L2 Learners’ Stylistic Pennchant and Perceived Aptness for Metaphors and Similes

Mahmood Hashemian
m72h@hotmail.com
Shahrekord University

Abstract
Aptness, defined as how the vehicle is well able to cover the salient features of the tenor (e.g., oil is like liquid gold vs. a train is like a worm), is claimed to be an important factor in the preference for metaphors over similes, or vice versa. This study was an attempt to test for the supposed correlation between the perceived degree of aptness and a priori stylistic preference for metaphors and similes by Iranian L2 learners. Participants, aged 20-25, were selected from 80 EFL Translation undergraduates. In the first place, they were asked to read 2 alternative lists of the same sentences (lists A and B) and to rate each sentence as to its appropriateness by filling in a number between 2 endpoints of 1 (very inappropriate) and 7 (very appropriate). In the second place, they were invited to consider the 2 alternative forms of the same sentence and to say which one they preferred: the metaphor or the simile form. Results revealed no such strong or moderate relationship between the perceived mean aptness ratings and the mean preference scores for the simile and metaphor versions ($r = -0.014$ for the metaphors; $r = 0.014$ for the similes). All things considered, the perceived degree of aptness failed to predict the stylistic preference for metaphors and similes. This comes to the rejection of the validity of the claim put forth by Chiappe and Kennedy (1999) regarding the predictive power of aptness to inform the preference of metaphors and similes by Iranian L2 learners.

Keywords: Aptness; Metaphor; Simile; Stylistic preference; Tenor; Vehicle

1. Introduction
Figurative language accounts for a large proportion of linguistic expressions one routinely uses in everyday verbal communication (Van Lancker, 2003). This, in main, manifests itself in such heavily figurative-laden tropes as similes and, in particular, metaphors, which are astonishingly believed “to work because the mind is metaphoric” (Lankton, 2002, p. 13). The abundant uses of metaphors, found in both oral and written discourse, would bear testimony to this forthright claim.
First, some technical terms need to be defined. A simile is an expression in which something is compared to something else by the use of such function words as like or as. In Tom eats like a horse, Tom’s appetite is compared to that of a horse. However, in a metaphor, no such function words are used. Something is described by directly stating another thing with which it can be compared. In Her words stabbed at his heart, the words did not actually stab, but their effect is compared to the stabbing of a knife. So, as is evident, metaphors and similes are structurally and lexically identical except for the presence of the explicit comparison marker like in the simile.

Both metaphors and similes feature two important components, namely tenor (topic) and vehicle. According to Kitty (as cited in Cameron & Low, 1999), tenor (topic) is defined as the entity, idea, or action which the figure of speech is talking about (e.g., That lie is a boomerang!). Also, vehicle is defined as the entity, idea, or action in terms of which tenor is compared (e.g., That lie is a boomerang!). Vehicle counts the object of comparison. What the tenor and the vehicle have in common is the ground (in this case, the backfiring of the action on its originator). It is this part of metaphor which must be properly understood in order for successful comprehension of metaphor to click into place for the listener and the reader (Harris et al., 2006). Tenor and vehicle can be both concrete and abstract. Frequently, tenor is abstract to make the comparison more vivid.

Given that metaphor and simile are of the essence, it is little wonder that there has been a vast body of research on figurative language (e.g., Cameron & Low, 1999; Roberts & Kreuz, 1994; Steen 2004). After all, those scholars intrigued by this amazing figurative dimension of language use are sure to come up with many questions in their minds. They may desire to know about the ways individuals can perceive and produce cognitively a metaphorical piece of language. Specifically, they may long for the adoption of a systematic approach to examine the steps language users take to perceive and produce a text, oral or written, with instances of metaphor.

As Harris, Friel, and Mickelson (2006) believe, studies on figurative language have taken two sides: comprehension and production. As for the former, the area of L1 and L2 research abounds with miscellaneous studies on the comprehension of figurative use of language, but regarding the latter, upon examining the relevant literature, one would spot areas of neglect in this research domain (Harris et al., 2006). These two research dimensions of metaphor are not on a par in terms of the amount of research done, and quantity of figurative research is weighted in favor of the comprehension side.

Of relevance to the aim of the present study is the question of the a priori stylistic preference for one figurative form over another, being metaphor and simile
in this case, and also of the perceived degree of agreement between the salient features of the vehicle and those of the tenor in question—aptness—on the part of the receiver of language.

Making comparative predications about similes and metaphors, some psychological models of figurative language (e.g., Miller, 1979) posit that similes, as opposed to metaphors, are comparative structures drawing more basic, direct, and explicit comparisons between tenors and vehicles. Their argument runs that it is the existence of the preposition like in a simile sentence, not a metaphorical one, which makes the comparison so mentally vivid and clear. On the other hand, these psychological models regard metaphors as short, or abbreviated, forms of simile which state an implicit comparison (Harris et al., 2006). Reading between the lines, these models seem to draw heavily on syntactic features and do not take the discourse context into any consideration.

Regarding the relationship between metaphors and similes, another model, eagerly embraced by Mac Cormac (1985), sets out to distinguish between these two figurative devices. Here, it is argued that the choice of metaphor is most probably motivated by a desire on the part of the speaker or the writer to highlight both the similarities and the differences simultaneously existing between the tenor and the vehicle. However, as maintained by this model, when the goal is to highlight just the similarities, similes will come to the fore. In case of novel metaphors, metaphors which are not highly conventionalized (Kövecses, 2002), the occurrence of uncommon tenors and vehicles together causes surprise, which in turn demands much more mental processing than the sort we would come up with in a simile (Ricoeur, 1977).

Yet, another intriguing model is attributive categorization or class-inclusion (Blasko & Connine, 1993; Gernsbacher, Keysar, Robertson, & Werner, 2001; Glucksberg, 1991, 1998, 2001, 2003; Glucksberg, Newsome, & Goldvarg, 2001; McGlone & Manfredi, 2001). According to this model of metaphor, the vehicle has “dual reference” (Jones & Estes, 2005). The argument is that the vehicle has the capability to carry both literal and figurative meanings. Take this illuminating example. When interpreted in a literal and metaphorical sense, respectively, gold can refer both to an example and to a class of objects (e.g., That bracelet is gold and Happiness is gold). In the second example, gold metaphorically refers to the class of valuable objects. Corollary to this argument, during the process of metaphorical comprehension, the metaphorical reference of the vehicle is bound to be cognitively triggered, and the literal reference is reserved for other times of literal interpretation (McGlone & Manfredi, 2001). Viewed from this angle, metaphor is a more direct and explicit statement of comparison, and therefore, is easier to comprehend than its simile counterpart.
A common quality that both metaphors and similes share is the degree of concreteness of the comparisons they make. It might be the case that, under certain linguistic circumstances, L2 learners of English stylistically prefer metaphors or similes. On the evidence of one research study (Gibb & Wales, 1990), it was suggested that when the vehicle was abstract, metaphors were preferred, and when it was concrete, simile forms were the prime candidates. In this connection, it was shown that the degree of concreteness of the tenor played no significant role. Also, Zelman et al. (as cited in Harris et al., 2006) found a positive correlation between the degree of concreteness and the degree of comprehensibility.

Also, being interested in the stylistic choice of wording in sentences, Harris et al. (2006) examined the preference for the choice of metaphors and similes on the part of native English-speaking students. Overall, similes were perceived over metaphors, with mean preferences of 72% and 28%, respectively.

Chiappe & Kennedy, 1999; Gibbs, 1993, 1994; Katz, 1989, 1992; Malgady & Johnson, 1976; Tourangeau & Sternberg, 1981, 1982). For example, a train is like a worm may not be a very apt comparison because the comparison fails to encompass many of the main features of a train like its typical strength and power. However, a figurative sentence such as Oil is like liquid gold may seem a very apt comparison, for the comparison does seem to capture many of the salient features of oil, such as its value and rarity. It is generally believed that how immediately people are able to figure out the perceived overlapping properties of the vehicle and the topic is a crucial factor in the comprehensibility of a comparison (D. Chiappe, Kennedy, & P. Chiappe, 2003).

In an empirical study, Jones and Estes (2006) investigated the role of conventionality and aptness in metaphor comprehension. The results of their study showed that aptness, not conventionality, affected the preference for the choice of metaphors over similes. Metaphor preferences correlated closely with high apt comparisons than for low apt comparisons. This replicates exactly the results of the study by Chiappe and Kennedy (1999), which gave confirmation to class-inclusion theory. The researcher found that aptness was conducive to the prediction of the form of the comparisons in question. The participants were asked to judge whether particular comparisons were best expressed as metaphors or similes. This created a range of preferences, with highly apt comparisons being more likely to be chosen as metaphors than as similes. However, in another study conducted by Harris et al. (2006), no significant relation was found between the perceived mean aptness ratings for the simile and metaphor versions and the Figure-Type mean preferences for metaphors and similes.
Given the above conflicting findings regarding the relative role of aptness in influencing the preference for the choice of metaphors and similes, the main purpose of the present study was to test for the supposed correlation between the perceived degree of aptness and a priori stylistic preference for metaphors and similes by Iranian L2 learners. The study, therefore, sought to test the following null hypotheses:

- H01: There is no significant difference between the mean preference scores by Iranian EFL undergraduates for metaphors and the mean preference scores for similes.
- H02: There is no significant difference between the mean aptness ratings by Iranian EFL undergraduates for metaphors and the mean aptness ratings for similes.
- H03: There is no significant correlation between the mean aptness ratings for simile and metaphor versions and the mean preference scores by Iranian EFL undergraduates for similes and metaphors.

2. Methodology

2.1 Participants

The participants were selected from a total population of 80 male and female EFL Translation undergraduates, aged 20-25, from Shahrekord University in their fourth year of study. In order to ensure the homogeneous entry behavior of the participants in terms of proficiency, the Oxford Placement Test (OPT), with reasonable measures of validity and reliability, was used to screen the students. The participants who scored lower than 50% of the total possible score (100, $M = 50$, $SD = 20$) were excluded from the study. This filtering process left the present researcher with 40 participants. The reason for the selection of these fourth-year participants was that they were already familiar with the concept of metaphor and simile in English, besides their native language knowledge, because they had passed courses on English Literature and Translation of Idiomatic Expressions. In this way, they might be considered advanced learners.

2.2 Materials and Procedure

This research study consisted of two experimental phases: Experiment 1 and Experiment 2, each with its own specific procedure.

In Experiment 1, the focus was on the participants’ a priori stylistic preference for one form over another. The participants were invited to mull over the two alternative forms of the same sentence and say which one they preferred: the
simile form or the metaphorical one (e.g., Research is like mountain climbing instead of Research is mountain climbing). They were given a sheet with two columns of 16 sentence-pairs. There was either the metaphor or simile version of a given sentence in each column, with the alternative version in the other column. The rational for the selection of the tropes was topic concreteness. Half of them contained concrete tenors/subjects (e.g., The giraffe is (like) a skyscraper), whereas half contained abstract tenors/subjects (e.g., Creativity is (like) a toaster). The task for the participants was to put a check mark (✓) for the alternative form they preferred in the small box beside the sentence.

As for Experiment 2, the main focus of attention was on how well the topic and the vehicle could jibe with aptness. In other words, a special attempt was made to investigate how well the vehicle would be able to capture the salient features of the topic in question. There were two alternative lists (A and B). Each list consisted of the same eight metaphors and eight similes as in Experiment 1. The participants received one of the parallel lists so that each participant rated only the figurative sentences of one of the two forms as to their appropriateness (congruity of the topic and the vehicle, aptness). Next to each sentence was a 7-point scale: 1 (very inappropriate), 2 (inappropriate), 3 (almost inappropriate), 4 (moderately inappropriate), 5 (almost appropriate), 6 (appropriate), and 7 (very appropriate). Participants were asked to circle one of the scales between the two end anchors.

3. Data Analysis

3.1 The Results of Experiment 1

As mentioned in section 2.3, the purpose of Experiment 1 was to assess the participants’ stylistic preference of one form over another, being metaphor and simile in this case. The data from the participants, who compared two alternative forms of the same sentence (i.e., metaphor or simile), were statistically analyzed to come up with the mean preference scores for both similes and metaphors:

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Sum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simile</td>
<td>40</td>
<td>351</td>
<td>8.78</td>
<td>3.051</td>
</tr>
<tr>
<td>Metaphor</td>
<td>40</td>
<td>289</td>
<td>7.23</td>
<td>3.051</td>
</tr>
</tbody>
</table>

As shown in Table 1, in general, similes were preferred over metaphors, with mean preference scores of 8.78 and 7.23, respectively.
Also, in order to ascertain if this statistical difference between the mean preference scores for metaphors and the mean preference scores for similes was statistically meaningful, a *paired samples t test* was run between the two mean scores:

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (2 tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
</tr>
</tbody>
</table>

As seen in Table 2, there is no statistically meaningful difference between the mean preference scores for metaphors and the mean preference scores for similes ($t = 1.607$, $df = 39$, $\alpha = 0.05$, $p = 0.116$). Because $p$ value is more than $\alpha$, the first null hypothesis is not rejected.

### 3.2 The Results of Experiment 2

As mentioned in section 2.3, the purpose of Experiment 2 was to gauge the degree of aptness as conceived by the L2 learners, becoming operationalized by requiring the participants to rate simile and metaphor sentences on a scale ranging from 1 (*very inappropriate*) to 7 (*very appropriate*). The data were subjected to descriptive statistics to acquire the mean aptness ratings for both similes and metaphors:

#### Table 3 Mean Aptness Ratings for Metaphors

<table>
<thead>
<tr>
<th>Score</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid N (listwise)</td>
<td>20</td>
<td>4.6969</td>
<td>.71355</td>
</tr>
</tbody>
</table>

#### Table 4 Mean Aptness Ratings for Similes

<table>
<thead>
<tr>
<th>Score</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid N (listwise)</td>
<td>21</td>
<td>4.8512</td>
<td>.45478</td>
</tr>
</tbody>
</table>
As shown in the Tables 3 and 4, the mean aptness rating for metaphors is equal to 4.696, whereas the mean aptness rating for simile is 4.851. At face value, there was marginal difference between these two aptness rating means.

To see if the obtained difference between the mean aptness ratings for simile sentences and the mean aptness ratings for metaphor sentences was statistically significant, a paired samples t test was run between the two mean scores:

Table 5 Paired Samples t Test for Mean Aptness Ratings for Metaphors and Similes

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig (2 tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure type - Score</td>
<td>1.888</td>
<td>6.325</td>
<td>.77</td>
<td>-3.01</td>
<td>4.501</td>
<td>40</td>
<td>.129</td>
</tr>
</tbody>
</table>

As depicted in Table 5, Figure Type variable had no effect on the degree of aptness ($t = 1.123$, $df = 40$, $\alpha = 0.05$, $p = 0.129412$). Because $p$ value is more than $\alpha$, and so no significant difference existed between the mean aptness ratings for metaphors and the mean aptness ratings for similes. So, the second null hypothesis was also maintained.

Moreover, to check any correlation between aptness and Figure-Type preference, the mean aptness ratings for the metaphor and simile versions obtained from Experiment 2 were compared with their preference ratings for metaphors and similes obtained from Experiment 1, and they were statistically analyzed, using the Pearson Product Moment Correlation:
Table 6 Correlation Between Aptness and Figure-Type Preference

<table>
<thead>
<tr>
<th></th>
<th>Pearson Correlation</th>
<th>Number of Similes</th>
<th>Number of Metaphors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>1</td>
<td>.014</td>
<td>-.014</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.933</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Number of Similes</td>
<td>.014</td>
<td>1</td>
<td>-1.000(**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.933</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>N</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Number of Metaphors</td>
<td>-.014</td>
<td>-1.000(**</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.933</td>
<td>.000</td>
<td>40</td>
</tr>
<tr>
<td>N</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Table 6 illustrates a weak relationship between aptness and Figure-Type preference \( (r = -0.014 \) for the metaphors; \( r = 0.014 \) for the similes). Therefore, the third null hypothesis was maintained, too.

4. Discussion and Conclusion

Overall, the two versions of the sentences, metaphors and similes, had no effect on the choice of one form over the other. This finding is not consistent with the one by Gibb and Wales (1990), who came to the conclusion that sentences featuring abstract vehicles were more often chosen by the participants in the metaphor form than in the simile form, and the one by Harris et al. (2006), who found a significant difference between the preference scores for metaphors (28%) and the preference scores for similes (72%). Apparently, the Iranian L2 learners considered both metaphors and similes to fulfill one single function: drawing the attention to the similarities existing between the tenor and the vehicle. Similes and metaphors, in this study, were conceived by the participants as conveying the idea of similarity rather than difference. This preference for both similes and metaphors as the definite indicators of similarity was clearly accounted for by the heavy bent towards the choice of the discourse goal of Compare Similarities for both metaphors and similes in the study by Hashemian and Iravani (2010). In this study, the discourse goal of Contrast Differences had no major role to play in informing the Iranian L2 learners’ choice of the possible discourse goals for metaphors and similes. This finding stands in contrast to that of MacCormac (1985), who maintained that the major function performed by metaphors is that of projection of both similarities and differences. His markedly different line of argument runs the explanation that the choice of metaphor, as against similes, exclusively characterized
by similarity, is most probably motivated by a desire on the part of the speaker or
the writer to highlight both similarities and differences simultaneously existing
between the tenor and the vehicle. Patently, the present study failed to come up with
this controversial finding.

As an alternative to the explanation for the obtained results, one might
attribute lack of significance between the mean preference scores for metaphors and
the mean preference scores for similes to the existence of no abstract vehicles in the
materials presented to the participants (Harris et al., 2006). All the simile and
metaphor sentences used in this study contained exclusively concrete vehicle. It
might be the case that the degree of concreteness of the vehicle could affect
stylistically the choice of metaphors or similes. In the study carried out by Gibb and
Wales (1990), it was metaphors that the participants mostly opted for regarding the
Abstract-tenor sentences and similes that most often chosen for Concrete-tenor
sentences. Therefore, the present study, due to inherent differences in stimuli, was
unable to find such an effect.

Regarding the second null hypothesis, Figure Type variable had no
significant effect on the perceived degree of aptness by the participants for
metaphors and similes. All in all, one could argue that both metaphors and similes
were rated as appropriate in terms of their degree of aptness. Extending this finding
to the findings obtained in the experiment by Hashemian and Iravani (2010), one
could see no relation between the dominant choice of the discourse goal of Compare
Similarities and the perceived degree of aptness for metaphors and similes. Were it
to suppose that the dominant choice of the discourse goal of Compare Similarities
for simile form would predict the degree of aptness for similes, the similes would
have been rated as more apt in this experiment. However, no such effect was found
in the study. Therefore, one might conclude that the reason for the participants’
strong tendency in the abovementioned study to choose similes as conveying the
idea of Compare Similarities is not due to the fact that similes, in general, are more
apt for making the comparisons. This finding (Hashemian & Iravani, 2010) offers
further support for that by Harris et al. (2006) who also found no significant
difference between the mean aptness ratings for metaphors (4.88) and the mean
aptness ratings for similes (5.08), both being rated as appropriate in terms of the
extent to which their vehicles were able to state and encompass many of the main
characteristics of the tenor.

Also, regarding the third null hypothesis, the possible relation between the
perceived mean aptness ratings for the simile and metaphor versions and the Figure-
Type mean preferences for metaphors and similes obtained from Experiment 1, no
direct relationship was found ($r = -0.014$ for the metaphors; $r = 0.014$ for the
similes). This finding is in line with that by Harris et al. (2006), who also found a
weak correlation between aptness and Figure-Type preference for metaphors ($r =$
0.11) and similes ($r = -0.11$). However, the results run contrary to the findings of the study by Chiappe and Kennedy (1999) who concluded that those comparisons rated as more apt turned out to be those comparisons that were mostly metaphors rather than similes. However, the present study does not support this finding. As such, the degree of aptness was not found to be a good barometer for the preference for metaphor or simile versions.

In sum, no significant difference was found between the mean preference for metaphors and the mean preference for similes. The difference between the mean aptness ratings for metaphors and the mean aptness ratings for similes did not turn out to be statistically meaningful. Also, no significant correlation was found between the perceived degree of aptness and a priori stylistic preference for metaphors and similes by Iranian L2 learners.

Like any other research study in this contemporary world, this empirical study was not able, mostly due to practical considerations, to fully capture all the intriguing and important aspects of the phenomenon under investigation. Certainly, there are methodologically fruitful and important insights for future metaphor researchers, hidden in this piece of work. The hope is that prospective L2 researchers keep these suggestions on the front burner and apply them to their forthcoming research studies.

Given the lack of any significant difference between the mean preference scores for metaphors and the mean preference scores for similes in Experiment 1, future researchers are strongly advised to also include abstract vehicles in their materials. All the simile and metaphor sentences used in this study contained exclusively concrete vehicle. It might be the case that the degree of concreteness of the vehicle could affect stylistically the choice of metaphors or similes.

Also, considering the possibility that no significant relation was found between Aptness and Figure-Type preference as a result of the materials being not sufficient, future research might consider the inclusion of more items as a way to tackle this possibility.

References


Ricoeur, P. (1977). The rule of metaphor: Multi-disciplinary studies of the creation of meaning in language (R. Czerny et al., Trans.). Toronto: University of Toronto Press. (Original work published 1975)


Mahmood Hashemian is an assistant professor at Shahrekord University. His area of research includes cognitive-semantic linguistics, sociolinguistics, and applied linguistics. He has published articles in academic journals such as IJAL, IJLS, JALS, Linguistik online, JLTR, TPLS, Iranian EFL Journal, and International Journal of Social Sciences. Also, he has given lectures in conferences such as TELSSI (4, 7, 8, &9), LDP2010, ELT in the Islamic World, and 2nd International Language Conference (ILC) 2011, Malaysia.