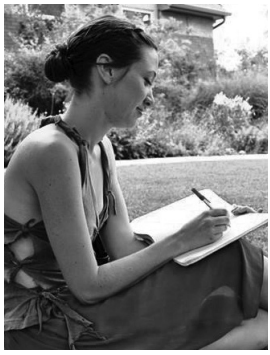
# Assessing the Debate between Embodied and Abstract Symbol Theories of Language Comprehension and Emotion

Celeste Hackenberg



Writer's Comment: *In an online interview with Chuck Smith and Sono Kuwayama, abstract artist Agnes Martin was asked how she felt about being a late bloomer. This is what she said: "If you have to do something new, it takes a long time to develop. If you just go on doing what's been done, you can start when you're twenty." I am decidedly a late bloomer as well, being that I am twenty-seven years old and finishing my undergraduate degree next year. Though I knew I wanted to be a writer at eighteen, I didn't know what job that was. Upon going to a career center, I was told, "Be a journalist. Here are all the different kinds," but this did not appeal to me. Though I had very limited experience of the world up until this point, I knew (or, more likely, felt) that there had to be something more I could do with my creative talent and interest in language, so I decided to feel things out for a while. I worked several different jobs and studied at four different colleges in both the United States and Canada, unknowingly creating patches that would eventually be sewn into a cohesive, if not uniform, quilt. Here at UC Davis, my final undergraduate destination, I am pursuing an honors degree in psychology with a self-created emphasis in psycholinguistics and poetry (unofficial), which I find to be the perfect combination of scientific and creative. Embodied cognition, the topic of the following paper, is not one normally explored by cognitive neuroscientists, but I think it is a very important one. Intuitively, many of us can see that there is a strong link between our bodies and our emotions because we have probably had experiences where our bodies were affected by how we feel or vice versa, but it is not so easy or desirable for many scientists to provide evidence for this, despite the potential positive implications of doing so. Possibly because the concept is so abstract (back to Agnes Martin). This is why artists should do some science—our quirky perspectives latch onto ideas more linear thinkers tend to ignore. The future as I see it is inter-and multi-disciplinary. After all, we scholars are all striving after the same thing: understanding and revolutionary insight, right?!*

Instructor's Comment: *Celeste Hackenberg is a Davis Honors Challenge (DHC) student who completed an honors project in conjunction with my Language and Cognition (PSC 132) class, an upper division core course which provides an introduction to the cognitive processes and brain mechanisms involved in language comprehension and production. She elected to review the debate between embodied and abstract symbol theories of language comprehension and emotion, a topic that has puzzled philosophers for many centuries and is currently hotly debated in the psycholinguistic literature. Traditionally, language has been viewed as a "module", with abstract "symbolic' representations of the meanings of words and initial mental operations that are impermeable to emotion. But recent empirical evidence suggests that language is actually embodied, and that emotional information is an important part of the representation of the meanings of words. Celeste beautifully lays out both sides of the debate and concludes that "symbols can, but do not have to be grounded in embodied states, such as emotion". She also suggests an interesting translational angle: if language can indeed be grounded in emotion then this would present promising opportunities for clinical applications, for example in improving memory and language performance in patients with depression. Celeste's review reveals her passion for research and her ability to integrate complex ideas and make them accessible to a more general audience. I really enjoyed reading her wonderful review.*

*—Tamara Swaab, Department of Psychology and Center for Mind and Brain*

**Abstract**

Traditionally, literature on language comprehension has viewed the organization of words in the human brain based on theories that view words as being represented as abstract symbols in an interconnected network (Fodor, 1975; Collins & Loftus, 1975; Landauer & Dumais, 1997; Burgess & Lund, 1997). However, more recently, an alternate view known as embodied theory has begun garnering support among cognitive psychologists. One basic tenet of this theory pertinent to the present review is that emotion is experienced in the body, thereby influencing cognitive processes including language comprehension (Niedenthal, 2007). To test embodied theory, researchers have used evidence from Kutas and Hillyard's (1980) ground-breaking event-related potential (ERP) experiment, which showed a particular component known as the N400 is sensitive to semantic incongruities. In adapting this research to study how language affects emotion, it was repeatedly found that N400 also responds to emotion, which supports the embodied view (Chung

et al., 1996; Chwilla, Virgillito, & Vissers, 2011; Egidi & Nusbaum, 2011). This review argues in favor of an adapted view of language comprehension and emotion, which includes both abstract symbol and embodied theories, considering the strengths of both the classical models and this newer experimental evidence. Some clinical implications of accepting this new model are also discussed.

## Introduction

For the past several years, amodal abstract symbol theories have been used to explain how meaning might be created in the brain. Fodor's theories, beginning in the 1970s, have had a strong influence on how cognitive scientists understand language representation up until the present day. In these theories, as Niedenthal (2007) remarks, information is taken in through the senses and converted into abstract symbols, which then become separate from the systems that encoded them (p. 1003). Building on Fodor's theories and Quillian's preliminary memory modeling (1967), Collins and Loftus (1975) showed that meaning could be drawn out of connections between abstract symbols. Later computational models such as Latent Semantic Analysis (LSA) and Hyperspace Analogue to Language (HAL) were also developed to further test Collins and Loftus's semantic network theory (Burgess & Lund, 1997; Landauer & Dumais, 1997). These abstract symbol models will be explained in more detail later in this review.

These theories might be fine for representing meaning *once the initial input has been acquired*, but, the question becomes, how can meaning be drawn out of meaningless abstract symbols, the form in which they are initially encoded? This disconnect between words represented as abstract symbols and their eventual registered meaning is known as the *grounding problem* (Barsalou, Simmons, Barbey, & Wilson, 2008). Furthermore, are these symbols really language-specific as Fodor originally claimed, or can the mental representation of these symbols be influenced by other factors such as perception, action, and emotion during word processing, as embodied theorists believe they can (Niedenthal, 2007)? A great deal of previous research on the embodied theory of meaning has focused on perception and action, but several recent studies have linked emotion to meaning as well. According to Havas, Glenberg, and Rinck (2007), emotion might influence language in the following two ways: first, the motor system may encode the meaning of words in terms of approach or withdrawal, and second, mood congruence might facilitate understanding.

The present literature review will outline the major theories and research behind both the abstract symbol theory and the embodied view of semantic comprehension. The question of whether and how emotion might influence the way language is understood is an important one because it could change the way scientists study the brain's representation of word meaning. Abstract symbol theories have been very useful in explaining many cognitive phenomena, but embodied theories have the potential to create a more comprehensive understanding of what is really happening in humans when they hear words and make meaning out of them, being that they show the body to have an effect in addition to the brain. In recent years, cognitive neuroscience methods of inquiry have made it possible for researches to empirically assess the debate between abstract symbol theories and embodied theories of language meaning. A great deal of evidence has amassed in favor of each respective theory. Considering both the abstract symbol theory *and* the embodied view of language comprehension appears to be necessary for the studies of emotion and language to progress. A combined view could have strong clinical implications since language likely has an impact on mood and vice versa.

## Abstract Symbol Theory

Fodor's *Language of Thought* (1975/2006) has been widely influential in cognitive science. In his hypothesis, thinking occurs in the brain as associations between syntactic and semantic concepts. His theory of *Modularity of Mind* (1983) goes further to posit that what makes a cognitive system modular is the fact that it is domain specific and informationally encapsulated, meaning that it "only responds to stimuli of a particular class" and does not have "complete access to a person's expectations, beliefs, presumptions, or desires" (Coltheart, 1999, p. 118-119). In sum, Fodor's theories maintain that emotion does not play much of a role in processing or representing information coming in through the senses (Niedenthal, 2007). This section will begin by examining Collins and Loftus's (1975) memory network model, followed by two of the most basic computational models of abstract symbol theory (LSA and HAL), and end with one of the few experimental studies that have been conducted to provide support for abstract symbol theory to understand its implications.

According to Collins and Loftus, who have spent years building on Quillian's famous spreading activation theory (1962, 1967), meaning is

derived from greater activation of related nodes in a memory network. Each concept in memory is represented by a node, and its properties exist in the form of links that connect to similar nodes. Links, or properties, are essential for representing the meaning of a concept, since they allow activation to spread from word representations activated by the input to word representations held in memory. Other nodes are more or less semantically related depending on the number and strength of links shared between them. Thus, semantic relatedness depends on number of shared properties, and not on other factors such as emotion.

In addition to Collins and Loftus's theoretical framework, two main computational models have also been created based on the abstract symbol theory. First, Landauer and Dumais (1997) found that their LSA model could offer an explanation to the centuries-old debate, dubbed "Plato's problem," concerning why it is that people seem to have more knowledge than they are actually able to access at any given time. The researchers explained that, while the brain might not function exactly like a computer, it *does* store and reprocess the input it has acquired through experience. Since words with related context appear in close proximity in high-dimensional space given a computer simulation of the human brain, a word that someone may think that they do not *know* may still exist in memory and become activated through induction (p. 226). In this model, after sensory input has been received, words are positioned using mathematical calculations that make use of information about word frequency and co-occurrence within a particular paragraph, with the key assumption that co-occurrence is a function of similarity between words. No knowledge about the world is required in the computer model, since, once encoded into the system, meaning is abstracted from associations between symbols representing the words alone. LSA has been used to model children's learning abilities, and the model is on par with test-takers of Foreign Language exams in its abilities, thereby demonstrating its real world relevance (p. 220).

The HAL model provides another solution to the problem of meaning representation in the mind. A 1997 article by Burgess and Lund outlines the creation and applications of their computational HAL model. To develop a matrix representing the semantic relationships between words, the researchers accessed Usenet, a database with a large, conversational vocabulary thought to accurately represent the types of language stimuli humans encounter naturally, through which they were

able to gather around 300 million words from internet discussions on a vast array of topics. These words were subsequently analyzed to measure their co-occurrence, without any interference from subjective human judgment. A vocabulary was then created, comprising 70,000 of the most commonly-occurring words. Words were ranked depending on their position relative to one another in 10-word "moving windows" and subsequently positioned as "vectors" or coordinates in a 70,000 x 70,000-word high-dimensional space. The researchers found that HAL was able to categorize words based on their semantic properties, given that they were arranged according to the context in which they commonly appeared in everyday language. Burgess and Lund argued that, since they used a methodology for encoding abstract words that was effectively the same as the one used for concrete words, they did not have the same grounding problem that other computational models have had. However, the researchers noted that limitations did exist, in that this computer program did not have access to all kinds of representations that exist in human experience apart from strictly linguistic input (p. 205).

Finally, although few empirical studies have been conducted on abstract symbol theory, there are a handful of examples that can be used to gain a better or more practically applicable understanding of abstract symbol theory. One such example comes from a 1971 experiment in which Bransford and Franks look at how semantically meaningful ideas were integrated in memory. The researchers conducted a series of experiments to test whether exposing participants to partially related ideas would result in their believing that they had heard a sentence that combined these components, when in fact they had not. Short simple sentences were presented, and participants were told they would need to remember them to answer questions later. Later, longer compound sentences were presented, which included segments that were semantically related to the short sentences, but never once actually expressed. For comparison, participants were also presented with "noncase" sentences, which did not have any semantic congruity between parts although some individual words may have repeated. Finally, participants were asked whether or not the sentence had appeared before and how confident they were about their decision. Experimenters believed not only that the participants would *think* they had heard these sentences before, but that they would also be confident about having done so. Their hypothesis was confirmed, revealing that meaning was abstracted from connections formed between

the shorter sentences, which, interestingly, did not exactly correspond to the mere presence of repeated words (p. 348).

**Embodied Theory**

One of the earliest arguments against the idea that computers could model human understanding comes from philosopher John Searle (1980). He likens the process of a computer creating meaning to that of a non-Chinese speaker formulating responses to Chinese questions by way of a reference book, which gives instructions on how to combine symbols. This thought experiment, known as the Chinese Room Argument, calls into question what it means to truly *understand* something, apart from being able to represent the correct answers as these computational modeling programs can do. Is symbol manipulation enough to say that one has true semantic understanding of a concept? This question lays the groundwork for an embodied theory of language comprehension.

In contrast to abstract symbol theory, in which most support comes out of a theoretical framework, a great deal of research on the embodied theory has been experimental in nature. In more recent years, cognitive neuroscience has proven to be a valuable tool in determining how emotion might influence language comprehension. In 1980, the N400 event-related potential (ERP) component was discovered by Kutas and Hillyard. This negative-going wave was a breakthrough for studying language using ERP data. It was first noted for its tendency to follow words showing semantic incongruity with the rest of a sentence. Potential explanations include the possibility that the negative peak at 400ms is a sign of "re-processing" incongruent information. In the original paper, Kutas and Hillyard note that the N400 could potentially be related to other kinds of violations, which might include emotional expectancies.

In 1996, Chung, et al. tested whether the N400 effect for semantic incongruity would extend to emotional incongruity. Their hypothesis was that mood state would have the same effect on word processing as linguistic context, and therefore, the positive or negative mood of the participant would determine whether or not the N400 effect would be observed for positive or negative words. The researchers induced moods in the participants by having them remember positive or negative life events. The participants were then presented with short paragraph-long stories where the last word could either be positive or negative, congruent or incongruent with their mood. The expected results were

generally observed with mood-incongruent endings eliciting higher N400 amplitude.

Building on this and previous research, Chwilla, Virgillito, & Vissers (2011) zeroed in to directly test the opposing hypotheses of the long-accepted abstract symbol theory of language processing against the newer, less understood embodied theory. The two theories have clear and opposing positions on whether or not emotion can affect language processing. Since, in the abstract symbol view, the representation of meaning is absolutely separate from processes that are active during encoding, such as motor action, perception, and emotion, proponents of this view would claim that there should be no interaction between emotion and language comprehension. Embodied theorists, alternatively believing that meaning is grounded in the very processes that abstract symbol theory sets aside as separate, would predict that an interaction between emotion and language would necessarily occur. Using the ERP method, Chwilla, Virgillito, and Vissers tested the claim that N400 amplitude is influenced by mood by investigating whether an interaction occurred between cloze (high or low probability words embedded in sentences) and mood (happy, sad, or neutral), which was induced using video clips. They found the unambiguous result that mood did in fact interact with cloze effect. Specifically, their original hypothesis that participants in a happy mood would show a larger N400 cloze effect than those in the sad mood condition was confirmed. Importantly, those in the sad mood condition seemed to re-analyze sentences that didn't make sense, and it appears that "standard cognitive phenomena" are most often associated with those in a happy mood, which may have implications for treatment in patient populations (p. 2411). In fact, by using a wide range of moods (happy, neutral, and sad), these researchers hoped that their study might be a first step towards studying the relationship between language and mood in patients.

In another 2011 study, Egidi and Nusbaum set out to test whether mood affects semantic processing in discourse. One aspect of their study looked specifically at whether meaning or semantic integration is facilitated by mood congruence or if incongruence causes difficulty in processing. Knowing that in discourse comprehension, the amplitude of the N400 wave increases as a function of difficulty in processing semantic incongruities, the researchers induced mildly positive, neutral, or mildly negative moods in their participants with video clips and then had them

read discourse passages with mood-congruent or -incongruent endings to see if mood incongruence would also produce this effect. The researchers found an interaction between mood and ending valence, with those in the mildly positive and neutral conditions showing greater N400 amplitude (difficulty in processing) for negative words but no facilitated processing for positive words, while those in the negative mood condition showed facilitation for the negative words as well as greater difficulty processing positive words. The researchers suggest that mood acts as a filter for semantic information in discourse and is not simply another aspect of linguistic context. Therefore, they argue, mood is an essential part of a detailed understanding of language.

## Conclusion and Potential Clinical Applications

There appears to be strong evidence in favor of both the abstract symbol theory and the embodied theory of meaning, so the question then becomes how to reconcile these two seemingly opposing viewpoints. Interestingly, as noted in Louwerse and Jeuniaux (2008), while each side strongly argues in favor of their own position, neither side wishes to exclude the other from a comprehensive account of language comprehension. Louwerse and Jeuniaux claim that a *symbol interdependency hypothesis* probably provides a better theory of processing and representation of meaning in the brain. Here, symbols *can*, but do not *have* to be grounded in embodied states, such as emotion. This hypothesis takes into account the possibility that symbol representation may be influenced by perceptual, motor, and emotional factors, but does not *require* grounding for *all* types of meaning representation. Sometimes it may be more efficient to process information using only abstract symbols, as in situations where deep levels of processing are not required (Louwerse and Jeuniaux, 2008, p. 314). Rather than continuing this debate that has extended from the field of philosophy, beginning with Socrates and Plato, to present-day cognitive neuroscience, future research should focus on developing models that take into account the role that factors like perception, motor responses, and emotion might play in determining how meaning is processed and represented in the brain without doing away with classical models that explain how symbols are sometimes manipulated once they have been encoded.

Practical applications of taking emotion into account when considering language comprehension extend to the world of clinical

psychology and could have a profound impact on the way various mood disorders are treated. For example, Jacobs et al. (2011) were the first to examine the effects of emotion on comprehension and memory for verbally presented discourse in patients suffering from pediatric bipolar disorder. They found that depressed mood had a significant impact on how many details (called *microstructure*) were remembered—much fewer compared to healthy controls, whereas no differences were observed in these children's ability to recall *macrostructure*, or the big picture. Studies such as these are important since proper treatment would include an understanding of why certain aspects of learning might be more difficult for patient populations. Another study, by Kwiatowski and Parkinson (1994), adds to the practical relevance of considering mood's impact on language comprehension. The researchers found that clinically depressed individuals tended to have better recall for negatively valenced words as compared to healthy controls. The study further considered whether mood induction procedures produced results equivalent to studies of individuals with actual depression and found disparities between them. This is significant because it means that future research will have to consider the ecological validity, meaning whether the results of mood induction are comparable to natural mood states, if the results are to be extended to clinical populations.

## References

Barsalou, L.W., Simmons, W.K., Barbey, A.K., & Wilson, C.D. (2008). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*(2), 84-90.

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

Burgess, C., & Lund, R. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes, 12*(2-3), 177-210.

Chung, G., Tucker, D.M., West, P., Potts, G.F., Liotti, M., Luu, P., & Hartry, A.L. (1996). Emotional expectancy: Brain electrical activity associated with an emotional bias in interpreting life events. *Psychophysiology, 33,* 218-233.

Chwilla, D.J., Virgillito, D., & Vissers, C.Th.W.M. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of Cognitive Neuroscience, 23*(9), 2400-2414.

Collins, A.M., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences, 3*(3), 115-120.

Egidi, G., & Nusbaum, H.C. (2011). Emotional language processing: How mood affects integration processes during discourse comprehension. *Brain and Language*, *122,* 199-210.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J.A. (2006). The language of thought. In Bermúdez, J.L. (Ed.), *Philosophy of psychology: Contemporary readings* (pp. 101-106). New York: Routledge. (Original work published 1975).

Havas, D.A., Glenberg, A.M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14*(3), 436-441.

Jacobs, R.H., et al. (2011). Negative emotion impacts memory for verbal discourse in pediatric bipolar disorder. *Bipolar Disorder, 13,* 287-293.

Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203-205.

Kwiatowski, S.J., & Parkinson, S.R. (1994). Depression, elaboration, and mood congruence: Differences between natural and induced mood. *Memory & Cognition, 22*(2), *225-233.*

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*(2), 211-240.

Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied *and* symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Embodiment and meaning: A debate* (pp. 309-326). Oxford: Oxford University Press.

Niedenthal, P.M. (2007). Embodying emotion. *Science, 316,* 1002-1005.

Quillian, M. R. (1962). A revised design for an understanding machine. *Mechanical Translation, 7,* 17-29.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science, 12,* 410-430.

Searle, J.R. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences, 3,* 417-457.