SECOND LANGUAGE ACQUISITION OF METAPHORS 
LATENT SEMANTIC ANALYSIS PREDICTIONS AND 
EXPERIMENTAL EVIDENCE

Tang Shimin  Yue Jianhui 
Graduate University of Chinese Academy of Sciences

Abstract

A major problem for SLA research is to account for a learner's acquisition of knowledge not available in the experience. A metaphor new to a learner has never been taught to him and this offers an ideal case to study the problem. The LSA theory predicts that the meaning of a metaphor can be derived from the knowledge of the vehicle and the topic. This prediction was tested by an experiment, the findings of which are consistent with LSA prediction that if a learner knows about the vehicle and the topic he can figure out the meaning of the metaphor. This study offers a solution to the problem of acquisition of knowledge.

Keywords
second language acquisition metaphor Latent Semantic Analysis

1 Introduction

Second language acquisition researchers face the same problem that first language acquisition researchers face: how do they account for the fact that second language learners seem to have knowledge that is not available in the linguistic data to which the learners are exposed? For example, just like native speakers of English, second language learners also have tacit knowledge of the difference between John is easy to please and John is eager to please although they have not been taught explicitly about it. This difference has been cited by Chomsky to support the view that much of the linguistic knowledge, especially syntactic knowledge is innate determined by the genetic code. Since the linguistic stimulus is not enough to develop the necessary knowledge needed in using a language, the problem has been termed "the poverty of the stimulus" by Chomsky

Recently Pinker 1994 has summarized the evidence on language acquisition and he concludes that language learning must be based on a strong and specific innate foundation a set of rules and principles that need parameter setting and filling in but not acquisition from experience. Although this "language instinct" position is debatable, it has an idea that we agree on, that in the minds of language learners there exists some powerful mechanism that can use the limited information they receive to turn themselves into competent users of language. What we want to study is what this mechanism is and how it works. In this respect, knowledge of metaphors such as My lawyer is a shark provides a dramatic case for research since novel metaphors are rarely taught to students in high schools or universities in China.

My lawyer is a shark

The study aims to answer two questions: Can second language learners understand metaphors new to them such as My lawyer is a shark? If they can, how do they figure out the metaphor meaning?

In this paper we suggest an account for the acquisition of knowledge of metaphors based on a
computational model of human knowledge representation and processing [the Latent semantic analysis (LSA) proposed by Landauer & Dumais 1997].

Our basic claim is that language learners understand metaphors by activating the associated features of the metaphor vehicle and topic and selecting the features that are closely related to the vehicle and the topic. For example, if a language learner knows much about the words lawyer and shark, then hearing metaphor [1] will activate features associated with lawyer and shark in the learner's mind and enable the learner to arrive at a correct understanding based on the combination of features that are related to both lawyer and shark. If no features are recognized to be related to lawyer or shark, then the learner would not be able to understand the metaphor.

Two basic assumptions undergird our view. First, a concept is constituted by a set of features or is associated with a set of features. Our understanding of a word or a sentence means that we have knowledge about the features of the concept. We understand the word shark because we know that a shark is associated with a set of features such as a shark is a fish, can swim, has sharp teeth, has leathery skin, has gills, is vicious, predatory, aggressive, tenacious etc. We not only know this information, we know it at the same time we hear the word. That is, we activate the features fast swim, mer, fish, sharp teeth, leathery skin, and gills, predatory, aggressive, tenacious when we hear the word.

Second, humans intuitively know that members of a concept can be described along certain specific dimensions and specific members of a concept usually have their dimensions unfilled, that is, the dimensions are empty. For example, the concept road can be described on the dimensions of shape, surface, width, cost and color, but other dimensions such as emotional arousal are not applicable to the concept of road. Thus, any one road is either straight, curved or twisting; it can be smooth or bumpy, narrow or wide, or cheap or expensive, black, white or gray. But none of roads can be said to be calm, or excited. And any two members of roads differ along these dimensions. We assume that when hearing the word road, we are likely to activate the dimensions such as shape, surface, width, and color or cost although the values of these dimensions are not specified.

In our view, sentence comprehension can be described as this. When we hear a sentence of the form "X is a Y," we activate the meaningful dimensions of the subject X and the associated features of the predicate Y. The sentence form itself makes us combine the features of Y and the dimensions of X, attributing relevant features of Y to the unfilled dimensions of X. If the predicate is a simple concept, if it has few limited features such as young or a teacher, young means having lived only for a short time. A teacher is one who gives lessons in a school, then the features of the predicate are attributed to the subject. When we hear the man is young or is a teacher, we simply activate the features associated with young or teacher and attribute these features to the subject the man. If the predicate is a complex concept, which has many different features such as a shark, a wolf, or a snake, then most of the associated features of the predicate may be activated, but not all the activated features are attributed to the subject only those appropriate for the subject. Thus, different features of snake play a role in understanding [2] and [3]. The subject selects those features of the predicate that are appropriate for it and inhibits the features that do not apply or apply less aptly, thus generating a contextualized word sense.

[1] The road is a snake.

We thus propose that comprehension is the activation of the features of the predicate and the dimensions of the topic followed by the linkage between the relevant features and dimensions. Consider the example from Glucksberg (1998). Glucksberg lists nine features of a shark: vicious, predatory, aggressive, tenacious, fast swim, mer, fish, sharp teeth, leathery skin, and gills. He argues for the view that metaphors are class inclusion assertions in which the topic my lawyer is assigned to a superordinate category and the vehicle shark refers to that category and is a prototypical exemplar of that category. On our approach, when we hear the metaphor, we activate the features of shark and the dimensions of my lawyer. We understand the metaphor because the dimensions of the topic are not specified initially and we select vehicle features that are appropriate for the topic my lawyer. The topic dimensions and the
vehicle features that are activated interacts with each other. Black 1962 so that of the features that have been activated only those that are appropriate for the topic that can be attributed to one of the dimensions of my lawyer can be strengthened and those inappropriate for the topic will be quickly deactivated. The strengthened features become the interpretation of the metaphor. Thus although the last five of the features mentioned above have a strong activation level they are deactivated quickly and get out of consciousness because they are not related to lawyer.

However, our proposal of comprehension as the activation of the features of the predicate and the argument does not specify computationally how the relevant properties are identified and how the irrelevant properties are inhibited. Thus we also need a model of comprehension that selects the right features automatically without having to be told what is relevant and what is not." Kintsch 2000 257.

Recently, Kintsch has proposed a computational model that can be used to understand how relevant properties are identified during metaphor comprehension. Kintsch 2000 2001. One major component of Kintsch’s Prediction Model is LSA which represents the meanings of words sentences and essays in terms of vectors in a semantic space and determines the relatedness of words and essays by computing the cosine between two vectors.

Gui Shichun 2003 is the first Chinese scholar to introduce the Latent Semantic Analysis or LSA and its potentials in language research to the linguistics community in China. Following his lead, other scholars have also explored LSA rationale and its possible applications. Wang Liu 2004 Li Zhang Li 2006. The research conducted so far does not however address the issue of second language acquisition of English by Chinese speakers. This is an issue that we examined in the present study.

LSA is a theory of knowledge induction and representation. Landauer Dumais 1997 which provides a method of determining the similarity of meaning of words and passages by analysis of large text corpora. After processing a large sample of written text, LSA represents the words used in it and any set of these words — such as a sentence paragraph or essay — as points in a very high dimensional semantic space. As a practical method for the characterization of word meaning, LSA allows us to closely approximate human judgments of meaning similarity between words and to objectively predict the consequences of overall word-based similarity between passages. Based on a powerful mathematical analysis the singular value decomposition SVD LSA is capable of correctly inferring much deeper relations thus the phrase “Latent Semantic Analysis” LSA indicates its representations of the meaning of words and passages from analysis of text alone. It is not a traditional natural language processing or artificial intelligence program. It uses no humanly constructed dictionaries knowledge bases semantic networks grammars syntactic parsers or the like and takes as its input only raw text parsed into words and separated into meaningful passages. LSA is capable of discovering deeper semantic relations between words or passages because it is proved that reducing the dimensionality the number of parameters by which a word or passage is described of the observed data from the number of initial contexts to a much smaller but still large number will often produce much better approximations to human cognitive relations.

To construct this semantic space it analyses word co-occurrences in a large number of written documents. Specifically, in this study we construct the semantic space by analyzing a corpus of some 1,000 documents containing over 8,000 word types — a total of about 250,000 word tokens. The following procedures are applied. From each article we took a sample consisting of the whole text or its first 250 words whichever was less for a mean text sample length of 246 words. The text data were cast into a matrix of 1,000 columns each column representing one text sample by 8,268 rows each row representing a unique word type. This matrix was then submitted to SVD singular value decomposition and the 400 most important dimensions were retained. The reduced dimensionality solution generates a vector of 400 real values to represent each word and each context. Thus the input to LSA consists of occurrence patterns over contexts LSA does not represent word meaning in terms of co-occurrence frequencies but as vectors in a semantic space of 400 dimensions. The meaning of a word or a combination of words a sentence is represented by a vector in the semantic space.

In the above section we discussed how a concept on being heard activates a set of features that are
closely associated with the concept but we did not say anything about how the activation level can be represented. Now that we have LSA we can solve this problem. The activation level corresponds to the cosine between the two vectors representing the two concepts. In the comprehension process initially the predicate concept will activate features that are closely related to it. Thus when we hear my lawyer is a shark our mind will activate features such as fish can swim has sharp teeth vicious aggressive etc. according to their relatedness to the predicate concept shark. Then the topic my lawyer will function to strengthen those features that are also closely related to itself. Thus although initially the activation levels of fish sharp teeth can swim are high these features will be deactivated and suppressed because they are not related to the topic my lawyer and cannot be combined with lawyer. But the features that are not strongly activated such as vicious aggressive will gain a strong positive activation value because they are at least somewhat related to lawyer. Kintsch 2001

Next features that attain a strong positive activation value are attributed to the topic thus entering the meaning of the metaphor. This means that if LSA knows about the topic and the predicate vehicle then based on the cosine between the topic and the vehicle features it can derive the meaning of the combination of the topic and the vehicle. The meaning of the metaphor is not based on LSA prior knowledge of the metaphor itself but on LSA prior knowledge of the individual topic and vehicle which LSA acquires through exposure to linguistic data. Thus it excludes the necessity for positing innate knowledge about the metaphor. This provides a solution to the problem of acquisition of knowledge much more knowledge than appears to be available in the data can be acquired when the deeper semantic relations are figured out by LSA. This is a major prediction of LSA about second language learners acquisition of the implicit knowledge necessary to understand metaphors. This is equivalent to saying that LSA acquisition can occur unconsciously based on the information available in the experience without being provided with the specific innate knowledge. We will test this prediction later in the paper. But we first describe how LSA algorithm characterizes the comprehension of metaphor.

2. LSA Simulations

The algorithm discussed by Kintsch 2001 is used in our calculation of the meaning of metaphors of the form X is a Y. The algorithm first selects features from the LSA space that are related to the vehicle and then selects from this set those features that are also related to the topic. The first step is achieved by computing the semantic neighborhood of the vehicle. The complete semantic neighborhood of a vehicle in the semantic space is a 400-dimensional hypersphere around the vehicle in which all 268 items in the semantic space are arranged according to their relationship with the vehicle. LSA can order all the items in the space according to their cosine with the vehicle generating a list of m words ordered in terms of their cosine with the vehicle. For metaphors m has to be fairly large 500 m 1500 because the vehicle and the topic often are quite unrelated.

The next step picks those terms from the neighborhood of the vehicle that are also related to the topic. The cosines between the m neighbors of the vehicle and the topic are computed and the k terms with the highest cosine are selected.

It is not necessarily the case that there exist terms related to both the vehicle and the topic. Thus in order to avoid introducing noise by selecting the strongest terms even when their absolute strength is low the terms selected must have a cosine with the vehicle and the topic above some minimum threshold. Only terms that have a cosine with the vehicle greater than two standard deviations above the mean for all words in the space used here will be included among the considered items. Similarly all items related to the topic with a below threshold cosine will be eliminated.

According to our theory of metaphor comprehension those features terms that are both related to the vehicle and the topic are attributed to the topic and enter the meaning of the metaphor. Figure 1 shows the features that are selected by the LSA program we have developed and that have a high activation level when the metaphor my lawyer is a shark is processed. Features like aggressive vicious predatory are attributed to my lawyer because they have a larger cosine value with lawyer.
As shown in Figure 1, the vectors for features swimmers, sharp teeth, fins and like dolphins, aggressive, vicious, and predatory were computed with k = 5 and m = 500. According to LSA, the first four irrelevant features are as strongly related to the vehicle shark as the metaphor relevant features; their average cosine with shark is \( \frac{365}{4} \) as the metaphor relevant features; their average cosine with lawyer is \( \frac{366}{4} \). But when they are combined with lawyer, the metaphor irrelevant features will be suppressed because they are unrelated to lawyer; their average cosine with lawyer is \( \frac{90}{4} \). And the metaphor relevant features will become more activated because they are more strongly related to lawyer; their average cosine with lawyer is \( \frac{310}{4} \).

The relevant features are moderately strongly related to both the vehicle and the topic. Thus, LSA selects these features aggressive, vicious, and predatory to attribute to the topic my lawyer, which means that LSA interpret the metaphor as meaning that my lawyer is aggressive, vicious, predatory.

As we discussed earlier, a predicate or a vehicle may have a set of features that are not all attributed to the topic. In the case of a metaphor, the calculation of the meaning of the sentence is the same. For example, the metaphors the road is a snake and the man is a snake, When the vehicle the snake is heard, it activates a number of features such as the Sshape, vicious, and wicked. These are features that are closely related to the vehicle snake. But our LSA program calculates the following results. The Sshape is closely related to the road, cosine road Sshape 056 vicious and wicked are not closely related to the road cosine road vicious 018. However, the Sshape feature is not closely related to the topic the man cosine man Sshape 025 and vicious and wicked are closely related to the topic the man cosine man vicious 065. Thus, according to our theory of metaphor comprehension, the Sshape is attributed to the road and vicious and wicked are attributed to the man.

These calculations show that our LSA program can derive knowledge of a metaphor based on its previous knowledge of the vehicle and the topic, without any previous knowledge of the metaphor. In other words, a mechanism that has understood two concepts, for example, a vehicle and a topic, can
understand the combination of the two concepts that is the metaphor. Does this prediction about the acquisition of metaphor bear out in humans?

3. Experiment

To test the LSA prediction that a mechanism can acquire the knowledge necessary to understand a metaphor without previous knowledge of it, we carried out an experiment in which participants' understanding of metaphors and their knowledge of the individual vehicles and topics are examined. The findings show that participants who have sufficient knowledge of vehicles and topics can indeed figure out the meaning of metaphors that they have not encountered before. This constitutes evidence for the LSA prediction.

3.1 Methods

3.1.1 Participants

22 postgraduate students at GUCAS all second language learners of English and all native speakers of Chinese took part in the experiment. None had participated in any previous studies of metaphor comprehension or language acquisition.

3.1.2 Materials and procedure

The experimental sentences were either cited from a well-known experimental paper by Glucksberg and Gildea (1982) or constructed in a similar way. However, the examples in Glucksberg et al. were all changed from plural forms. Some salesmen are bulldozers to singular form. The salesman is a bulldozer because many of the words involved were low frequency words and a preliminary analysis showed that LSA knew more about their singular forms than their plurals. bulldozers do not appear in the space used but bulldozer does. The 9 metaphors used in the experiment were

1. This job is a jail
   The salesman is a bulldozer
   The flute is a bird
   The girl is a bird
   The road is a snake
   The man is a snake
   The man is a lion
   My surgeon is a butcher
   My lawyer is a shark

   The analysis was performed exactly in the same way as described above for my lawyer is a shark that is with m 500 0 k 50.

   Each participant was invited into a computer room and seated in front of a computer screen where the following instructions and questions appeared

   All questions can be answered either in English or Chinese.
   What do you know about a lawyer? Please write as many features as you know about a lawyer.
   What do you know about a shark? Please write as many features of a shark as you know about a shark.
   What is your understanding of my lawyer is a shark? Please write out your understanding of it.

3.2 Results and discussion

Participants' understanding of the vehicle and topic and the metaphor is then coded on a scale of 5. Each feature provided by a participant is coded as one point. so a participant who gave five features of a shark (for example, has gills, has sharp teeth, can swim) is aggressive, predatory gained 5 points. And participants' comprehension of metaphors is also coded on a scale of 5 points. A participant who gave the correct interpretation the lawyer is aggressive was given 5 points. One that gave the incorrect interpretation the lawyer is a fast swimmer was given 2 points.
As shown in Table 1, the mean knowledge rating for the vehicles of the 9 metaphors was 3.27 for the topics 368. The mean comprehension ratings for the 9 metaphors was 4.04 SD 0.84 indicating that participants generally understood the metaphors. As reported in Table 2, a correlation for the data revealed that participants’ knowledge of the vehicle and their understanding of the metaphor are significantly related r 0.768 p < 0.005 (two tails). Similarly, participants’ knowledge of the topic and their understanding of the metaphor are significantly related r 0.652 p < 0.001 (two tails).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The mean knowledge ratings for the terms and metaphor comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td>knowledge ratings for topic</td>
<td>3.68</td>
</tr>
<tr>
<td>knowledge ratings for vehicle</td>
<td>3.27</td>
</tr>
<tr>
<td>comprehension of metaphors</td>
<td>4.0455</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Correlations between knowledge ratings for the terms and metaphor comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metaphor Comprehension</td>
<td></td>
</tr>
<tr>
<td>knowledge ratings for vehicle</td>
<td>0.768</td>
</tr>
<tr>
<td>knowledge ratings for topic</td>
<td>0.652</td>
</tr>
</tbody>
</table>

Pearson correlation is significant at the 0.01 level (two-tailed).

These findings suggest that participants who have a solid knowledge of the vehicle and the topic are generally able to figure out the meaning of the metaphor. The experiment findings are consistent with the LSA predictions that the meaning of a metaphor can be derived from knowledge of the vehicle and of the topic alone. This shows that participants’ knowledge of metaphors which is not taught to them explicitly can be acquired through a process of exposure to large amount of linguistic data.

4 General Discussion

This study provides at least one solution to the problem of “poverty of the stimulus” in language acquisition. How people acquire as much knowledge as they do on the basis of as little information as they get? Second language learners are not exposed to metaphors such as The road is a snake or The flute is a bird yet they can figure out the correct meaning by themselves without any help from others. Many linguists solution to the problem is to posit innate knowledge. Children are born with specific knowledge and need only hints to trigger it. Chomsky 1991 and Pinker 1994 assert that a child’s exposure to adult language provides inadequate evidence from which to learn either grammar or lexicon. Their language develops from a set of innate specific principles and parameters that needs only to be set on the basis of available data. Jackendoff 1982 claims that even the notion of possession is innate based on a comparison between the constructions of ‘have’ and the constructions of ‘get’.

Another group of researchers faced with the problem of origin of knowledge have proposed a constraint-based solution to the problem. Theorists such as Carey 1985, Clark 1987, Keil 1989, and Markman 1994 have hypothesized constraints on the assignment of meanings to words. For example, it has been proposed that early learners assume that most words are names for objects that any two words usually have two distinct meanings: that the same object is given only one name etc. However, there are several problematic aspects to constraint-based resolutions of the problem. One is whether a particular constraint exists as supposed. Is it true that young children or beginners of second language learning assume that the same object is given only one name and if so is the assumption correct about the adult language to which they are exposed? It is not in English or Chinese. Films, movies, motion pictures and flicks all refer to the same thing and anchors host announcers can all be translated into dog and collie.

In addition, suppose that young children or SL learners assume that there are no synonyms. How much could that help them in learning the lexicon from the language to which they are exposed? It could even pose a threat rather than a help to the learner by tempting the learner to assign too much or the wrong difference to our dog, the collie, and Fido. Thus, the constraint-based solution is not completely
satisfactory Landauer Dumais 1997

The solution we provide here does not suffer these problems. We do not need to postulate UG-type knowledge or innate constraint to account for the acquisition of knowledge. Knowledge is derived from mere exposure to very large amount of text. Evidence points to this conclusion: It is an obvious fact that all professors, researchers, reporters, the kind of people that achieve sort of “success” in language have definitely read a lot in either high schools or universities. An illiterate person rarely becomes a professor or a researcher. Why can knowledge develop from exposure to large samples of text? The answer given by LSAs is that when a human or a mechanism learns a lot about local relations between words or weakly related facts, deeper or latent semantic relationships or patterns can form from these local relations when they are aggregated by themselves. This formation of deeper patterns does not need external or internal force to help, they just form by themselves.

In the case of the acquisition of metaphors by second language learners, the meaning of a metaphor can be determined completely by the combination of the vehicle and the topic. If the learner knows about the vehicle and the topic, he can derive the meaning of the combination of the vehicle and the topic by activating the features of the vehicle and the dimensions of the topic and by selecting the features that can fill the dimensions of the topic. The meaning of a metaphor comes from the interaction of the vehicle and the topic and the vehicle and the topic are enough to produce the meaning of the metaphor. Second language learners’ acquisition of metaphors provides at least an answer to the knowledge origin problem.

References
Black M 1962 Models and Metaphors Ithaca NY Cornell University Press
Carey S 1985 Conceptual Change in Childhood Cambridge MA MIT Press
Glucksberg S 1985 Understanding metaphors Current Directions in Psychological Science 7 39
Kintsch W 2000 Metaphor comprehension A computational theory Psychonomic Bulletin Review 7 257
Markman E M 1994 Constraints on word meaning in early language acquisition Lingua 92 199
Pinker S 1994 The language instinct How the mind creates language New York William Morrow and Co

116