Part IV: Critique

How cognitive is cognitive poetics? The interaction between symbolic and embodied cognition

Max Louwerse and Willie Van Peer

1. Introduction

The field of literary studies has long suffered from a methodological identity crisis. For decades strongly motivated researchers have been fighting an uphill battle to make the study of literature empirical to get the field out of this crisis. Societies like the Association for Empirical Studies of Literature and Media (IGEL), the Poetics and Linguistics Association (PALA), and the International Association for Empirical Studies of the Arts (IAEA) have played an important role in this process. It is somewhat surprising that a methodological identity crisis had to be diagnosed and treated, since few other scientific disciplines have been questioning fundamental methodological issues. Difficult to imagine is that a discipline like Physics would doubt the value of empirical research. Equally difficult to imagine is that a discipline like Psychology would have any hesitations with empirical research. To most literary scholars, however, the empirical study of literature was (and still largely is) considered questionable at best. To a small elite group of researchers investigating literature empirically is not only self-evident but also unavoidable. How else could one get answers to questions like what writers do when they write, what readers do when they read, what the literary culture does when it reacts to new developments, and what the characteristics are of a literary text. Instead, unobservable magic was supposed to be part of the "scholarly" process.

It could be countered that the kind of objects studied in literary studies, or the kind of mental activities involved in processing them demand a degree of subjectivity on the part of the scholar that is not easily amenable to empirical methods of study. It is then argued that what readers engage in when they encounter literary texts are processes very much tied to individual and social norms and values, to identification or to the interrogation of social practices or the ideological bases upon which they rest. And indeed this kind of interminable self-reflexivity is the daily practice one can observe in most programs of literary studies. Many people would argue that under such conditions it is not easy to employ empirical methodology, because the processes going on are invisible to the eye, they are subjective in nature, and depending on value orientations. We do not deny such characteristics of these phenomena, but we do disagree that they in any way prohibit the use of empirical methods. To begin with, traditional literary scholars often confound difficulty with impossibility. True, it is not always easy to come up with strict operationalizations of such mental activities that go on during reading. But neither is it easy to measure the speed of light, or to reconstruct the history of the earth. "Not easy" is not the same as "not possible". It seems to us that most literary scholars have given up the idea of empirical research *even before* they have given it a try. In that way, they are involved in a self-fulfilling prophecy: because scholars declare the task to be nigh impossible, it is never tried, and this is subsequently used as evidence against any future effort to do so.

Fortunately, literary studies does, however, sometimes show signs that it wishes to liberate itself from its methodological identity crisis. Cognitive poetics has played an important role in this process, either as a facilitator or as the end product. Volumes like Semino and Culpeper (2002), Steen and Gavins (2003), and Stockwell (2002), as well as the current volume serve as (empirical) evidence that "cognitive" is the way to go in literary matters. As becomes clear in these volumes, cognitive poetics applies the principles of cognitive science to the interpretation of literary texts. Cognitive science, which is the scientific study of mind and intelligence, is a highly interdisciplinary enterprise, as it incorporates the fields of psychology, linguistics, anthropology, education, neuroscience, and computer science. It may be obvious to the reader of this volume what cognitive poetics entails, what cognitive science entails, and how these can go together. But as we will show, there are some basic issues in serious need of further discussion.

Gavins and Steen (2003: 2) describe cognitive poetics as follows:

> [Cognitive poetics] suggests that readings may be explained with reference to general human principles of linguistic and cognitive processing, which ties the study of literature in with linguistics, psychology, and cognitive science in general. Indeed, one of the most exciting results of the rise of cognitive poetics is an increased awareness in the social sciences of the special and specific nature of literature as a form of cognition and communication. What is noted at the same time, however, is that this special position of literature is grounded in some of the most fundamental and general structures and processes of human cognition and experience, enabling us to interact in these special and artistic ways in the first place.

It could be debated (although the issue is not central to our argument) that the social sciences have become increasingly aware of the special and specific nature of literature as a form of cognition and communication, thanks to cognitive poetics. For instance, literary discourse and figurative language have repeatedly shown up in handbooks on cognitive psychology such as, for instance, Gernsbacher (1994), Graesser, Gernsbacher, and Goldman (2003), Traxler and Gernsbacher (2006), Van Dijk and Kintsch (1983), Louwerse and Van Peer (2002) and Wilson and Keil (1999). This may suggest then that cognitive poetics has not changed the social sciences, but that cognitive poetics borrows theories and principles from the cognitive sciences. That would mean that the unidirectional relation is opposite to what is stated in the description quoted above. This is an important observation, because of the second and more pertinent issue we would like to raise, that of what is *not* said in the above description. Cognitive poetics does not just apply principles found in the cognitive sciences. It has, on the contrary, a built-in bias because it carefully selects what to borrow (and what not). For instance, it tends to not borrow principles from computer science and computational linguistics, despite the appropriateness of these principles for the study of literary text and discourse. Instead, cognitive poetics relies for its concepts and methods heavily (if not almost exclusively) on cognitive linguistics. Clearly, a field cannot borrow from each and every area of an interdisciplinary conglomerate like cognitive science. But by not choosing computational linguistics and by emphasizing cognitive linguistics, it has made an essential decision with regards to cognition and language (and their interaction). It has come to assume that language comprehension is strictly embodied, as is held in cognitive linguistics. Understanding the words in a literary text has to involve activation of embodied experiences we have with these words. Against this view we argue that this embodiment bias does neither justice to the topic of investigation (literature) nor to the field from which cognitive poetics borrows (cognitive science).

To gain understanding of the embodiment bias, we need to discuss some recent developments in the cognitive sciences with regards to the nature of language comprehension, which show language comprehension as both symbolic and embodied. This is followed by computational analyses of examples from cognitive poetics that were considered embodied. More specifically, examples will be taken from chapters in Stockwell (2002) to show that symbolic approaches can also capture aspects of meaning.

2. Symbolic and embodied aspects of language comprehension

The cognitive revolution in the 1950s, a response to behaviorism that ruled psychology since the early 20th century, started the cognitive sciences. Whereas behaviorism emphasized the study of observable behavioral processes and dismissed the study of inward mental processess, the cognitive sciences emphasized the importance of human mental processes. The invention of the computer undoubtedly played a crucial role in understanding these processes and provided researchers with a tool to model them. These days marked the start of artificial intelligence and computer science. These fields consider human thinking being not much different than computational thinking. Indeed, computational models allowed researchers to test hypotheses even without human experiments.

Probably as a reaction against these symbolic approaches to language comprehension, researchers in the 1980s started to conjecture that symbolic processes alone could not explain language comprehension, because symbols are not grounded in bodily experiences. Thought experiments related to this symbol grounding problem illustrated the limitations of symbolic processes (Harnad 1990; Searle 1980). For instance, imagine attempting to read a literary text in an unknown language. Somebody hands you a dictionary that allows you to translate the words on the page into that of another foreign language. A dictionary of that foreign language translating the words into yet another language will not help much either. The reason is that the symbols remain ungrounded. This thought experiment shows that symbols remain meaningless, like the symbols in a computer, until the symbols get grounded or embodied into the physical world. A word like “spoon” means an eating or cooking utensil with a shallow bowl attached to a relatively long handle only because we have physical experiences with spoons in our world. That is, we can pick up spoons, we can throw them, can bend them, can use them as a miniature mirror and can even use them to eat cereal or soup. This embodied approach to the meaning of “spoon” is different from a symbolic approach. In the latter approach the meaning of spoon comes about through the interrelations of the word “spoon” with other words. For instance, we know what a spoon is because the linguistic context of the word with the words “fork” and “knife,” or with a word like “eating”.

Two competing approaches to language comprehension can be distinguished in the 1990s: a symbolic approach emphasizing the computational nature of symbols and an embodied approach emphasizing the grounding of symbols in the physical world. The symbolic approach gained impetus with computational models like Latent Semantic Analysis (LSA; Landauer and Dumais 1997) and Hyperspace Analogue to Language (HAL; Lund and Burgess 1997). Take LSA (http://lsa.colorado.edu), for instance. Meaning is captured by mapping initially meaningless words into a continuous high dimensional semantic space, which more or less simulates cognition (Landauer 2002). More specifically, a first-order process associates stimuli (words) and the contexts they occur in (documents). Based on their contiguity or co-occurrence, stimuli are paired. These local associations are next transformed by means of Singular Value Decomposition (SVD) into a small number of dimensions (typically 300) yielding more unified knowledge representations by removing noise. Like language comprehension, memory for the initial local associations (surface structure) becomes memory for more global representations (the central meaning). LSA can thereby be seen as a theory of knowledge representation, induction and language acquisition (Landauer and Dumais 1997; Landauer 2002; Louwerse and Ventura, 2005). Let's illustrate this with the following three sentences.

(1) The dog barked against a tree in the park.
(2) The cat climbed into a tree in the park.
(3) The squirrel jumped from branch to branch

Based on first-level co-occurrences, "dog" and "cat" in the first two examples are semantically associated, because their contexts share the lexical items "tree" and "park". However, the semantic relatedness in LSA is not (only) determined by the relation between words, but also by the words that accompany a word. This means that a semantic association can be found between "cat" and "squirrel" even though they do not share any context. The fact that "branch" and "tree" may share a context in another document, however, or even the fact that the semantic neighbors of "branch" and "tree" share a semantic context (or the neighbors of the neighbors of the neighbors of the neighbors of "branch" and "tree"), allows for a semantic association. This means that words may never occur in the same document for LSA to still compute a semantic association.

The method of statistically representing knowledge has proven to be useful in a variety of studies. It has been used as an automated essay grader, comparing student essays with ideal essays (Landauer, Foltz, and Laham 1998) and performs as well as students on the TOEFL (Test of English as a Foreign Language) tests (Landauer and Dumais 1997). More recently, LSA has also been used for several other applications. First, it plays an important role in Coh-Metrix (Graesser, McNamara, Louwerse, and Cai, 2004; Louwerse, McCarthy, McNamara, and Graesser 2004), a web-based tool that analyzes texts on over 50 types of cohesion relations and over 200 measures of language, text, and readability. LSA measures the semantic relatedness between sentences, paragraphs and texts. LSA has also been used in intelligent tutoring systems like AutoTutor and iSTART. AutoTutor engages the student in a conversation on a particular topic like conceptual physics or computer literacy. AutoTutor uses LSA for its world knowledge and determines the semantic association between a student answer, and ideal good and bad answers (Graesser et al. 2004). iSTART uses LSA in its teaching of reading strategies to students by providing appropriate feedback to students' self-explanations (McNamara, Levinstein, and Boonthu 2004).

Obviously, LSA is not synonymous to a symbolic approach to language understanding. At the same time, it can be seen as a model (both theoretical and applied) of language comprehension. Moreover, it has been considered by many as the example model of symbolic language comprehension (Glenberg and Robertson 2000; Landauer and Dumais 1997). Whereas other corpus-based models of word meaning may provide similar results as LSA (Louwerse, Lewis and Wu, in press), LSA is insensitive to sparsity problems that are present in many other corpus linguistic approaches.

But many psychologists and cognitive scientists have argued that corpus-based models of word meaning can simply not be the whole story. For instance, embodied theorists (Barsalou 1999; Glenberg and Robertson 2000) claim that word meaning can never be fully identified by associative models using only amodal symbols. Without grounding the words to bodily actions in the environment we can never get past defining a symbol with another symbol. Indeed, a wealth of information shows language comprehension is fundamentally embodied. For instance, comprehenders' motor movements match those described in the linguistic input. Klatzky, Pellegrino, McCloskey, and Doherty (1989) showed comprehension of verbally described actions (e.g. the phrase "picking up a grape") to be facilitated by preceding primes that specified the motor movement (e.g. grasp). Glenberg and Kaschak (2002) found similar evidence measuring how much the sensibility of a sentence is modified by physical actions. When subjects read sentences like "Mark gave you a pen" and used a congruent action (press a button close to the body of the subject), reaction times were lower than when an incongruent action (press button away from the body of the subject) was applied. Zwaan, Stanfield, and Yaxley

(2002) measured response times for pictures matching the content of sentences and pictures that did not. For instance, they used sentences about a nail being pounded either into the wall or into the floor. Response times for pictures matching the sentence (e.g. vertically oriented nail for sentence in which nail pounding into the floor) were faster than for mismatching pictures, leading to the conclusion that subjects simulated the scenes described in the sentence. Similarly, Zwaan, and Yaxley (2003) showed that spatial iconicity affects semantic judgments (the word "attic" presented above the word "basement" resulted in faster judgments than the reverse iconic relationship), suggesting that visual representations are activated during language comprehension.

3. Symbolic interdependency

With evidence concurrently supporting the symbolic approach and the embodied approach, how can language comprehension be explained? We have argued elsewhere that language comprehension is both embodied and symbolic (Louwerse 2007; Louwerse [under review]; Louwerse and Jeuniaux [in press]; Louwerse, Cai, Hu, Ventura, and Jeuniaux 2006). We proposed a Symbol Interdependency Hypothesis based on Deacon's (1997) hierarchy of signs, which is in turn based on Peirce's (1923) theories. Peirce identifies three types of signs: icons, indices and symbols. In icons a direct relation between the sign and what the sign represents can be observed. For instance, a picture of one of the editors of this book represents the editor of this book. There are physical similarities between picture and person. In indices that relation is not as direct. Instead, indexical relations are situated in space and time. An example is the fingerprint of one of the editors. The fingerprint itself does not represent the editor, but represents the presence of that editor at some point in time at some particular place. When time goes by, the interpretative link between presence of editor and his footprint tends to weaken. Finally, there are symbols. In symbols the relationship between the sign and what the sign represents is determined by convention. The wedding ring of the editor represents that he is married. But different cultures could have a different symbol for marriage, as long as the cultural community agrees on the symbol. Language is also an example of symbols. Deacon (1997), based on Peirce (1923), argues that icons, indices, and symbols have a hierarchical relationship with each other, whereby indices are built from combinations of icons, and symbols are built from combinations of indices. Moreover, relations between these signs can operate at one level (symbols being related to other symbols) and at different levels (symbols being related to indices and icons). Deacon claims that this hierarchy of different levels of signs can help us explain why humans have language, but other species do not. Humans are symbolic species – they can make links between symbols and between symbols and indices and icons – whereas other species cannot build the bridge between indices and symbols (although higher species like chimps can get very close).

According to the symbol interdependency hypothesis, language comprehension can be symbolic through interdependencies of symbols, but it can also be embodied through the dependencies of symbols on indices and icons. It thereby makes two important predictions. First, language comprehension is typically symbolic, although there are conditions under which embodied representations are activated, as is the case in deep processing or when comprehenders are cued to activate other modalities. This does not mean language comprehension is solely symbolic, because comprehenders can always activate embodied representations (indices and icons). It merely means comprehenders can often rely on the symbols in language to bootstrap meaning. Second, because language has evolved to become a communicative short-cut, language structures now represent relations in the physical world: language has encoded embodied relations. As examples of these encoded structures, Louwerse (2007) lists examples like subjects always preceding objects (Greenberg 1963), and categories being determined by the way we perceive the structure of the world (Rosch 1978).

What does this exposé on language comprehension mean for cognitive poetics? Considering the reliance of cognitive poetics on theories in cognitive science and thereby on theories of language comprehension, one would expect the evidence found in favor of the symbol interdependency hypothesis to be extended to cognitive poetics. That is, the examples which cognitive poetics uses to show that the understanding of literary language is embodied, can also be applied in a computational approach and be given alternative explanations. In the remainder of this chapter we will follow Stockwell's (2002) *Cognitive poetics. An introduction* and take examples from the first chapters of this book, "Figures and grounds," "Prototypes and reading," "Cognitive deixis" and "Conceptual metaphors". We will apply a symbolic approach to complement his embodied analyses.

bro_002.pod 431
07-11-10 06:57:28 -mt- mt

4. Figure and ground

Figure and ground are central concepts in cognitive poetics (Stockwell 2002) taken from early 20th century Gestalt Psychology. Similar elements are grouped together (figure) and contrasted with dissimilar elements (ground) in order to perceive an image coherently. For instance, supporters distinguish soccer teams on the basis of the colors of their shirt, grouping the same color shirts and contrasting them with different-color shirts. In literary studies figure and ground are used in the notion of foregrounding, the process by which something is given prominence by the reader of a text (Van Peer 1986). Foregrounding allows the author to defamiliarize the reader by foregrounding certain aspects of the text. Foregrounded stylistic traits in the text give prominence to the figure, differentiating it from the ground (e.g. the prominence of the main character Hamlet in the title of the play). Stockwell argues that readers activate image schemas, i.e. mental pictures that readers use as basic templates to understand common situations. These image schemas are supposed to be embodied and consist of a trajector (figure) that has a grounded relationship with a landmark (ground) through a path. Note that this may be seen as an improvement over traditional methods of text analysis, as it allows a clearer formulation of hypotheses. As an example, Stockwell uses the following lines from Shakespeare's *Midsummer Night's Dream*:

> Over hill, over dale,
> through bush, through briar,
> Over park, over pale,
> through flood, through fire,
> I do wander everywhere

Stockwell argues that readers interpret these lines by outlining the "I" in this fragment as the trajector/figure who takes a path flying above the landmark/ground (hill, dale, park, pale). Two questions arise, however. First, what is the evidence that readers activate these image schemas in examples like the one above? Secondly, the claim is that these image schemas are embodied. What is the cognitive evidence for this claim? Moreover, does the claim entail that non-embodied approaches will not be able to grasp the gist of the passage? Does the linking of the last line (trajector/figure) to the remainder of the passage (landmark/ground) help to form the backdrop of this summary? Does this really require an embodied representation?

Table 1. LSA cosine matrix of Shakespeare lines

	1.	2.	3.	4.	5.
1. over hill, over dale	1	0.2	0.87	0.12	0.19
2. through bush, through briar	0.2	1	0.24	0.66	0.2
3. over park, over pale	0.87	0.24	1	0.16	0.2
4. through flood, through fire	0.12	0.66	0.16	1	0.13
5. I do wander everywhere	0.19	0.2	0.2	0.13	1

To answer these questions, we entered the five lines of the Shakespeare text in an LSA matrix comparison using the literature-with-idioms LSA space (528 factors). This space consists of English and American Literature from the 18th and 19th century from the Project Gutenberg page. The space is composed of 104,852 word types, 57,092,140 word tokens and 942,425 paragraphs, with 338 dimensions. Cosine values were computed between each of the five lines, resulting in a 5x5 matrix, presented in table 1.

Following the method used in Louwerse et al. (2006) and Louwerse (2007) the LSA matrix of cosine values was next supplied to an ALSCAL algorithm to derive a Multidimensional Scaling (MDS) representation of the stimuli (Kruskal and Wish 1978). That is, the matrix of LSA cosine values was transformed into a matrix of Euclidean distances and these distances were scaled multidimensionally by comparing them with arbitrary coordinates in an *n*-dimensional space (low cosine values correlate with large distances, high values with short distances). The coordinates were iteratively adjusted such that Kruskal's stress is minimized and the degree of correspondence maximized. The fitting of the data was good (Kruskal's stress =.067, $R^2 = .993$) with a two-dimensional scaling. The graphical representation of the two-dimensional output is presented in figure 1. What becomes apparent in this figure is how the line considered to be the trajector/figure ("I do wander everywhere") is differentiated from the other lines, considered to be the landmark/ground.

Thus, the results Stockwell obtained from his analysis using an embodied approach is much the same as those obtained from an analysis using a symbolic approach. The fact that 18th and 19th century texts were used for the LSA space will not affect the results. First, the strength of LSA lies in the higher-order dependencies, whereby results are not dependent on individual words but on higher-order relations between the co-occur-

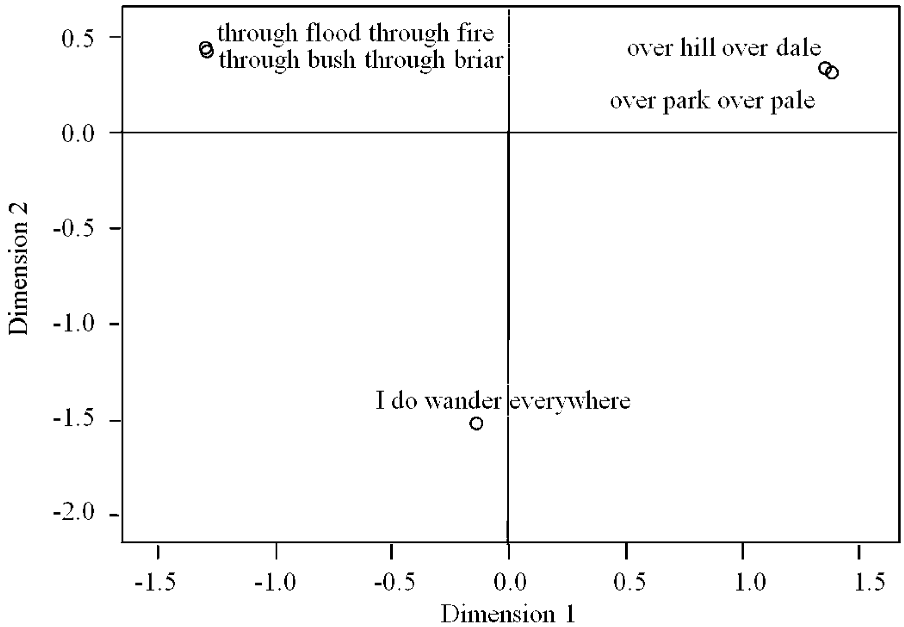

Figure 1. MDS representation of LSA cosine matrix of Shakespeare lines

rences (of the co-occurrences, etc.) of words. Secondly, the comparison presented here is between Shakespeare texts only. Even with a less than ideal LSA space, the comparison between a non-manipulated Shakespeare text and a manipulated Shakespeare text stays.

Admittedly, the text sample is short, and may have been more convincing (still) if we had used a longer extract from the play. However, let's continue this foregrounding example by butchering the Shakespearean lines into something that looks less aesthetically appealing like the following lines.

I wander over hill and dale
and through bush and briar
I wander over park and pale
and through flood and fire
I wander over grass and trees

The last line does not seem to represent the trajector/figure anymore. When the same method is used and the LSA cosine matrix (Table 2) is applied to an MDS ALSCAL algorithm, a representation emerges that is

Table 2. LSA cosine matrix of butchered Shakespeare lines

	1.	2.	3.	4.	5.
1. I wander over hill and dale	1	0.11	0.87	0.05	0.7
2. and through bush and briar	0.11	1	0.14	0.65	0.16
3. I wander over park and pale	0.87	0.14	1	0.09	0.75
4. and through flood and fire	0.05	0.65	0.09	1	0.06
5. I wander over grass and trees	0.7	0.16	0.75	0.06	1

presented in figure 2. The prominence of the trajector/figure that is missing from the butchered passage does not show up in the MDS analysis.

We do not argue that foregrounding can be explained simply by dumping lines of text in a computer. We do claim, however, that it is wise not to put all one's eggs in the embodiment basket, and instead also consider alternative approaches, particularly if these approaches are complementary.

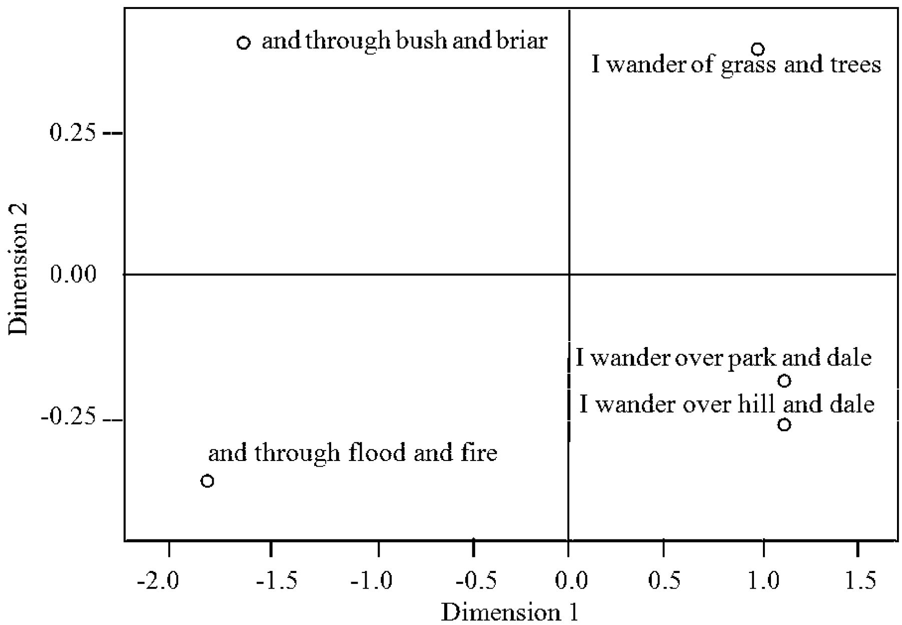

Figure 2. MDS representation of LSA cosine matrix of butchered Shakespeare lines

5. Categorization and prototypes

In the following chapter Stockwell (2002) shows how categorization and prototypes, key concepts in cognitive science, are important in literary language. How do we know that an "eagle" is a good member of the family "bird," but "ostrich" and "penguin" are not? Stockwell argues that we know this because of our interaction with the environment, our embodied experiences. The question is whether a symbolic approach would yield similar results.

Louwerse et al. (2006) conducted a number of studies whereby they used LSA to categorize concepts. They found that LSA was able to represent words used by Collins and Quillian (1969) into a hierarchical framework, clustering words like "wings," "feathers" and "fly" with the word "bird," but "fin," "gills" and "swim" with the word "fish". The clustering analysis also showed how "animals" "eat," have "skin," "move," "bite," are "dangerous" and "edible". Similarly, Louwerse et al. (2006) conducted a number of studies investigating category membership using Rosch's (1973) data. In Rosch's study participants rated the typicality of members of a category on a scale, showing that participants were consistent in their ratings (e.g. "robin" is a more typical member of the category "bird" than "chicken" is). Furthermore, participants were faster in judging whether a picture belonged to a certain category when the picture showed a typical member than when it was a non-typical member (Rosch 1975). Eight categories (fruit, science, sport, bird, vehicle, crime, disease and vegetable) with six members in each category were taken from Rosch (1973) (see also Akmajian, Demers, Farmer, and Harnish 2001). LSA results showed a significant correlation between experimentally obtained rank of category and LSA cosine value. This correlation per category showed that five out of the eight categories (bird, crime, fruit, sport, vegetable) had significant correlations. The remaining three categories (disease, science, vegetable) did have the expected pattern, though these did not receive the significance level, most likely due to the small number of cases (6 per category).

The categorization found in results from human experiments can also be found in the results from computational analyses, showing that categorization is not solely based on embodied factors, as Stockwell argues. In addition to referring to Rosch's studies on categorization, Stockwell argues that categorization also applies to literature, for instance in the way we categorize genres. We can take this a step further and determine how literary authors can be classified. Take for example "Dante," "Dickens,"

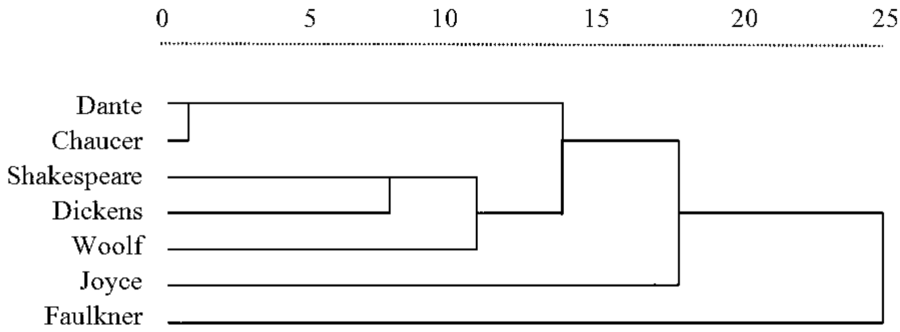

Figure 3. Hierarchical cluster analysis of LSA cosine matrix of literary authors

"Faulkner," "Joyce," "Shakespeare," and "Woolf". These words were entered in an LSA analysis using the Touchstone Applied Science Associates (TASA) semantic space. The TASA corpus consists of approximately 10 million words of unmarked high-school level English texts on Language Arts, Health, Home Economics, Industrial Arts, Science, Social Studies, and Business. This corpus is divided into 37,600 paragraphs, (averaging 166 words per paragraph) and is considered one of the benchmark corpora in computational linguistics, because it approximates the language familiarity of a college level student (Kintsch 1998; Landauer and Dumais 1997). We did not use the literature space from the previous study, because we are not interested here in literature per se, but in texts about literature. The LSA cosine matrix was next supplied to a hierarchical cluster analysis, resulting in the clustering presented in figure 3.

A simple semantic analysis of these words results in a cluster suggesting a close relation between Dante and Chaucer, as well as Shakespeare and Dickens. Woolf is closest to these latter two, and Joyce and Faulkner, particularly the latter, is furthest away from the other groupings. The higher-order relationships alone between these words can inform us how they can be clustered.

Again, the argument we are trying to make is not that any classification of literary authors can simply be done by computers. Leaving the question aside how literary authors could be grouped through an embodied approach, the argument we do want to make is that alternative approaches other than embodiment should be considered.

6. Cognitive deixis

In the next chapter, Stockwell discusses cognitive deixis, the language feature anchoring meaning to a context. Examples of deictic terms are personal pronouns like "I," "you" and "we," and adverbials like "here" and "there" for which the meaning crucially depends on the point of view of the speaker; and on the concrete spatial and temporal context and situation in which these words are used. For instance, "I" has a different referent for me than for you (let alone the referents of the other personal pronouns expressed in this sentence). Stockwell argues deixis as being central to embodiment. To some extent it is true that deixis is problematic for a symbolic approach, because it is difficult to capture the meaning of deictic items that are by definition pointing to referents in the physical world. The question, of course, is whether this is a weakness of the symbolic approach per se or if this is due to the limited information LSA is given access to. To determine whether LSA is able to capture meaning in deictic items, the best we can do is compute whether it is able to group these deictic items on the basis of their textual information alone. We use personal pronouns ("I," "you," "we"), possessive pronouns ("mine," "yours," "ours") and possessive adjectives ("my," "your," "our") to test this. Our prediction is that LSA is able to cluster singular items separately from plural items (e.g. "I" vs. "we"), thereby clustering person (e.g. "I," "my," "mine" vs. "we," "ours," "ours").

Two things are worth mentioning. First, based on the way LSA works, we have no reason to believe it will be able to cluster these items, since the higher-order relationships for words co-occurring for one (e.g. "I") are likely to be identical to the contexts of the other (e.g. "you"). Secondly, and more importantly, if a symbolic approach is unable to cluster deictic items, embodiment processes must always be active in language processing or at least in processing deixis. On the other hand, if a symbolic approach is able to cluster these items based on their textual occurrence, the argument made by the Symbol Interdependency Hypothesis can be extended: some of the meanings can be derived from the text, but ultimately we can ground language to referents in the physical world. As before, an LSA matrix was computed, this time using the 18 personal pronouns, possessive pronouns and possessive adjectives, and applied to an MDS ALSCAL algorithm. Results are presented in figure 4.

It is obvious from this Figure that singular items are clustered separately from plural items. The pronoun "it" and the possessive adjectives form an exception, possibly because of the ambiguous status of the word

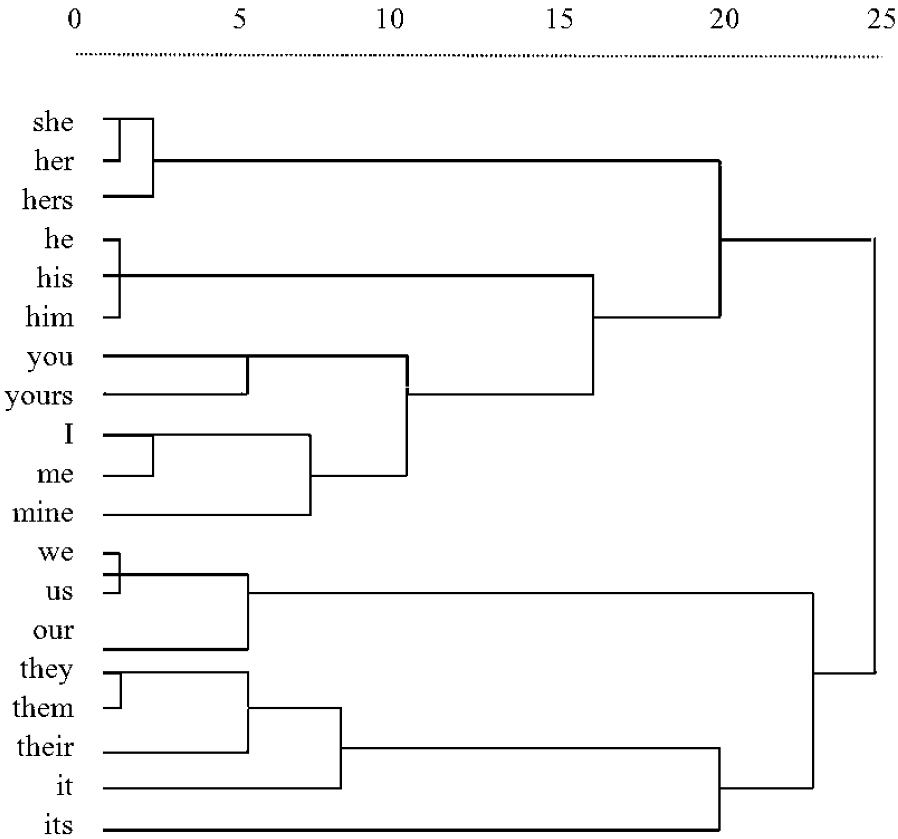

Figure 4. Hierarchical cluster analysis of LSA cosine matrix of personal and possessive pronouns and possessive adjectives.

"it" in its referring, anticipatory, cleft, and prop uses. For the singular cases, third-person items are clustered and first- and second-person items are clustered.

The analysis here does not demonstrate that deictic words are not embodied. It merely aims to show how deictic items cluster in intuitively accurate ways. Cognitive deixis can therefore be explained from an embodiment but also from a symbolic point of view.

7. Conceptual metaphor

This paper cannot be ended without addressing conceptual metaphors, since they form a central topic of interest in cognitive linguistics (Lakoff, 1987; Lakoff & Johnson 1980, 1999). Conceptual metaphors are formed by the blending of two conceptual domains, a source domain and a target domain. The source domain is the conceptual domain from which metaphorical expressions are drawn, whereas the target domain is the domain we try to understand. Take, for instance, the metaphor "my lawyer is a shark." My lawyer is the "vehicle" or "target," the new element to be described and "shark" is the "tenor" or "source" of the familiar element. Central to the cognitive linguistic approach to conceptual metaphors is that metaphors are not a matter of language but a matter of thought (Lakoff & Johnson 1980, 1999). Moreover, researchers in cognitive linguistics, as well as those in cognitive poetics, argue that conceptual metaphors are embodied. According to them, understanding conceptual metaphors requires linking target and source, the latter more than the former, to embodied experiences. For instance, up-and-down serves as a source domain for a variety of targets, as in "his spirits rose / his spirits sank". If conceptual metaphors are strictly embodied, there is little hope for a symbolic approach explaining metaphors.

Kintsch (2000), however, proposed a predication model that extends LSA and can approximate the meaning of metaphors by constraining the meaning of words that are compared. Consider, for instance, the sentence "that man is a shark". Using Kintsch's proposal, semantic associates are computed such as "fish," "jaws," "dangerous" and "fin". These words are then compared to "man" whereby only those words associated with "man" are kept (in the above example the semantic associate "dangerous"). The elegance of the predication model is that it can distinguish between "that man is a shark" and "that shark is a man".

For illustrative purposes here, we ignore this order effect and take the first three metaphors presented in Stockwell's (2002) conceptual metaphor chapter and compute those words that form the nearest semantic neighbors of these metaphors. That is, we are asking LSA to pick up the word that is semantically closest to the target word. Since we are looking for adjectives (e.g. that man is a shark – that man is dangerous), we are selecting the first adjective that is returned by LSA. Results of the metaphor and their nearest neighbor are presented in table 3. However, the argument can of course be made that these adjectives are simply the closest words to either one of the two words (target or source). To rule out this

Table 3. Metaphors/target/source and their nearest semantic neighbors with the cosine value

Metaphor, target and source	Nearest adjective neighbor	Cosine value
the man is a shark	dangerous	.56
man	honest	.41
shark	amphibious	.41
that man is a wolf	dangerous	.58
man	honest	.41
wolf	brute	.46
Juliet is the sun	glorious	.54
Juliet	Artistic	.48
sun	golden	.51

explanation, Table 3 also lists the nearest neighbor for target and source separately, showing that the semantic nearest neighbor for the metaphor (target *and* source) is different than that for either target or source.

These simple examples show the use of LSA as a tool to approximate the meaning of conceptual metaphors, and that it is not only an embodiment approach that can do this. Even more recent developments (e.g. Fauconnier and Turner 2002) do not annul the results obtained through our LSA showing that a symbolic approach can account for the data just as well. We do not argue that LSA can explain all metaphors. Nor do we propose that one just needs to enter metaphors in it and out comes the meaning. Instead, we claim that symbolic and embodied approaches are complementary. Apparently, cognitive poetics has not made a careful analysis of symbolic approaches, but simply latched on to the (embodied) assumptions of cognitive linguistics, ignoring complementary approaches.

8. Conclusion

We started this chapter by welcoming the empirical approach in cognitive poetics, but observing a certain fundamental bias based on theories of language comprehension in support of strict embodiment. We countered this bias by showing how language comprehension is both symbolic and embodied: embodied representations do not always have to be activated, and language has encoded many embodied relations. To show how a symbolic approach can be applied in a field like cognitive poetics, we took examples from Stockwell (2002) and analyzed them, using LSA. It needs to be emphasized here that we chose Stockwell's text with respect, as it is considered a milestone in the field, as one of us has stated in writing. We selected four chapters (figure and ground, prototypes, cognitive deixis, conceptual metaphor) and illustrated how LSA analyses can shed light on the processes of meaning construction just as well as embodiment theory does. We thereby believe we have shown that a symbolic approach is at least worth considering in interpreting language and literature. Adding a symbolic approach augments the theoretical and methodological validity of cognitive poetics.

Acknowledgments

This research was supported by grant NSF-IIS-0416128 Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding institution. Correspondence concerning this article should be addressed to Max M. Louwerse, Institute for Intelligent Systems / Department of Psychology, University of Memphis, 202 Psychology Building, Memphis, Tennessee 38152–3230.

References

Akmajian, Adrian, Richard A. Demers, Ann K. Farmer and Robert M. Harnish
1990 *Linguistics. An Introduction to Language and Communication*. Cambridge, MA: MIT Press.
Barsalou, Lawrence W.
1999 Perceptual symbol systems. *Behavior and Brain Sciences* 22: 577–660.
Collins, Allan. and Ross M. Quillian, M.
1969 Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8: 240–248.
Deacon, Terrence
1997 *The Symbolic Species: The Co-evolution of Language and the Human Brain*. London: Allen Lane.
Gavins, Joanna and Gerard Steen (eds.)
2003 *Cognitive Poetics in Practice*. London: Routledge.
Fauconnier, Gilles and Mark Turner
2002 *The Way we Think: Conceptual Blending and the Mind's Hidden Complexities.* New York: Basic Books.

Gernsbacher, Morton Ann
1994 *Handbook of Psycholinguistics*. San Diego, CA: Academic Press.
Glenberg, Arthur M.
1997 What memory is for. *Behavioral and Brain Sciences* 20: 1–55
Glenberg, Arthur M. and Michael Kaschak
2002 Grounding language in action. *Psychonomic Bulletin & Review* 9: 558–565.
Glenberg, Arthur M. and David A. Robertson
2000 Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language* 43: 379–401.
Graesser, Arthur C., Morton Ann Gernsbacher and Susan J. Goldman (eds.)
2002 *Handbook of Discourse Processes*. Mahwah, NJ: Erlbaum.
Graesser, Arthur C., Shulan Lu, George T. Jackson, Heather Hite Mitchell, Mathew Ventura, Andrew Olney and Max M. Louwerse
2004 AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments, and Computers* 36: 180–193.
Graesser, Arthur C., Danielle S. McNamara, Max M. Louwerse and Zhiqiang Cai
2004 Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers* 36: 193–202.
Greenberg, Joseph H.
1963 Some universals of grammar with particular reference to the order of meaningful elements. In: Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, MA: MIT Press.
Harnad, Stevan
1990 The symbol grounding problem. *Physica* D 42: 335–346.
Kintsch, Walter
1998 *Comprehension: A Paradigm for Cognition*. New York: Cambridge University Press.
Kintsch, Walter
2000 Metaphor comprehension: A computational theory. *Psychonomic Bulletin and Review* 7: 257–266.
Klatzky, Roberta L., James Pellegrino, Brian P. McCloskey and Sally Doherty
1989 Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory and Language* 28: 56–77.
Kruskal, Joseph B. and Myron Wish
1978 *Multidimensional Scaling*. Beverly Hills, CA: Sage University Series.
Lakoff, George
1987 *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
Lakoff, George and Mark Johnson
1980 *Metaphors We Live By*. Chicago: University of Chicago Press.
Lakoff, George and Mark Johnson
1999 *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.

Landauer, Thomas K.
2002 On the computational basis of learning and cognition: Arguments from LSA. In: Brian H. Ross (ed.), *The Psychology of Learning and Motivation* 41: 43–84. Amsterdam: Elsevier.

Landauer, Thomas K. and Susan T. Dumais
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, Thomas K., Peter W. Foltz and Darrell Laham
1998 An introduction to latent semantic analysis. *Discourse Processes* 25: 259–284.

Louwerse, Max M.
2007 Iconicity in amodal symbolic representations. In: Thomas Landauer, Danielle McNamara, Simon Dennis and Walter Kintsch (eds.), *Handbook of Latent Semantic Analysis*, 107–120. Mahwah, NJ: Erlbaum.

Louwerse, Max M.
In press Embodied representations are encoded in language. *Psychonomic Bulletin and Review.*

Louwerse, Max M., Zhiqiang Cai, Z., Xiangen Hu, Mathew Ventura and Patrick Jeuniaux
2006 Cognitively inspired natural-language based knowledge representations: Further explorations of Latent Semantic Analysis. *International Journal of Artificial Intelligence Tools* 15: 1021–1039.

Louwerse, Max M. and Patrick Jeuniaux
In press Language comprehension is both embodied *and* symbolic. In: Manuel de Vega, Arthur Glenberg, and Arthur C. Graesser (eds.), *Embodiment and Meaning: A Debate*. Oxford: Oxford University Press.

Louwerse, Max M., Gwyneth Lewis, and Jie Wu
In press Unigrams, bigrams and LSA. Corpus linguistic explorations of genres in Shakespeare's plays. In: Willie Van Peer and Jan Auracher (eds.), *New Directions in Literary Studies.* Cambridge: Cambridge Scholars Publishing.

Louwerse, Max M. and Willie Van Peer (eds.)
2002 *Thematics: Interdisciplinary Studies.* Amsterdam/Philadelphia: John Benjamins.

Louwerse, Max M. and Mathew Ventura
2005 How children learn the meaning of words and how LSA does it (too). *Journal of Learning Sciences* 14: 301–209.

Louwerse, Max M., Philip M. McCarthy, Danielle S. McNamara and Arthur C. Graesser
2004 Variation in language and cohesion across written and spoken registers. In: Kenneth Forbus, Dedre Gentner & Terry Regier (eds.), *Proceedings of theTwenty-Sixth Annual Conference of the Cognitive Science Society,* 843–848. Mahwah, NJ: Erlbaum.

Lund, Kevin and Curt Burgess
1996 Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instrumentation, and Computers* 28: 203–208.

McNamara, Danielle S., Irwin B. Levinstein and Chutima Boonthum
2004 iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, and Computers* 36: 222–233.

Peirce, Charles S.
1923 *The Collected Papers of Charles Sanders Peirce* (C. Hartshorne, P. Weiss, & A. Burks, eds.). Cambridge: Harvard University Press.

Rosch, Eleanor
1973 Cognitive representation of semantic categories. *Journal of Experimental Psychology: General* 104: 192–233.

Rosch, Eleanor
1978 Principles of categorization. In: Eleanor Rosch and Barbara B. Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale New York: Lawrence Erlbaum.

Searle, John R.
1980 Minds, brains, and programs. *Behavioral and Brain Sciences* 3: 417–57.

Semino, Elana and Jonathan Culpeper (eds.)
2002 *Cognitive Stylistics: Language and Cognition in Text Analysis.* Amsterdam/Philadelphia: John Benjamins.

Stockwell, Peter
2002 *Cognitive Poetics: An Introduction*. London: Routledge.

Traxler, Matthew J. and Morton Ann Gernsbacher (eds.).
2006 *Handbook of Psycholinguistics.* San Diego, CA: Elsevier.

Van Dijk, Teun A. and Walter Kintsch
1983 *Strategies of Discourse Comprehension*. New York: Academic Press.

Van Peer, Willie
1986 *Stylistics and Psychology; Investigations of Foregrounding.* London: Croom Helm.

Wilson, Robert A. and Keil, Frank C. (eds.)
1999 *The MIT Encyclopedia of Cognitive Sciences.* Cambridge, MA: MIT Press.

Zwaan, Rolf A. and Richard H. Yaxley, R.H.
2003 Spatial iconicity affects semantic-relatedness judgments. *Psychonomic Bulletin & Review* 10: 954–958.

Zwaan, Rolf A., Robert A. Stanfield, and Richard H. Yaxley
2002 Language comprehenders mentally represent the shapes of objects. *Psychological Science* 13: 168–171.