

**Meaning and context:
the implications of LSA (latent semantic analysis)
for semantics.**

[John R. Skoyles](#)

CPNSS, LSE

Abstract

Until recently, cognitive science understood semantics in terms of its using the information contained in first-order (or direct) associations. In Psychological Review (1997), Landauer and Dumais in an approach they called latent semantic analysis (LSA), showed that the acquisition and comprehension of word meaning depended upon the processing and extraction of a previously unexamined kind of information (hidden in word context and past word usage): higher-order (or indirect) associations. Higher-order association arise from the totality of associations in past usage that every word has with every other one. A diverse and heterogeneous collection of meanings are also context dependent (and so based upon higher-order associations). (For example, those that underlie the social meaning of human interactions and intentionality .) Before LSA, the ability to theoretically and empirically explore the links between meaning and context through such associations was severely limited. LSA has now corrected this situation by providing an objective, and in the case of language, a tested model of the relationship between meaning, usage constraints, and context. I explore the implications of this for how we understand our capacity to experience meaningfulness not only of words but social and intentional existence.

Introduction

Recently, a novel approach has been created, latent semantic analysis (LSA) (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998; see also the related HAL model, Burgess & Lund, 1997) that provides both a computer-based and a scientifically tested account of the role of context in the development and functioning of semantics. It is a two part model. First, it provides a mathematical account of the presence and extraction of a previously unstudied form of information (higher-order associations) contained in the context surrounding words and general past word usage. The mathematics linking higher-order associations and context in LSA, though applied to linguistic meaning, are applicable (as its authors note) to the information processes that underlie potentially a wide range of other types of semantic experience. Second, confirming the relevance of such information to context processing, LSA offers a successfully tested computer simulation model of how this context information is responsible for the development and comprehension of word and sentence meaning. A wide range of empirical data exists upon such semantic activities and processes -- the computer simulation of LSA in many cases not only provides successful models of them but gives them for the first time a mathematical foundation. Notably, this is the case for one aspect of child development that has until now been mysterious: how children learn the meaning of unknown words. Formerly seen as an enigmatic stage in development, LSA has demonstrated that children, given the context information provided by surrounding known words and past word usage, can readily guess their unknown word meanings. LSA's success in accounting for this, and other diverse measured aspects of sentence and word comprehension, strongly suggests that the information contained in the higher-order associations is crucial to the human experience of what it is for some event, circumstance or thing to have meaning.

Context and Meaning

Historically, LSA originated in the need of computer scientists to find an automatic means to retrieve documents by keywords (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990). Computer scientists faced the problem that the link between keywords and the words in sought-after documents are context dependent. If a searcher types in, say, the keywords, “film” and “Marilyn Monroe”, they want to retrieve not only documents that mention “film”, but also related synonyms found in the same context (such as “movie”, “Hollywood” and “motion picture”). Further, they want to retrieve documents that only contain the word “film” that fits in with the context of “Marilyn Monroe”, and not homograph meanings found in different contexts (such as “film” meaning “paint surface”). Computationally, however the context sensitivity needed to identify such synonyms and homograph meanings cannot be reduced to the information contained in the circumstance of a word and its direct associations with adjacent ones (Haugeland, 1989). To solve this, computer scientists developed methods to determine the synonyms and homographs of keywords by extracting from texts a type of information that had been not previously investigated -- the higher-order (or indirect) associations words have with each other. (Previous work upon the information in texts [due to limits upon computer power] had been confined only to first-order or direct associations.) Computer scientists using information extracted from higher-order associations have been able to develop methods to detect synonyms and homographs efficiently (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). These techniques have been used to create powerful keyword document retrieval programs (Deerwester, Dumais,

Furnas, Landauer, & Harshman, 1990).

Meaning -- the thing that LSA computerizes -- until recently has been quixotic to scientific and mathematical investigation. After all, there is the problem of what does it “mean” for a word or sentence to mean something? Philosophers since Socrates have studied this for two and half thousands years without being able to provide a clear answer. One approach to understanding its nature is to explore what might be called its “behavioral properties” (in contrast to the “semantic” ones studied by philosophers that concern reference or sense).

First, all words have meaning, and different words have various degrees of similarity and distance of meaning. “Vast” and “big” have a great deal of similarity and closeness, “vast” and “small”, a smaller amount, while “vast” and “six” are distant. What might underlie this distance/similarity property, however, is uncertain: the resemblance or lack of meaning between words cannot be inferred from their identity or their close association with other words (the level of first-order associations). Nothing, for example, about the perceptual identity of “vast” and “big” -- its letters -- tells us that they are synonyms. Nor is this information provided by their immediate associations with surrounding words (Haugeland, 1989). In spite of this, humans have an intuitive grasp of which meanings are similar and which are not. This suggests that our experience of words might in some unknown way be providing us with the information with which to make such judgments. Meaning has the property that words that have roughly the same meaning -- synonyms -- are context intersubstitutable. Consider “vast” and its synonyms, “big”, “large” or even “gargantuan”. Each of these three words can be substituted in most sentence contexts without significant change of sentence meaning. “The big rock broke the wagon”, for example, means

roughly the same as “The vast rock broke the wagon”, and even, “The gargantuan rock broke the wagon”. Like the distance/similarity of word meaning, while humans can understand intuitively which words can be swapped, it has proved impossible by philosophical or other analysis to specify why.

Linked to context is another property, this time of words and texts: we do not expect words to appear randomly in grammatically correct sentences. Expectations and constraints exist -- word usage patterns -- about which words tend to appear with which other ones. This can be experimentally shown with sentences in which words are omitted and subjects are given multi-choice options to pick the missing word -- cloze sentences (Taylor, 1953). If given the incomplete sentence, “The X stone broke the wagon” and asked to pick the most likely word out of “big”, “forgetful”, and “sweet”, you would pick “big”. “The big stone broke the wagon” has an expected pattern of usage, that is absent in the sentences, “the forgetful stone broke the wagon”, and “the sweet stone broke the wagon”. Indeed, even without suggested word choices, if single words are cut from writing, readers can usually guess, using the remaining words before and after them, about half of them (Miller & Coleman, 1967). Further, if the next word in a piece of ordinary writing is covered and a person has not read ahead, they will be able to guess it about one in four times (Miller & Coleman, 1967: experiment 3; Gough, Alford & Holley-Wilcox, 1981: 91-92). The ability to fill in missing words from sentences is an important means of testing the comprehension progress of language learners (Storey, 1997) being used to assess language competence, for example, in TOEFL (Test of English as a Foreign Language) certification. This argues word usage contains information central to the comprehension of word meaning.

LSA shows that the complex patterns of association contained in word usage can be extracted from many tens of thousand of sentences. When they are analyzed, the information they provide is sufficient to enable the LSA model to identify synonyms with a competence equal to humans. For example, applicants with English as a second language to US colleges if given a word and four possible synonyms get 64.5% correct; the LSA model, 64.% (Landauer & Dumais, 1997, p. 220).

The information that enables LSA to do this comes from various sources of associations present within word usage. These include already studied sources such as word frequency (how common a word is), and word frequency associations (how commonly a word is found next to another one). LSA, however, depends (and it turns out, much more so) upon information that exists at the higher level in which words associate through indirect (or higher-order) associations -- that is associations based upon links between first-order associations (for what these are see below). While the role in language meaning of direct associations has already been studied by philosophers and scientists, due to the difficulties of computationally extracting and modeling higher-order associations, this had not been previously done for this higher-order kind of information. LSA has, however, now started investigating its contribution to the processing of language meaning. This has fundamentally changed scientific ideas about the importance of the information contained in word usage. When first-order associations were studied, this was found to be insufficient, and as a result, it has been thought that the associations contained in word usage did not contain the information required to model word meaning (Haugeland, 1989). LSA by extracting higher-order associations and modeling them into successful and working language simulations has shown that this conclusion

was premature.

Higher-order Associations

To discuss this recently discovered source of information -- higher-order associations -- it is best first to briefly describe first-order ones. The word “autistic” provides an example of a first-order or direct association: this word is often closely followed by the word “individual” or “child”. If the word “autistic” appears in a sentence, then it is likely it will be followed by the word “child”. Since it is statistically easy to count different words and find how often they appear near to each other, these frequency associations have been studied in the past (Haugeland, 1989). Indeed, the mere familiarity with them in writing tells us that such associations commonly exist. Such readily appreciated associations, however, do not contain the only type of information that exists amongst words; nor do such associations necessarily hold the most important information. Two words not only have a direct association with each other, but also many indirect ones via separate associations with further words. Consider the synonyms “big” and “vast”: they are strongly associated (they are synonyms) yet they rarely appear in the same sentences. People do not usually say “the big and vast stone broke the wagon” -- if you use one word, you will not repeat yourself and redundantly use the other. Instead of first-order associations, what links “big” and “vast” is occurring in the same kind of sentences, and so with the same kind of associations to other words. On balance, it is often a matter of indifference as to whether “big” or “vast” gets picked for use in any particular sentence to indicate large size. Because “big” and “vast” appear in the same kind of sentence contexts, they have separate but parallel (and so similar) associations with other words. If “big” appears frequently with another

word, say, “stone”, so predictably will “vast”, if “big” does not with another word, then neither will “vast”. As a result, at a higher level, there is a strong link between the associations “big” has with other words, and the ones “vast” has. The information contained in this link reflects a higher level association between them. Higher-order association exist not just in regard to “stone” between “big” and “vast” but also with each of the many tens of thousands of other words in vocabulary. Moreover, such high-order associations occur not just between “big” and “vast”, but every other pair of words that can be combined together (though most such associations contain little information). Higher and even more indirect levels of word association also exist. Some of the words with which “big” associate, associate with it more or less strongly in the presence of yet further words. If the word “big” appears in a sentence with “stone”, it likely that “heavy” will also appear. Thus near innumerable associations of a further and even higher level exist. Together they in subtle and complex ways shape a web of information that permeates our entire vocabulary. As in a spider’s web, where a pencil distorts immediate silk strains, that in turn transmit tension throughout the web and so determines its final deformation, the actual information defining the context of a word in a sentence depends upon associations that exist far away. This web of hidden associations LSA has shown underlies the information responsible for word meaning.

The word constraint information provided by higher-order associations is extracted by LSA by analyzing tens of thousands of different episodes of past sentences and paragraphs (a process described below). This extracted information enables computers to generate meaning similarity and distance judgments. Significantly, for establishing the importance of this information, these computer judgments match those provided by human subjects

(Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). If higher-order associations extracted by LSA from word context constraints were a mere epiphenomenon, this would be surprising. As a consequence, the success of the LSA model argues that higher-order associations hidden in word usage constraints are responsible for the meaning of words (as reflected in similar and dissimilar judgments). The LSA computer model further tells us that three quarters of the information needed to make such semantic judgments comes from higher-order associations not first-order ones (Landauer & Dumais, 1997, p.226). Theoretically, this might seem surprising since compared to first-order ones, most higher-order associations provide only weak information about word usage. However, higher-order associations are by many orders of magnitude much more numerous than first-order associations and so when added up in total contain much more information.

Semantic Space

The raw context information contained in higher-order associations is difficult to manipulate directly. To be useful, it needs to be mathematically transformed. This is done by converting it into a multi-dimensional space in which hundreds of dimensions locate words in regard to their context substitutability. In this space synonyms occupy the same locality (they are little different in their substitutability in any context), while homographs with their different meanings are widely separated (they produce different meanings when put in different contexts). More generally, the more similar the meaning of different words, the closer they are in this space. In this multi-dimensional space, locations are not only confined to words: places are also given to complete and incomplete sentences. A word can be synonymous, after all, not only with another word, but a group of words -- for example, the

definition -- “being of extreme size” is synonymous and substitutable with “vast”. Sentences and other groups of words gain a location in this multi-dimensional space by having all the various locations of their individual words mathematically added together (see below). This provides a mathematical “center” for all their different individual locations.

This multi-dimensional space enables keywords to be linked to the documents in which they appear. By checking the location of keywords provided by a searcher, a program can find words with similar meaning (since they are found in similar contexts). Moreover, the different homograph meanings of the keywords such as “film” can be separated so that the locations of accompanying keywords (“Marilyn Monroe”, “paint”) can be used to select which of their meanings (movie, chemical surface) the searcher intends. A further use of this space arises from the above noted ability to add the locations of words in a sentence to create a new location. LSA shows that when this is done the known words of a with sentences containing an unfamiliar word, this provides some rough information about the unknown word’s meaning. Though in the case of only one sentence, this information may be not precise, it might still be sufficient to decide between two competing possible meanings for the unfamiliar word. With increasing numbers of sentences, however, the combination of the rough locations goes further and can identify its meaning.

This ability to infer unknown elements from the context provided by known ones makes the high-dimensional space (generated from higher-order associations) useful as model of many previously unexplored or enigmatic aspects of how we experience and acquire meaningfulness. In particular, it is relevant to one puzzling aspect of child development: how children acquire the meaning of unfamiliar words. Children learn on average 10 to 15 new

word meanings each day, but only one of them can be accounted for by direct instruction. The other nine to 14 word meanings are picked up in some other unknown way. Landauer and Dumais (1997) showed that when children meet an unfamiliar word, its context enables them to guess its meaning. They can do this since they can estimate the average location of the words surrounding it (the location of its context), and this suggests (via its position in the multi-dimensional space) its potential meaning. In effect, words are learnt backwards from their context. To understand this, consider you do not know the meaning of “gargantuan” but you know many sentences that contain “vast”. You are then given many new sentences that contain “gargantuan”. Since “gargantuan” and “vast” are synonyms you could swap “vast” into the ones that contain “gargantuan”. But, of course, you will not easily recognize that the sentences that fit “gargantuan” also fit the ones that fit “vast” (and vice versa). If you could, you guess “gargantuan” has a meaning close to that of “vast”. (It can be swapped into the same sentence contexts). This is what the information in the higher-order associations of the words surrounding “gargantuan” do, in effect, when they yield the same location in the multi-dimensional space as “vast”. The same principle applies when the two words are not synonymous but have closely related meaning – though, of course, the further they are apart in meaning the fewer context associations they will share.

This problem of context identification is not confined to words: a common problem faced by the mind is to determine the nature of something -- an unusual object, an unfamiliar response or an unidentified aspect of a situation -- given the information of surrounding events and entities which gives it meaning and relevance. As discussed later on, this is particularly the case with our sense of social meaning and our understanding of

intentionality. If suitable past experience exists to extract higher-order associations from previously encounter context associations, such information can be used via its conversion into a multi-dimensional space to identify information about a present unfamiliar entity from its surrounding context. Since this is the situation with many experiences, many entities (social situations, minds) fall into this information processing situation, suggesting that the extraction of higher-order associations from context might be widely exploited by the mind to experience widely different types of meaningfulness not just that of words.

LSA

The above mathematicisation of context was introduced in 1997 in a paper in Psychological Review called, “A Solution to Plato’s Problem: The Latent Semantic Analysis Theory of Acquisition, Induction and Representation of Knowledge” authored by Thomas Landauer (of Colorado University) and Susan Dumais (then of Bellcore and now Microsoft). In this paper they show that the LSA account of context in terms of extracted higher-order associations can be used to simulate successfully the measured performance of many aspects of language that use context such as word sorting and category judgments, estimations of passage coherence, and the quality and quantity of knowledge contained in student psychology essays (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998).

While LSA is a computational model, it is important to note that it is quite different to other computer models familiar to cognitive psychologists. To appreciate this, I will first give here these a brief and a simplified outline of its mathematical processes (Landauer & Dumais, 1997); and then list some of its differences from other approaches.

LSA extracts from a very large corpus of words-- nearly five million words

from a child's encyclopedia (other researchers in a related approach called HAL [hyperspace analogue to language] have used ones of even greater size -- 300 million words from Internet newsgroups, Burgess & Lund, 1997). After preliminary processing, this text is then turned into a 60,000 columns of word types by 30,000 rows of paragraphs matrix made up of the different vocabulary words against the different paragraphs in which they appear. This matrix is then converted by a mathematical technique related to factor analysis called singular value decomposition into a multi-dimensional space of 300 eigenvectors. In this multi-dimensional space, position is located mathematically with vectors, and their closeness is measured in terms of the cosines between them (since these details are not conceptually important they are not discussed further). As presently formulated, these dimensions lack obvious meaning (the dimensions while describing semantic space do not themselves represent meaningful semantic dimensions). However, mathematical approaches exist that suggest extractions might in the future be made that provide meaningful dimensions (Lee & Seung, 1999). In doing this extraction, LSA is radically unlike previous computer models in cognitive psychology. First, the process is computationally intensive, not like the calculation of word association frequencies that are done by hand or with more basic computers. This explains why, even though the study of associations and meaning has a long history, that is it is only recently that higher-order associations have been investigated. Its mathematics, moreover, go far beyond the scope of normal insights about how entities associate and the analysis with which philosophers and psychologists are familiar: if a single cell in such a matrix is square millimeter, than a 60,000 by 30,000 matrix is the size of a 60 meter by 30 meter piece of one millimeter graph paper – the processing done across such matrixes is unimaginable to ordinary

intuitions into what might constitute an actual or efficacious information process.

Second, since it is based upon the extraction of pre-existing but “latent” information, it is not analogous to programming accounts of cognition (such as those associated with Artificial Intelligence, AI). Nothing from human knowledge is added by way of programming: all the information it contains comes from extracted associations.

Third, it does not incorporate human judgments such as with WordNet (Fellbaum, 1998) about the nature of the information that underlies semantic associations -- again all the information is there in past usage -- the model only extracts it and converts it into an usable form.

Forth, it does not involve the training of networks with external error correction adjustment such as with supervised neural network models: the information is instead extracted by mathematical techniques. (Though as Landauer and Dumais note, 1997, p.217, this information need not be potentially different to that acquired when such networks are appropriately trained.)

The success of LSA does not mean that it is a perfect or a finished model of context-based cognition. Its extraction of higher-order associations is done in a crude manner without exploiting any of the potential advantages that normally aid the understanding of words offered by syntactic, logical, discursive, situational, sensory or motor associations. It simulates word meaning instead at the limited level of treating texts as collections of “bags of words” (Landauer & Dumais, 1997, p.226). However, the fact that it so successfully overcomes these limits indicates that the LSA model, and its use of the information contained in higher-order associations catches (in spite of such limitations) something crucial and previously not investigated about

word meaning.

An important question not asked by Landauer and Dumais is what happens to word learning when the processes described by the LSA model are impaired and children cannot use context either to learn words or comprehend them. Do individuals, for example, exist with abnormalities in how they understand words because they are unable to use context as described by LSA? The clinical literature contains many cases where retardation causes children to be delayed in learning the meaning of words, but actual abnormalities in what is learnt are rare. Interestingly, the one condition in which they do frequently occur is autism. This raises the possibility that LSA might provide possible insights into their language and their other difficulties.

The Synonymy and Homography of Situations

LSA investigates the information contained in the associations between elements (words) in episodes (sentences). As shown above, LSA has established that most of that information is derived from hidden higher-order associations. LSA does not ask where this information comes from: it merely extracts from context and shows that it generates the capacity to make meaning judgments. The source of that information however is almost certainly associations in the real world that arise when we describe its elements (words) in terms of episodes (as depicted in sentences). This is not to suggest that these associations have an independent existence: episodes, after all, do not exist in the real world separate to us, but result from, our need to describe and process the world in terms of units.

At a level more abstract than language, things can have relevance or meaningfulness attributes akin to “synonymy” and “homography”¹. A car, a

donkey, and a sedan-chair may look different, but they provide similar means when on holiday to get from one place to another. They -- at least if our desire is holiday transport -- can substitute for each other (much as the words, “vast”, “big” or “being of extensive size” can in the same sentence substitute for each other). Here, however, what they share is not the same word meaning but the same functional relevance as required by our needs and our appreciation of those needs. “The tourist went in a car to the shops”, or “The tourist went on a donkey to the shops”, or “The tourist went in a sedan-chair to the shops” while different as physical activities are similar in terms of their being relevant solutions to our need to travel. (And, reflecting this, the sentences describing them mean roughly the same). Likewise, in terms of function, an entity can be “homographous” in regard to different relevant concerns: a car in one context can be a means of transport, in another a place to escape rain, and in yet a further one, a status symbol. Such synonymy and homography are important since we seek similarities and dissimilarities between the elements (such as events, processes and entities) in episodes, not in regard to their perceptual commonalties but in terms of their commonality in regard to our concerns. Such relevances may derive their meaning from how we can use things, or how they might affect us; for example, that a particular means of transport in a given context is functionally the same for getting from A to B as another. But how do we learn this similarity in regard to our concerns when nothing perceptually links them? Clues can be gained by extrapolating from what LSA has found about word meaning. We possess tens of thousands of episodes relevant to these concerns known directly (in past personal experience), or known indirectly or vicariously (in what has happened to other people). Much as with words across tens of thousands of sentences, (which contain higher-

order associations), these tens of thousands of episodes also contain higher-order associations. These higher-order associations when extracted provide (as with language) context information with which we can create a multi-dimensional space in which our sense of relevance can locate events, processes and things in regard to our concerns. Such context plays a key role in the flexibility and sophistication of our ability to experience things and situations as having meaning.

Sociability

One area of meaning is sociability. Humans are a social primate. Due to this we possess emotions that bind us with our kin, friends, and other humans. One aspect of this involvement with other people is that we do not interact with them as we might with physical things. Instead, we interact in terms of social meanings. For example, when we meet a friend, how we interact will happen within the context of our past relationship (there might be past obligations to repay); it will happen in terms of our and their relationship to others (they might know a mutual friend whose impressions we value); in terms of what they feel (if our friend is sad, we try and cheer them up); in terms of social recognition (our friend wants to be valued as a special friend), and in terms of social mores, rules, manners, and morals (there are boundaries upon our behavior). All these concerns enrich social interaction with multiple levels of context and meaning; nothing analogous occurs with our relationship with physical things.

One particularly important source of meaning is that carried within each person in the form of their mental states. These goals, desires, beliefs and knowledge shape what people do without being directly perceivable (unlike their behavior). Due to this, we cannot predict what people will do merely by

reflecting upon their behavior (unlike with objects) and have as a consequence to rely upon surrounding context which might illuminate them. In information processing terms, this involves cognitive processes inferring the presence of unknown elements from a context of known ones -- a cognitive task that is akin to the identification of the meaning of unknown words from the known ones of the context of their surrounding sentence. Here, however, the unknown meaningful element – a person’s intentions -- is guessed from observed behavior and its situation.

Anticipation of unknown elements from known ones in a social context is also central to when we interact. Take the example of a greeting: though it has some usual first-order elements such as saying, “hello”, smiling, and hand shakes, these depend upon an interpreted context. A greeting done indifferently causes us offence. One done in another way, might cause us to ask what is troubling someone. The elements involved, moreover, might be much the same, but small contextual subtleties can change radically their social meaning. For example, suppose an individual is carrying shopping bags and they see a friend across the road; that individual cannot make the usual first-order signs of greeting such as, “hello” (they cannot hear you above the traffic), and they cannot wave (their hands are tied). They, however, will improvise this information. They might, for example make an exaggerated smile while mouthing words of greeting, or they might slightly raise one of their shopping bags, and wave it and their body. Stopped from making a normal greeting, they generate behavior which others due to its context can interpret as a greeting rather than annoyance or merely odd behavior. An individual is not concerned with creating a particular “welcome” behavior, but concerned with creating behavior that given the circumstances in which it appears, will be understood as having a particular

social meaning such as “hello”.

Sociability builds upon previous context like sentences build upon previous ones. As a result, like them, sociability inter-chains context. In a text, this interchaining is reflected in the fact that when a sentence follows another, it expands and develops its meaning. Because of this, when you scramble up the sentences of a text, they do not make sense (their meaning depends upon the sequential flow of context). Likewise, sociability involves a dependent flow of context: imagine the episodes of a social situation scrambled up; they would cease to be recognizable as a meaningful social interaction. One action in one context means one thing, at another moment, a quite different thing. Due to this context dependence, if we miss one element in a social interaction, we can guess it from its surrounding ones, much as in a cloze sentence, an omitted word can be guessed from the context provided by surrounding ones (its social narrative).

Social psychological theory gives a central position to social roles, sense of personal identity, self-presentation, and awareness of in-group and out-group boundaries – all of which share a common dependence upon the ability of people to understand or experience events as having meaning within an interpretative context. While well described, social psychologists have not however provided an information processing framework in which to understand them at a more basic information processing level. Phenomena such as “self”, “personal identity” and “social group perception” from an information processing perspective are nebulous since they are proposed to exist without an underlying mechanism that could support their internal generation or perception of their experienced meaningfulness. Social psychologists describe such phenomena as being carried out by processes that are interpretative, sensitive to context and concerned with what is loosely

referred to as ‘meaning’: this suggests we can extrapolate from what LSA has found about the processing of language meaning to such phenomena.

LSA tells us that the higher-order associations contained in the usage of words gets turned into a multi-dimensional space of meaning. By extrapolation, it is reasonable to suggest a similar process underlies the context information processing of these components of our sociability. After all, we experience many tens of thousands of episodes of social interaction, and so have the appropriate higher-order associations needed to generate multi-dimensional spaces in which to process social information and generate social phenomena (such as “self” , “personal identity” and “social group perception”). Given that such information is available, that it is known to be used by one meaning capacity already (language), and that there exists a need for such information processing system, parsimony suggests that we should assume (unless we have good reason to think otherwise), that this information processing does indeed underlie sociability. If we do not, we have to explain why we have an experience (sociability) that bares similarity in information processing terms to another (language), that has the information (thousands of episodes), and yet uses some different and unknown process. Ockham in his razor requires that we do not to multiply entities beyond necessity. Thus, even though we lack precise models, it is reasonable to suggest social meaning depends upon LSA-like processes.

Social psychology tells us that emotions deeply link to these above discussed context aspects of social meaning. Limits of space preclude upon this any but the most preliminary comments. However, one needs to note that the interaction between emotions and social context are likely to be important. For example, consider this description of our need for social acknowledgement by William James:

We have an innate propensity to get ourselves noticed, and noticed favorably, by our kind. No more fiendish punishment could be devised, were such a thing physically possible, than that one should be turned loose in society and remain absolutely unnoticed by all the members thereof. If no one turned round when we entered, answered when we spoke or minded what we did, but if every person we met “cut us dead,” and acted as if we were non-existing things, a kind of rage and impotent despair would ere long well up in us. (James, 1891, pp. 292-293).

People not only exist in a social context, but James suggests here that our emotions make us concerned that others experience us as a meaningful part of their social world.

Theory-of-Mind

In spite of minds being hidden, we have good abilities to appreciate how the behavior of people is shaped by their intentionality reflected in what they plan, their motivation and what they know. For example, we can predict if a person believes a toy is in a box, that they will act in one way, and that if they believe it has been moved, that they will act in a different one. Skill in guessing how mental states shape behavior is called theory-of-mind. As noted earlier, this ability is much like guessing the meaning of unknown words from the context of known ones. This skill will thus depend upon our ability to create a multi-dimensional space in which to locate the observed elements of behavior that can be used to infer hidden intentionality. The higher-order associations needed to create this space exists since we constantly experience different situations of behavior that results from different intents, beliefs and desires. Over a few years, this means we have tens of thousands of situations

from which to extract higher-order associations. Such information will enable us to create a multi-dimensional space in which situations can be redescribed in terms of contexts and hidden intentional states. This space thus will enable the known elements of a behavioral situation to provide contextual information about intents, beliefs and desires.

Another source of information for creating this space is language with its terms for intentional states. The multi-dimensional space LSA extracts represents conceptual knowledge -- an aspect that has resulted in LSA models being successfully used to automatically mark the content of student essays in regard to their knowledge of psychology textbooks (Landauer, Foltz, & Laham, 1998). The acquisition of theory-of-mind skills is known to link to language (Astington, & Jenkins, 1999), particularly exposure to mental state words (Peterson, & Siegal, 1999). This suggests it is likely to be extracted by LSA from word usage. Theory-of-mind awareness has not as yet been modeled by computer simulation, nor has the LSA model been tested as to whether it can simulate it. According to the theory proposed here, appreciation of minds derives from extracting and using higher-order contexts. LSA, given the dependence of theory-of-mind on language, and its proposed dependence upon higher-order context, should have the capacity to model it.

Limitations

A problem, reflecting the novelty of the LSA approach, is that there is no general model of the role of context or semantics in cognition and development. The work of Landauer and Dumais has yet to inspire theoretically related analysis in other areas where meaning arises such as sociability, and how the capacity for such experiences are acquired. The

comments made above in regard to them are preliminary aimed to suggest the plausibility of such analysis. However, without further theoretical development, the ability to link meaning to LSA-type processes will be limited. For example, without such theory, the tests and tasks that will enable the role of context and meaning to be analyzed will not be developed. Lastly, the LSA model itself needs considerable development. At present, it ignores much information available to word learning such as syntax, and real world associations. At present, it models the extraction of higher-order associations in terms of a single extraction event, however, it is more biologically realistic to assume the brain does this incrementally as new episodes are increased. How these can be incorporated is unclear though they will be critical in developing LSA into a more accurate model of linguistic and context cognition.

The argument of the above theory might be further criticized as insufficiently specific. Extrapolating what is known about the mathematics of the context processing underling word meaning to the context processing underlying other areas of meaningfulness is unavoidably conjectural, particularly given so little is known experimentally or theoretically (for example, we know little about how children develop a sense of sociability). However, we have good grounds to make such extrapolations: the fact that in one domain – that of word meaning – we find a process that is not intrinsically specific to that domain, suggests it could potentially be found in other ones. That possibility is strengthened when examination of such kinds of meaningfulness shows that they too like word meaning depend upon context processing and that this has so far gone unexplained and unmodeled. In new territory we cannot provide all the evidence we want but logic and plausibility can guide us. Knowing that cognition in one domain solves an information processing

problem in one way, and that this information processing is not necessarily domain specific enables us to think about other domains as possessing similar processes if they share similar underlying information processes. By parsimony we can infer such process are likely even when not proved. To suggest otherwise would be to claim that information processing that solves the problems of one domain and is suitable for another is not used, but rather instead done by some yet further and unknown second one. While this is conceivable, this goes against the requirement of Ockham's razor not to multiply entities beyond necessity. We explore simple possibilities before more complex ones.

Conclusion

Pessimism at present exists as to whether information processing can explain meaning. However, our current lack of insight into semantics might reflect the incompleteness of our knowledge about the variety of information processing operations that underlie context and the experience of meaningfulness. LSA shows that until recently this missed the key role of context in the creation of meaningfulness (derived from extracted higher-order associations); an understandable omission given that its investigation requires computer technology that has only become recently available. In doing this, LSA offers new insights into the problems that people face when learning and comprehending words. Moreover, the information processing that LSA shows exists for word meaning is one that which would also be needed for the development and functioning of many other kinds meaningfulness such as underlies sociability. Thus, LSA provides a new framework to account for why humans experience meaning in social events, interpersonal situations, as well as words. This should enable a new approach

to semiotics.

The theory, as proposed here, is of course, still preliminary. Many issues due to space have been omitted or simplified. Moreover, the LSA model upon which it is based is still in its early stage. But what eventually emerges to explain the meaningfulness of things and circumstances will not be born fully made and will have a strong basis in extrapolation from known to unknown. However, in spite of these limits, this theory suggests a new beginning to understanding them based on established science that is independently motivated of the phenomena – meaningfulness – that it seeks to explain.

Footnote

¹ Due to the lack of appropriate terminological alternatives, I these words, even though normal usage at present limits them to language.

References.

- Astington, J.W. & Jenkins, J.M. (1999). A longitudinal study of the relation between language and theory of mind development. Developmental Psychology, 35, 1311-1320.**
- Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. Language and Cognitive Processes, 12, 177-210.**
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing By Latent Semantic Analysis. Journal of the American Society For Information Science, 41, 391-407.**
- Fellbaum, C. (1987). Wordnet. MIT Press.**
- Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textural coherence with Latent Semantic Analysis. Discourse Processes, 25,**

285-307.

Gough, P. B., Alford, J. A., & Holley-Wilcox, P. (1981). Words and context. In O. J. Tzeng, & H. Singer (Eds.), Perception of print: Reading research in experimental psychology, pp 85-102, Hillsdale, NJ: Erlbaum.

Haugeland, J. (1989). Artificial Intelligence: The very idea. Mass.: Cambridge, MIT Press.

James, W. (1891). Principles of psychology, vol. 1. London: Macmillan.

Landauer, T.K., & Dumais, S. (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., Foltz, P., & Laham, D. (1998). Introduction to Latent Semantic Analysis. Discourse Processes, 25, 259-284.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401, 788-791.

Miller, G. R., & Coleman, E. B. (1967). A set of thirty-six prose passages calibrated for complexity. Journal of Verbal Learning and Verbal Behavior, 6, 851-854.

Peterson, C. C., & Siegal, M. (1999). Representing inner worlds: Theory-of-mind in autistic, deaf and normal hearing children. Psychological Science, 10, 126-129.

Taylor, W. (1953). Cloze procedure: a new tool for measuring readability. Journalism Quarterly, 30, 415-433

Storey, Peter. (1997). Examining the test-taking process: a cognitive perspective on the discourse cloze test. Language Testing, 14, 214-231.