concrete sentences and more reversed-type gist responses for abstract sentences. The interaction terms are largely due to a very large number of reversed verbatim responses to concrete similes and reversed gist responses to abstract similes.

The analysis of the omitted responses showed only a main effect of concreteness, F(1,109) =617.7, MSe=1.29, with many more omitted responses to abstract than concrete sentences. In fact, about two-thirds of all possible responses to abstract sentences were omitted.

Table 1

Per Cent Recall Responses by Category in All-English Study (Experiment 1)

Recalled as

Metaphor Simile Error Omitted

Input Sentence Verbatim Gist Verbatim Gist

Metaphor

Concrete 59 10 8 1 2 20

Abstract 15 11 2 3 2 67

Simile

Concrete 50 7 15 3 1 24

Abstract 14 12 4 4 2 64

Note: Sentence recalled as metaphors and similes include both verbatim and gist recall.