Symbol systems and perceptual representations

Walter Kintsch[⊗]

University of Colorado

The relationship between verbal mental representations and
action, perception and emotion is discussed. It is argued
that symbols are not directly grounded in the real world but
are made meaningful because of their relationship with
other symbols. Symbols are not defined by the real world,
but symbols mirror the real world. An isomorphism exists
between symbolic representations and the world of action
and perception, but the meaning of symbols cannot be
reduced to perception and action. This argument is
illustrated with some examples from LSA and related
models of verbal meaning.

The ability to represent the world symbolically, and hence abstract thought,
evolved from other, more concrete forms of representation. The ways in which people
model the world in their mind follows an orderly sequence, both phylogenetically and
ontogenetically. Donald (1991), for instance, distinguishes a sequence of cultures, each
characterized by a particular form of mental representation. All animals learn and have
procedural memories. At some level (certainly at the level of primates) animals are able
to represent the world in terms of generalized records of past experience that allow them

to react directly to situations in a way that summarizes their experience with that situation. This is also true for young children (Nelson, 1996, calls this general event memory). As biological evolution continues, representational systems based on intentional imitation emerge. Speech appears at this stage, but its purpose is social communication, not representation. Individuals at this level are characterized by representation-through-action; cultures are characterized by crafts, arts, and ritual. At this point biological evolution is replaced by cultural evolution: at the sensorimotor level humans are simply primates, but the use of language distinguishes them more and more from their animal ancestors. Donald makes an important distinction between narrative language and theoretic language use.[1] Narrative language is based on mental models that are fundamentally linguistic. Flexible computations can be performed with models of this kind, giving humans a degree of control over the world that is impossible without language. Words and thoughts are inseparable at this level. Most human cultures and human individuals are capable of working at the level of narrative culture. However, few cultures, and by no means every individual in these cultures, achieve the level of formal representation, that Donald terms theoretic culture. The engrams of memory are aided by the exograms procured by our technology at this level. Most of what we do in schools is trying to bring students up to that level of cognitive functioning.

This is not the place to fill in the details of this conceptual framework (see Donald, 1991, and Nelson, 1996). However, a few comments are in order. First, it must be pointed out that when an individual or a society moves from one level of representation to the next, the earlier level does not disappear, but becomes embedded in the new forms as a new kind of representation is added to the behavioral repertoire and integrated with the earlier ones. Educated individuals in our society who function at the level of formal reasoning at the same time are also capable of direct action, participate in art and ritual, and share the narrative traditions of our culture. No wonder psychologists have a hard time analyzing human behavior. Thus, although event memory is not basically linguistic, it may become linguistically encapsulated: what we remember is

---

[1] Several authors have made the same or similar distinctions, e.g. Bruner (1986) with respect to language, or Vygotsky (1987) with respect to everyday and academic concepts.

sometimes not the event itself but our (explicit or implicit) linguistic description of the event. Furthermore, when the brain develops new forms of mental representation, it not only does not discard the older ones, but in fact uses the old structures for the new purposes. Thus, cognition is indeed embodied.

How the different levels of representation interact in detail is not known at present, but much research, including the present workshop is directed at this goal. The question that I want to address here is how to model these overlapping, intertwined mental representations, particularly linguistic representations.

If we take seriously the levels of representation analysis that was sketched above, there seem to be two kinds of approaches to the question of mental representations. On the one hand, one could try to model a system of mental representation that combines elements from all of these levels. Thus we would have images, concepts, somatic markers, S-R links, features, schemata, abstractions, words, etc., etc., all linked and somehow organized in one grand system, at least in principle, for I have no idea how one would go about constructing such a model. Alternatively, one could model each level of representation separately. This has the advantage that at least for linguistic representations, we do have an idea how to begin to construct such a model: I'll describe Latent Semantic Analysis and its variants shortly. But first I must address an obvious problem created by the divide-and-conquer strategy. If we model separately the level of linguistic presentation, how can it be coordinated with other levels of representation? In behavior and the brain these levels function all together and are not easily separated.

One can think of the different levels of mental representations as a set of coordinated maps. Consider the relationship between the level of linguistic representation (as captured by LSA, for example) and other levels of representation, which we can lump together here as the "real world". The relationship between the symbolic representation and the real world could be a fist-order isomorphism, where the symbol in some way (as by a picture, feature system, or dictionary definition) captures meaning directly. The meaning of a word would be its referent in the real world, as in the

perceptual symbol system of Barsalou (1999) and similar proposals. This is what I believe people are saying when they claim that language must be "grounded" (Harnard, 1990). An alternative way of thinking about maps and their relation to the real world is in term of what Shepard (1987) called a second-order isomorphism. In a second-order isomorphism the meaning of a symbol is not defined through its reference to another level of representation, but through its relationship with other symbols. The meaning of a word lies in its relationship with other words. Specifically, in LSA the meaning of a word is a vector of 300 numbers that are entirely vacuous by themselves and acquire meaning only because they specify how this vector is related to all the other words and texts in the LSA space. In this scheme the real world does not provide the meaning of a symbol, but symbols and referents are nevertheless coordinated as a second-order isomorphism. On a two-dimensional map, I can trace a path from Boulder to Denver that is isomorphic to the trip from here to there in the real world. In the 300-dimensional LSA space I can compute the semantic distance between these words and expect it to correspond to features of the real world (the cosine between Boulder and Denver is .21, way above the cosine between Boulder and farther away locations such as Tenerife (.01) and Spain (.02) (for similar examples, see the contributions of Zwaan and Louwerse to this volume).

Calling something a second-order isomorphism does not solve the problem of how symbols and the real world are related. That question remains an open research question. But it does alter the way one approaches the problem. The meaning of symbols is not to be reduced to perceptual features or actions; raher the question becomes one of correspondence between different levels of representation.

I shall first sketch a proposal for modeling verbal meaning. This proposal has two components. First, we must describe how people induce verbal meaning – the representation of meaning in memory. That is the goal of LSA and related systems. Second, we must model how stored meaning is used to construct contextually appropriate meanings of words, sentences, and texts in general. It is of course possible that all we need for language is already stored in memory and that meanings simply have to be retrieved from the semantic store ready-made. The alternative presented here is that the

semantic store only provides the raw material for the construction of meaning, and that meaning emerges when words, sentences, and text are used in context.

After this sketch of symbolic verbal meaning, I shall discuss the proposition that, although language is not grounded in the sense of a first-order isomorphism, language mirrors perceptual features of the world with a high degree of fidelity.

Memory for verbal meaning: LSA and its Extensions.

Our goal is to construct a representation of the meaning of words and texts that can serve as a model for the kind of representation the human mind builds. People generate such representations while interacting with the world and other people, listening, speaking and reading, mostly without explicit instruction. A computer simulating this process is restricted to reading texts, but must be able to generate a semantic representation without guidance, as people do.

The kind of representation generated depends in part on the precise mathematical algorithm used to generate it, but more significantly, upon the language input used. We start with a large set of texts (e.g., 50K documents, containing 100K different words and over 10M words total) that is analyzed into a document-by-word matrix, whose cell entries are the frequencies with which each word appears in every document. Note that this is by no means all the information in the corpus – all we look at is word co-occurrences, neglecting word order, syntax, as well as discourse structure. Our input matrix is huge, $50,000$-by-$1000,000$, and sparse, that is, most entries are 0. We then reduce the dimensionality of this matrix, down to about 300 dimensions. That is, we express the meaning of a word by a vector in a space with far fewer dimensions. This has two effects. First, it is a process of abstraction. We are not interested in the particular documents used to construct the semantic space, the particular topics the authors wrote about, their particular word choices; instead we are interested in the general semantic properties of words. Reducing the dimensionality of our corpus has the effect of discarding a great deal of idiosyncratic information about word use in this particular

corpus, while retaining the essential semantic information about word meanings. Secondly, dimension reduction is a process of generalization. While the original input matrix was sparse, the reduced matrix is filled in, so that we have a measure of the semantic distance for every word pair, even though most pairs by far have never co-occurred together in any of our documents. Thus, *doctor* and *physician* have a high cosine of .61, even though they have rarely been used together in a single document.

There are various ways to perform this dimension reduction. LSA (Landauer & Dumais, 1997; Landauer, McNamara, Dennis, & Kintsch, 2007) starts with a term x document matrix M of size $n \times m$ that is decomposed via singular value decomposition. Typical values of n and m are 100,000 and 50,000, respectively. All eigenvectors but the ones corresponding to the d (300 or so) largest eigenvalues are discarded. A word or document is thus represented by a vector of 300 real numbers. This vector is not by itself interpretable but its cosine (or other measure) with other vectors defines its position in the semantic space: hence meaning is defined by its relationship with other vectors.

Independent Component Analysis (e.g., Stone, 2004; for an application of ICA to language analysis see Mangalath, 2007) assumes that each observation (word, document) is a mixture of independent semantic elements (components, topics). Several constraints are used to unmix the observations: components are chosen so that they are statistically independent (not merely uncorrelated). This is achieved by searching for components that are maximally non-gaussian, and/or the least complex ones, that is, the most compact or predictable ones.

The ICA model is conceptually related to the Topics Model (Steyvers & Griffith, 2007) where meaning is represented as a mixture of topics. The distribution of topics over words and documents is based upon a Bayesian learning algorithm. An appealing feature of both models is that components or topics are individually interpretable. Thus, the linguistic corpus we are working with most of the time can be analyzed into a set of topics, such as *drugs, colors*, or *doctor visits*, or, alternatively into a mixture of independent components.

A quite different approach has been taken by Jones and Mewhort (2007). Word meaning can be represented as a composite distributed representation by coding word co-occurrences across millions of sentences in a corpus into a single holographic vector per word. This representation is a pattern of elements that stores the word's history of co-occurrence and usage in sentences.

Which one of these methods - LSA, ICA, Topics, or Holograph - yields the best semantic representation, or the right one? Experience from working with these models suggests that they produce very similar results. Each has certain advantages for certain purposes, but basically they paint the same picture: what is related in one way in one model is similarly related in the other. There have been no formal comparisons among these approaches, however, so one must be careful with this conclusion, but that is the impression I have at this point. I take this as a positive indication that whatever semantic space we generate is not an artifact of the method used, but truly reflects the semantic information in the corpus.

What does make a striking difference, in contrast, is to enrich the input to the analysis, so that it contains not only word co-occurrence information, but also word order information. This is done in the Holograph model by adding to the item vector all convolutions of the word with the other words in the sentence. This allows the system to use order information and to infer grammatical categories. It also provides a much better account of human priming data (Jones, Kintsch, & Mewhort, 2006). If only context information is used in the holographic model (as well as in all other such models, like LSA) associative priming (*bee-honey*) is reasonably well predicted, but not semantic priming (*deer-pony*). On the other hand, if only order information is used in the holograph model (or in n-gram models, or in HAL; Lund & Burgess, 1996) the models handle semantic priming but not associative priming. But if order and context information is combined in the holograph model, a wide range of experimental data involving both associative and semantic priming can be accounted for.

Including order information in a latent semantic structure allows us to model a far greater range of phenomena than before. But the basic unit of analysis in such an expanded model still remains the word, and it has often been argued that the unit of language and cognition is the proposition, not the word (among others, by Kintsch, 1974, 1998). Recent advances in machine learning have made it possible to construct LSA like systems that represent propositional information in the form of dependency trees. Linguists analyze the syntactic structure of a sentence as a phrase structure tree. A dependency tree (Yamada & Matsumoto, 2003) is a weak form of a phrase structure tree, lacking the phrasal nodes, and thus does not represent all the relevant linguistic distinctions, but the dependency model does retain information about dependencies among words, i.e., about propositional units. I illustrate this relationship with a simple example. Figure 1 shows the dependency tree for the sentence "*Rolls-Royce said it expects its US sales to remain steady at about 1,200 cars*," after Yamada and Matsumoto (2003)[2]. In Figure 2, I show the proposition list corresponding to that sentence according to Kintsch (1998), and in Figure 3 I superimpose the propositional structure on the dependency tree. The dependency tree shows which words in the sentence belong together as a propositional unit. To obtain a dependency tree, a dependency parser is trained on a large corpus of sentences with support vector machines (Yamada & Matsumoto, 2003; Nivre, Hall, & Nilsson, 2006). This approach is a departure from the purist path of unguided learning, but that is the only way we can do this analysis at present. Mangalath, in our lab, has begun to explore how the information provided by this dependency parse can be incorporated into the systems discussed above to generate a semantic structure that represents word information as well as propositional information. Basically, we analyze, in addition to the word-by-document matrix, a word-by-dependency matrix. Normally, a sentence like "*The ferocious lion killed an antelope*" would make both *lion* and *antelope* a little bit more *ferocious*; however, since *ferocious* is dependent on *lion*,   only the *lion* should become a more *ferocious*, not the *antelope*.

Figure 1, 2 & 3

---

[2] The actual dependency tree includes parts of speech tags which are neglected here.

Thus, Mangalath uses syntax (in the form of dependency trees) to guide what is learned. But syntax can also constrain interpretation. It is possible to use the syntactic structure of a sentence to guide the construction of its meaning. We no longer have to add up the vectors of all the words in a sentence to arrive at the sentence meaning, but we can use the syntactic information to generate an interpretation of a sentence in much the same way as we use it in the construction of the semantic space in the first place. Thus, we turn to the construction of meaning – word meanings, sentence meanings, and text meanings.

The Construction of Meaning

Words have often more than one meaning and several different senses. In LSA as well as in the other methods discussed above, a single vector represents the meaning of each word. It is like a dictionary that lists each word, but is silent about its different meanings and senses.

Instead of listing word meanings and senses, as in a conventional dictionary or in a mental lexicon, meanings can be generated in context. While the semantic representation of a word is a single vector that combines all meanings and senses, a context appropriate meaning is generated every time the word is used in a different context. I have described such a model for a generative lexicon in Kintsch (2001; 2007; in press). Briefly, it allows the context to modify word vectors in such a way that their context appropriate aspects are strengthened and irrelevant ones suppressed. In the Construction-Integration model of Kintsch (1998), discourse representations are built up via a spreading activation process in a network defined by the concepts and propositions in a text. Meaning construction works in a similar way: a network is constructed containing the word to be modified together with its semantic neighborhood and linked to the context; spreading activation in that network assures that those elements of the neighborhood most strongly related to the context become activated and are able to modify the original word vector. Consider the meaning of *bark* in the context of *dog* and in the context of *tree*. The semantic neighborhood of *bark* includes words related to the *dog*-meaning of *bark*, such as *kennel*, and words related to the *tree*-meaning, such as

lumber. To generate the meaning of *bark* in the context of *dog*, all neighbors of *bark* are linked to both *bark* and *dog* according to their cosine values. Furthermore, the neighbors themselves inhibit each other in such a way that the total positive and negative link strength balances. As a result of spreading activation in this network, words in the semantic neighborhood of *bark* that are related to the context become activated, and words that are unrelated become deactivated. Thus, in the context of *dog*, the activation of *kennel* increases and the activation of words unrelated to dog decreases; in the context of *tree*, the activation values for *lumber* increases, while *kennel* is deactivated. The contextual meaning of $bark_{dog}$ is then the centroid of the bark and its dog-activated neighbors, such as *kennel*; that of $bark_{tree}$ is the (weighted) centroid of bark and neighboring words like *lumber*. $Bark_{dog}$ becomes more *dog*-like and less *tree*-like; the opposite happens for $bark_{tree}$. Two distinct meanings of *bark* emerge: using the 6 most highly activated neighbors to modify *bark* from a neighborhood of 500, the cosine between $bark_{tree}$ and $bark_{dog}$ is only .03. Furthermore, $bark_{dog}$ is no longer related to *tree*, cos = -.04, and $bark_{tree}$ is no longer related to *dog*, cos = .02.

Figure 4

Figure 4 illustrates a different example, where the meaning generated is metaphorical rather than literal (Kintsch and Bowles, 2001; Kintsch, in press): the meaning of *shark* in the context of *My lawyer is a shark*. The items that *lawyer* selects from the neighborhood of *shark* make $shark_{lawyer}$ more aggressive, and vicious, and less fish-like. Thus, context appropriate word meanings can be generated within a system like LSA, in spite of the fact that the vectors constructed by LSA are context-free word vectors. The use of context-free representations in memory allows us to deal with polysemy. We no longer have to decide how many meanings and senses a word has, or how they are to be retrieved when needed. Instead, we only have to deal with a single, context-free word vector plus a process (predication) that generates emergent contextual meanings from that vector.

Predication models how word meanings are constructed in context. How about sentence meaning, and the meaning of texts? LSA has proven to be so useful because it has a very simple and effective way of constructing the meaning of texts: it sums the vectors of all the words involved. This neglects syntax and neglects text structure – but if the texts are long enough it produces excellent results, as attested by the commercial success of Pearson Knowledge Technologies (www.knowledge-technologies.com) and the educational success of *Summary Street* (Wade-Stein & E. Kintsch, 2004; Franzke et al., 2005). However, the vector sum of the constituent words is not a useful representation when we are dealing with single sentences or brief texts. The syntax plays a large a role in this case, whereas it apparently averages out over longer text, so that its neglect does not cause fatal problems.

We can use syntax to guide the construction of sentence meanings. As in the construction-integration model, we can use the dependency analysis to construct propositional units and then we integrate these units as required by the task at hand. I shall illustrate this model with two simple examples.

In a cloze test students are given a sentence with a word missing and are asked to fill in the missing word from a given set of alternatives. For example, the sentence might be

If you try, you can _____ any problem.

with the alternatives *solve, quack, flour*, and *seem*. One way to model this task is to compute the cosine between the sentence fragment and the four response alternatives in LSA. As Table 1 shows, this leads to the incorrect choice of *seem*. However, we get a different prediction if we use the dependency analysis of the sentence to focus our comparison on the relevant proposition rather than on the whole sentence. Figure 5 shows the dependency tree for this sentence, indicating the proposition containing the target word. This analysis suggests computing cosines between the four response alternatives and *problem*, the relevant content word in the proposition containing the target word. We neglect function words in this comparison because they would only add noise. Note, however, that function words were crucial in the construction phase: they

determine what kind of dependency tree is constructed and hence, which content words enter into the integration phase. Table 1 shows that the correct alternative has the highest cosine with the relevant proposition fragment.[3]

Table 1; Figure 5

A second example involves the implications sentences may have. Suppose we have in our corpus the sentence "*Oracle fought to keep the forms from being released*". Our task it to determine whether the inferences "*Oracle released a confidential document*" and "*Oracle fought to keep the forms released*" are true or false. Both inference sentences have a high cosine with the original sentence, cos = .24 and .93, respectively, and hence appear to be implied. However, the dependency trees for these sentences shown in Figure 6 and 7 suggest otherwise. The inference sentence "*Oracle released a confidential document*" has dependencies *Oracle-released, released-document, document-confidential* which do not match any of the dependencies in the original sentence. For the second inference sentence, "*Oracle fought to keep the forms released*", the dependencies between *oracle, fought, keep* and *forms* are the same as in the original sentence, but not the rest: in the original sentence there are dependencies between *keep-from* and *from-released*, which are not matched by the *forms-released* dependency in the inference sentence. Hence both inferences can be rejected as invalid. Function words play a crucial role in this process: they do not enter the comparison process (function words like *from* in LSA or ICA tend to be uninformative, having been experienced in all kinds of different contexts), but they determine what is compared with what in a sentence.

Figures 6 & 7

The point I am making with these examples is this. We have fairly good models of how people store verbal symbolic information in memory, and of how this information

---

[3] It should be noted that there are other methods that allow LSA to solve cloze problems, using n-gram or word-order information.

is used in constructing the meaning of words, sentences, and texts. The question is, do we have to throw away these models, because "No matter how LSA, HAL and other ungrounded symbol theories are extended and modified, ungrounded arbitrary symbols cannot be an adequate basis for human meaning" (Glenberg & Robertson, 2000, p. 397). My answer is "not at all" – this claim and similar claims by others, is ungrounded and based on a misconception about the relationship between symbolic and non-symbolic representations.

Language as the mirror of the mind

When people remember a list of words, when they read a sentence, when they listen to a story they do not function solely at the symbolic level. They use imagery to make words more memorable (Paivio, 1969), they construct situation models that are multimedia creations (Bransford & Franks, 1971; Zwaan , Madden, Yaxley, & Aveyard, 2004; Zwaan, this volume); and perceptual and motor areas in the brain become activated (Pulvermüller, this volume), as was demonstrated by several authors in this conference. All levels of representation humans are capable of appear to be functioning in a well coordinated chorus. How can a purely symbolic model, such as the one outlined above, do justice to these multilevel processes? The brief answer is that it does not and can not. The more interesting answer is that language has evolved in the service of perception and action, and that it is capable of mirroring aspects of perception and action. Yes, words can distort and bias meaning, but language would not be the great human success story it is if it were incapable of expressing what we experience faithfully.

There have been a large number of claims, such as that of Glenberg and Robertson (2000), that purely verbal, in other words, symbolic representations of meaning are insufficient to model meaning. One the one hand, this is a truism because human meaning representations are multilevel and not purely symbolic; no one can doubt that we live and act in a non-symbolic world. On the other hand, the demonstrations of the shortcomings of verbal meaning are often misleading, based on much too crude a

conception of verbal meaning. Language does a better job than what it is given credit for by the critics of symbolic representations. Louwerse (2007; this volume) has argued this point with some compelling examples. I would like to add two more examples, one involving only word meanings, where LSA provides an account that closely mirrors perception, and one involving sentence meanings, where the CI-model is shown to mirror affordances.

Salomon and Barsalou (2001) convincingly demonstrated that what things look like is an important factor in verbal tasks of property verification. Concepts, certainly concrete concepts, are represented both perceptually and symbolically, which plays an important role not only in perception and action, but also when we talk and think about a concept. English has only a single word *wing*, but a *wasp wing* and a *butterfly wing* look different, and this difference is reflected in the mental representation of these concepts. Solomon and Barsalou showed this in a priming study, where participants first verified either a concept-property pair such as *wasp wing* or *butterfly wing* and then a target pair such as *bee wing*. The perceptually similar *wasp wing* significantly primed *bee wing*, but the perceptually dissimilar *butterfly wing* did not. Furthermore, Solomon and Barsalou showed that perceptual similarity mattered in this case, not merely conceptual similarity, because no priming differences were observed when there was no perceptual discrepancy. In the case of our example, *butterfly body* and *wasp body* primed *bee body* equally well – presumably, because the bodies of a wasp, bee and butterfly look more or less alike, unlike their wings.

LSA cannot see wings, but it has indirectly encoded some of the perceptual information needed to know that *wasp wing* and *butterfly wing* are really quite different, whereas *wasp body* and *butterfly body* are not. If we compute a vector for *wasp wing*, *butterfly wing*, and *bee wing*, using the predication procedure described above, we can compute how close to each other these vectors are in the LSA space. The cosine between *wasp wing* and *bee wing* is .66, whereas the cosine between *butterfly wing* and *bee wing* is only .48. Hence, *wasp wing* could be expected to be a better prime for *bee wing* than *butterfly wing*. On the other hand, both *wasp body* and *butterfly body* are equally close to

bee body in the LSA space. The cosines are .94 and .91, respectively – LSA really can't tell the difference between these concepts (and neither can most of us.)[4]

I am not saying that the participants in the Salomon and Barsalou experiment did not employ perceptual representations when verifying properties; they probably did. My claim is simply that even a purely verbal symbolic system like LSA would behave in much the same way because the verbal information mirrors perception to a considerable degree. This is also the case with my second example from Glenberg & Robertson (2000), who claim that purely symbolic, verbal representations cannot recognize perceptual affordances. Perceptual affordances are obviously based on perceptual and action experience, but these experiences are mirrored in our language, so that even a purely symbolic system like LSA or ICA can detect that some of the sentences below are more sensible than others in the context of a story of someone being caught in the rain:

   *(a) Being clever, he walked into a store and bought a **newspaper** to cover his face.*

   *(b) Being clever, he walked into a store and bought a **matchbox** to cover his face.*

   *(c) Being clever, he walked into a store and bought a **ski-mask** to cover his face.*

Both (a) and (c) are possible ways to cover one's face, (c) much better than (a); (b) does not make sense. A ski-mask or even a newspaper afford some protection from the rain, whereas a matchbox does not. The affordances inherent in an object are based on experiences, experience in the real world with newspapers and matchboxes, and experience with how the words *newspaper* and *matchbox* have been used. Glenberg and Robertson present an LSA analysis that fails to account for the difference between these and related sentences: if one computes the cosine between the bold-faced target words and the rest of the sentences in the examples above, they do not differentiate between sensible and nonsense alternatives (Table 2). However, when these sentences are analyzed in terms of their propositional constituents, the difference in affordances emerges quite clearly.

Figure 8; Table 2

---

[4] A more detailed analysis of the Solomon and Barsalou (2001) results is given in Kintsch (2007).

Figure 8 shows the dependency tree for *(a)*. The figure shows that the relevant proposition involves the target words and *(to) cover (his) face*. If you simply take the cosine between the three targets and the whole sentence, you would conclude that (b) was the best alternative and (c) the worst. Glenberg and Robertson have obtained sensibility ratings on a scale of 1 to 7 which show that (c) is highly sensible, (a) is less sensible, but (b) is clearly not. An ICA (or LSA) analysis that respects the propositional units obtained by a dependency analysis yields similar results, simply because, according to our experience with language: *cover-face* has a high cosine with *ski-mask*, not quite as high a cosine with *newspaper*, but a low cosine with *matchbox*.

Table 2 also shows two other examples from the Glenberg and Robertson study. For *She gave him a ——— to play with*, people accept the target items *red beanbag* and *plastic spoon* as sensible, but not *large refrigerator*. The dependency analysis suggests that the target item is to be compared with *(to) play (with)*. This analysis correctly yields a high cosine for *red beanbag*, and a low cosine for *large refrigerator*, but the cosine for *plastic spoon* is not as high as it should be according to the sensibility rating people provided.

For *Adam pulled out of his golf bag a _____ and used that to chisel an inch of ice off his windshield*, the target items are *seven iron*, *ham sandwich*, and *screwdriver*. The dependency analysis shows that *and* dominates both *pull*-the-target-item and *used that to chisel an inch of ice off his windshield*. This comparison correctly selects *screwdriver* as the most sensible completion, and *seven iron* as also possible, but rejects the *ham sandwich*, much like people do. Overall, ICA correctly selects the completion that is both related and afforded in all of the experimental examples and rejects the unrelated, unafforded alternative. The cosines for the afforded but unrelated alternative are more variable, as we have seen in the examples above. Perceptual representations obviously add something to symbolic representations, but purely symbolic representations do a remarkably good job of discriminating between perceptually afforded and unafforded alternatives.

Let me briefly comment on some other criticisms of symbolic representations that I think are misguided. The objection is not only to ungrounded symbols, but also to arbitrary symbols. Apparently, icons are regarded as somehow better, easier, more natural than symbols. To comprehend meaning entails understanding the difference between a symbol or icon and the real thing, that is, dual representation. That in fact may be easier when the symbol is arbitrary (as words are) than when it is not (as pictures are). Children as old as 2.5 years are apt to mistake photographed objects for the object itself and treat small toys as if they were much larger (DeLoache, 2004). They do not confuse an arbitrary sound with the object it represents, but do attempt to sit down on a miniature toy chair. Babies confuse pictures with the objects they depict – American babies who are familiar with pictures as well as Ivory Coast babies who are not. It is not at all trivial to learn that concrete, non-arbitrary icons represent something that they are not. Teaching math with manipulatives can be counterproductive because the children learn to manipulate objects without ever inferring abstract mathematical principles. The consequences of the confusion between objects and their representation can be serious, as when dolls are used in courtrooms to represent a child's body – but the child is unable to think of the doll as both a doll and a representation of herself (Ceci & Bruck, 1995). Icons and analog presentations are not necessarily easier than digital representations or arbitrary symbols.

The fact that symbolic processes involve the same brain areas as action and perception (e.g. Pulvermüller, this volume) does not imply that symbolic processes and sensory-motor processes are the same. In the course of evolution nature discards very little, the new is generally built upon the old, fins turn into legs, and sensory brain areas become involved with symbol processing. But symbol processing is still at a different level than sensory-motor processes. Other primates and humans are very similar in their sensory-motor processes, but differ radically in their symbolic processing. Humans operate at multiple levels of representations, from the most primitive ones shared by all animals to the symbolic level. But that does not justify reducing symbolic systems to sensory-motor processes.

17

Rejecting a reductionism that denies the autonomy of symbolic processes does not imply that we should not make every effort to gain a better understanding of how symbolic and non-symbolic representations interact. To learn more about how perceptual representations and symbolic representations are coordinated and how they interact, is a very important goal, and considerable progress has been made as the papers presented in this workshop attest, both at the level of brain processes, as in the work of Pulvermüller already cited, as well as the behavioral level (e.g., Zwaan, this volume). This problem has also been addressed at the level of computational modeling. Goldstone, Feng & Rogosky (2005) explicitly address the problem of connecting words to each other as well as to the world. Howell, Jankowicz, and Becker (2005) have described a neural network model that simulates the acquisition of language based on prelinguistic concepts. A group of linguists (Talmy, Lakoff) and computer scientists (Feldman, Narayanan) at Berkeley have developed what they call a "simulation semantics" to reflect their belief that much of language understanding involves embodied enactment (e.g. Feldman & Narayanan, 2004; for more information see http://www.icsi.berkeley.edu/NTL). In simulation semantics the mind simulates the world while functioning in it. Concrete action schemata, for instance, become symbolically extended. Understanding at the symbolic level may involve this kind of simulation, but it can also remain at the purely symbolic level – the world that LSA describes. Research like this serves as an existence proof that it is possible to model the interface between symbolic and non-symbolic representations.

The sensitivity of language to perceptual information (as well as emotion and action) should surprise no one. Most of the words we know we have learned from reading, which precludes a direct perceptual association. And for the words that we know both at the verbal level and the perceptual and action level, the crucial anchor words that link these different levels of representation, language has encoded in its own way the information it needs to mirror the world. I conclude with two quotes that Chomsky used in his discussion of language (Chomsky, 1966) and that I have also cited in Kintsch (1998):

Les langues sont le meilleur miroir de l'esprit humain.

(Languages are the best mirror of the human mind.)

<div align="right">

Gottfried Wilhelm Leibniz, 1765

</div>

Der Mensch lebt mit den Gegenständen hauptsächlich, ja, da Empfinden und Handeln in ihm von seinen Vorstellungen abhängen, sogar ausschliesslich so, wie die Sprache sie ihm zuführt.

(We interact with objects mainly as they are represented by language; indeed, exclusively so, since perception and action depend upon memory images.)

<div align="right">

Wilhelm von Humboldt, 1792

</div>

REFERENCES

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-660.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331-350.

Bruner, J. S. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: The scientific analysis of children's testimony*. Washington, D.C.: American Psychological Association.

Chomsky, N. (1966). *Cartesian linguistics*. New York: Harper.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*, 66-70.

Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.

Feldman, J., & Narayanan, S. (2004). Embodied meaning in a neural yheory of language. *Brain and Language, 89, 385-392*.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary
Street®: Computer support for comprehension and writing. *Journal of
Educational Computing Research, 33*, 53-80.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A
comparison of high-dimensional and embodied theories of meaning. *Journal of
Memory and Language, 43*, 379-401.

Goldstone, R. L., Feng, Y., & Rogosky, B. (2005). Connecting concepts to the world and
each other. In D. Pecher & R.A. Zwaan (Eds.) Grounding cognition: The role of
perception and action in memory, language, and thinking. Cambridge:
Cambridge University Press. (pp. 292-314).

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*, 335-346.

Howell, S. R., Jankowicz, D., & Becker, S. (2005) A model of grounded language
acquisition: Sensorimotor features improve lexical and grammatical learning.
Journal of Memory and Language, 53, 258-276.

Jones, M. N. & Mewhort, D. J. K. (2007). Representing word meaning and order
information in a composite holographic lexicon. *Psychological Review*, 114, xxx-
xxx.

Jones, M. N., Kintsch, W., & Mewhort, D. J. K. (2006). High-dimensional semantic
space accounts of priming. *Journal of Memory and Language*, 55, 534-552.

Karmiloff-Smith, A. (1992). *Beyond modularity*. Cambridge, MA: MIT Press.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge
University Press.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*, 173-202.

Kintsch, W. (2007). Meaning in context. In T. K. Landauer, D. McNamara, S. Dennis, &
W. Kintsch (Eds.), *The handbook of Latent Semantic Analysis*. Mahwah, NJ:
Erlbaum. Pp. 89-106.

Kintsch, W. (in press). How the mind computes the meaning of metaphor: A simulation
based on LSA. In R. Gibbs (Ed.), *Cambridge Handbook of Metaphor and
Thought*. New York: Cambridge University Press.

Kintsch, W., & Bowles, A. R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, *17*, 249-262.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211-240.

Landauer, T. K., McNamara, D., Dennis, S., & Kintsch, W. (Eds.). (2006). *The handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum.

Louwerse, M. M. (2007). Symbolic and embodied representations: A case for symbol interdependency. In T. K. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.) *The handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum. Pp. 107-120.

Lund, K., & Burgess, C. (1996) Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instrumentation, and Computers*, 28, 203-208.

Mangalath, P. (2006). *Beyond Latent Semantic Analysis: Cognitive component resolution with independent component analysis*. Thesis, University of Colorado.

Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind*. New York: Cambridge University Press.

J. Nivre, J. Hall, & J. Nilsson (2006). MaltParser: A data-driven parser-generator for dependency parsing. In *Proceedings of the fifth international conference on Language Resources and Evaluation (LREC2006)*, May 24-26, 2006, Genoa, Italy.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*, 241-263.

Salomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive Psychology, 43*, 129-169.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*, 1317-1323.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. K. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *The handbook of Latent Semantic Analysis: A road to meaning*. Mahwah, N.J.: Erlbaum. Pp. 427-448.

Stone, J. V. (2004). *Independent component analysis*. Cambridge, MA: MIT Press.

Wade-Stein, D., & Kintsch, E. (2004). Summary Street: Interactive computer support for writing. *Cognition and Instruction, 22*, 333-362.

Yamada, H., & Matsumoto, Y. (2003). Statistical dependency analysis with support vector machines. *Proceedings of IWPT*, Nancy, France, 195–206,

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science, 28*, 611-619.

Table 1. Cosines between four response

alternatives and the sentence fragment

"If you try, you can _____ any problem."

	LSA(sentence)	LSA(dependency)
solve	.49	**.88**
quack	.18	.08
flour	**.**15	.03
seem	**.58**	.26

Table 2. Three examples from Glenberg & Robertson (2000) with LSA cosines between the sentence frames and the target items, sensibility ratings of the target items for each sentence frame, and ICA cosines between the target words and the relevant portion of the sentence frames.

	LSA cos (sentence)	Sensibility Rating (Glenberg & Robertson)	ICA cos (dependency)
Being clever, he walked into a store and bought a _____ to cover his face.			
newspaper	.35	5.08	.18
matchbook	**.42**	**1.16**	.03
ski-mask	**.**.40	6.04	**.53**

She gave him a ——— to play with.			
plastic spoon	.48	5.58	.07
large refrigerator	**.49**	1.08	**-.03**
red beanbag	.39	**6.29**	**.59**
Adam pulled out of his golf bag a _____ and used that to chisel an inch of ice off his windshield.			
seven iron	.50	4.50	.42
ham sandwich	.56	1.00	.16
screw driver	**.60**	**5.00**	**.63**

List of Figures

Figure 1. Dependency tree for the sentence "*Rolls-Royce said it expects its US sales to remain steady at about 1,200 cars*" (after Yamada & Matsumoto, 2003).

Figure 2. Propositional analysis of the sentence in Figure 1(after Kintsch, 1998).

Figure 3. Propositional structure mapped on the dependency tree.

Figure 4. A spreading activation network showing the 8 most strongly activated neighbors of SHARK by LAWYER (ICA, TASA corpus).

Figure 5. Dependency tree for the sentence "*If you try, you can solve any_____*" indicating that the target word must be part of the [SOLVE-ANY-____] proposition.

Figure 6. Dependency tree for the sentence "*Oracle fought to keep the forms from being released*".

Figure 7. Dependency trees for the inference sentences "*Oracle released a confidential document*" and "*Oracle fought to keep the forms released.*"

Figure 8. Dependency tree for the sentence "*Being clever, he walked into a store and bought a **newspaper** to cover his face.*"

Appendix: Discussion

Glenberg: I agree with you that there's an awful lot of structure that's available in covariance and analyzed covariance as you presented it. But, in the spirit of this workshop, what I'd like to hear from you is how we might determine whether people are using those sorts of representations; whether people are extracting information from language in similar ways.

KINTSCH: I make predictions about what people do, I observe the behavior and see how well the predictions fit. That's all I can say.

Glenberg: And I agree with you, that the predictions do fit nicely. I think I need to formulate my question more incisively; and the question really should be, how can we tell if this is a better theory than perhaps the embodied theory? So, where are the predictions that are different from an embodied theory?

KINTSCH: Well, I don't think that that's so much the question - what's a better theory. I think that what we talked about yesterday was, and most people are agreed that, cognition has symbolic aspects and has embodied aspects, and that's worthwhile to study, and it's not an either/or kind of thing. This is clearly what I wanted to show.

Pulvermüller: Thanks very much for an excellent and very interesting talk. it is a very good thing to have these measures of relatedness of semantic distance between words and concepts, or family resemblance. However, the information that would not be included in this relationship, distance measures, would be for example, that *bark* is in one case an action and in the other case an object. And one wouldn't have directions from this very abstract representation about in which brain areas one would have to look. Of course one could take a very abstract strategy and just say, I take this description and make a correlation with brain activation. But on the other hand it might be a good thing to connect these descriptions, these semantic descriptions with some information that links

the representation to a particular cognitive domain, domains and brain systems as well. A priori. Obviously, an embodied description might have this advantage, to provide that. These measures we are discussing, these covariance measures, they would also incorporate important information which a purely embodied theory might lack, so there might be mutual, well, both sides might learn from each other.

KINTSCH: Well, I think that suggestion is a very interesting one, and that's something we can actually look into. Because, if I have a lot of nouns and a lot of verbs, I can get them to cluster in very distinct clusters, so the verbs are here, and the nouns are over there. Now if you have a word like *bark*, which could be either one, I don't know what would happen. But if I can construct, now, the dog meaning of *bark*, which is a verb, and the tree meaning of *bark*, which is a noun, those two vectors should be in different clusters.

Pulvermüller: Sure, yes.

KINTSCH: If I'm right. And it might also be interesting from a brain standpoint to see what happens with homonyms that have two different functions, how they are represented in two different places.

Pulvermüller: Right. Two overlapping styles is what we would propose, or have proposed. May I make another little comment on the syntax part. I've tried to argue that, using a model consisting of sequence detectors, it is actually possible to formulate dependency grammar, so I have a feeling that if you have both the basic sequencing of words, and on top of that a separate grammar dependency, grammar processor, that there might be some redundancy.

KINTSCH: I know, I was thinking yesterday when Andreas was speaking about all the work on sequencing, whether that couldn't be put together. A very interesting thought.

Pulvermüller: Yes, thanks.

Figure 1. Dependency tree of the sentence "*Rolls-Royce said it expects its US sales to remain steady at about 1,200 cars*" (after Yamada & Matsumoto, 2003).
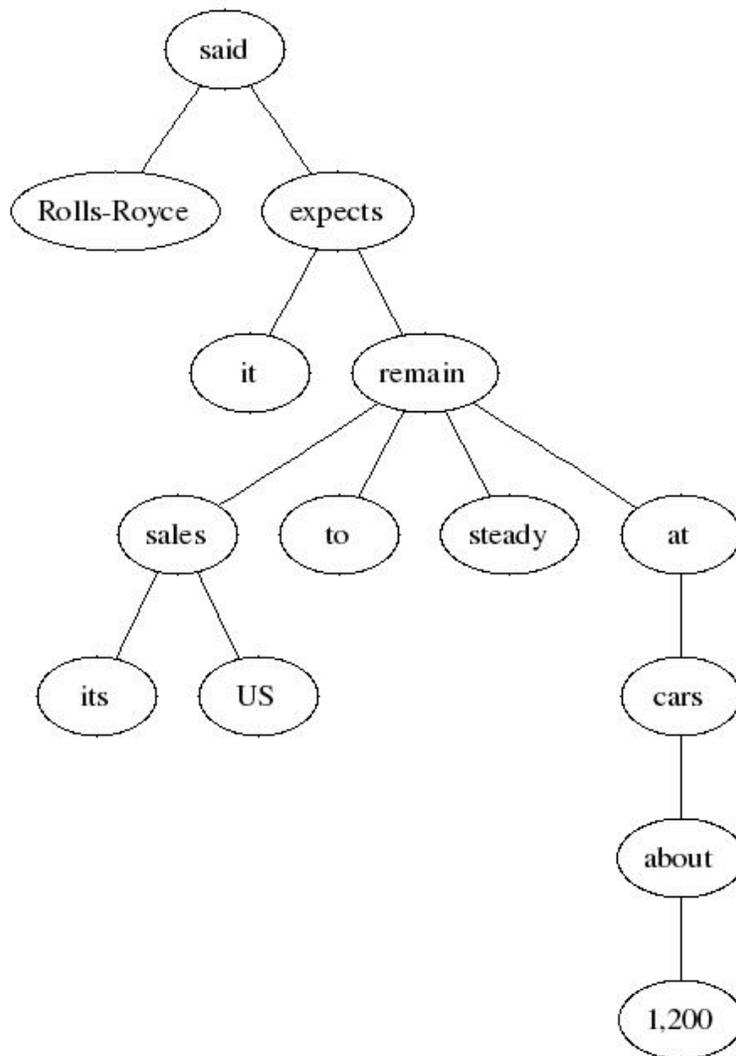
Figure 2. Propositional analysis of the sentence in Figure 1 (after Kintsch, 1998).

    (1)   SAID [ROLLS-ROYCE, 2]
    (2)   EXPECT [ROLLS-ROYCE, 3]
    (3)   REMAIN [SALES, STEADY]
    (4)   OF [ROLLS-ROYCE, SALES]
    (5)   US [SALES]
    (6)   AT [STEADY, 8]
    (7)   1,200 [CARS]
    (8)   ABOUT [7]

Figure 3. Propositional structure mapped on the dependency tree.
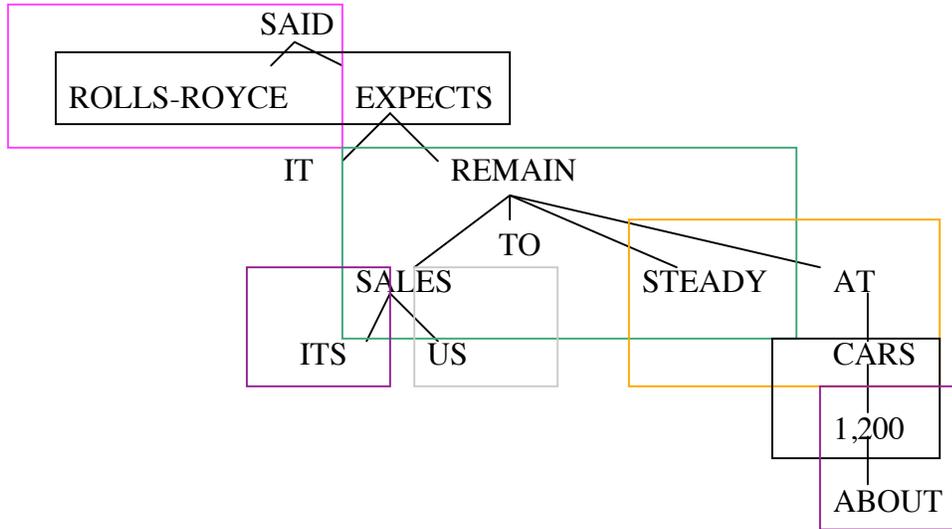
SAID
ROLLS-ROYCE   EXPECTS
IT   REMAIN
TO
SALES   STEADY   AT
ITS   US   CARS
1,200
ABOUT

Figure 4. A spreading activation network showing the 8 most strongly activated neighbors of SHARK by LAWYER (ICA, TASA corpus).
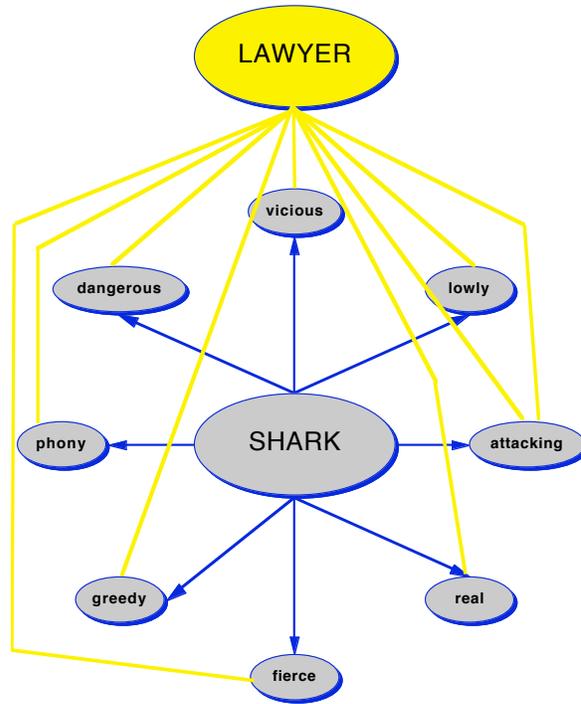
Figure 5. Dependency tree for the sentence "*If you try, you can solve any problem*"
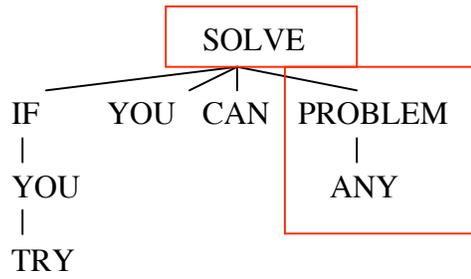indicating that PROBLEM must be part of the [SOLVE-ANY-_____] proposition.

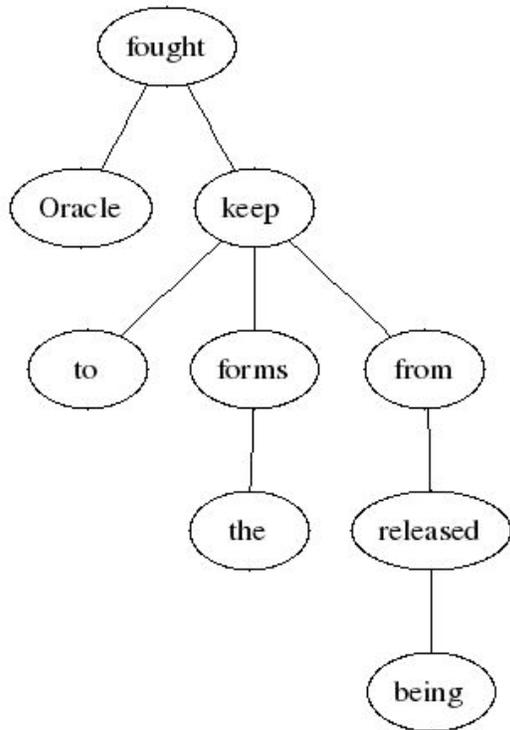Figure 6. Dependency tree for the sentence "*Oracle fought to keep the forms from being released*".

Figure 7. Dependency trees for the inference sentences "*Oracle released a confidential document*" and "*Oracle fought to keep the forms released*"
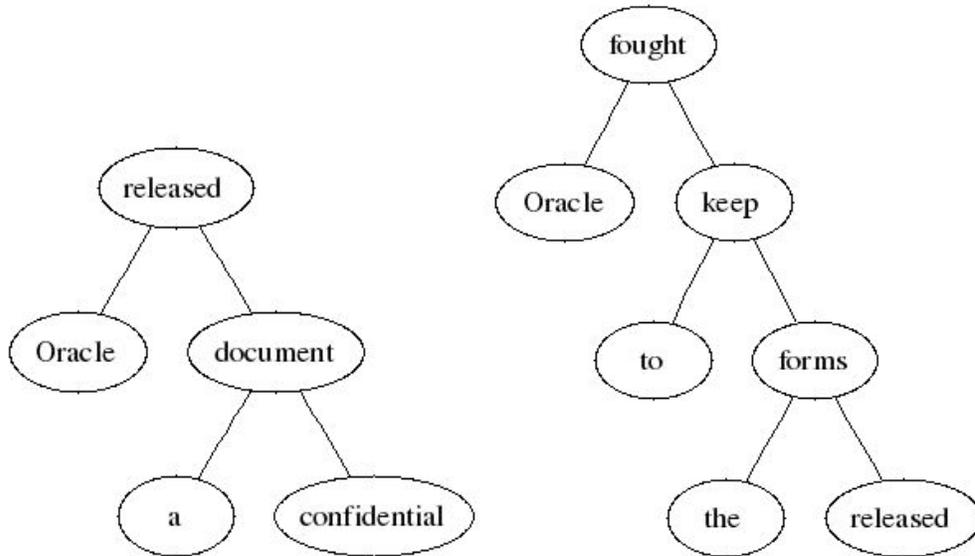
Figure 8. Dependency parse for the sentence "*Being clever, he walked into a store and bought a **newspaper** to cover his face* "