anuscripts should be sent
stylistic guidelines of the
sses and brief
Another separate pag
for use in indexing.
being reviewed by
view process in a timely

nd suggestions to:
iew Editor

Guilford Publications,
1993 (four issues),
individuals, $30 U.S.,
l, Visa, or American
.S. dollars through a
rica must be placed with
, SG6 1HN, United
ing offices. CHANGE
: present mailing label
months after mailing
failure to notify
o Communication
' 10012.
stracts, International
rnals in

ate University,

alifornia,

paper.
tances not falling within
gistered with the
fee of $3.00 per copy is
ion code for

Communication Theory

Three: Three

August 1993

Pages: 183–213

Kathleen M. Carley
David S. Kaufer

Semantic Connectivity: An Approach for Analyzing Symbols in Semantic Networks

We argue that the notions of "symbol" and "symbolic connectivity" can be rigorously developed both from the point of view of the theoretical literature on the symbol and from the point of view of semantic network theory. The theoretical literature, inspired mainly by the literary metaphor, typically takes interpretive density as the chief dimension underlying symbolic expression. Density measures have also dominated the analysis of semantic networks. There is now a sizeable amount of work on the generation of such networks from linguistic data. The majority of that work locates the network and displays it. When attempts are made to analyze the network, the focus is typically on the density (i.e., the number of links) of particular concepts (which serve as the nodes) within the network and on the inferences that can be made about the communicative prominence of such concepts in light of their density. While density is a useful way of analyzing the communicative "connectivity" of a symbol in a message, it provides only one dimension for analyzing connectivity within a semantic network. In this article we offer two further dimensions—conductivity and consensus—with which to analyze semantic networks for connectivity. We illustrate a typology based on these three dimensions. These dimensions and the associated typology form a useful conceptual device that enables the researcher both to specify and differentiate semantic objects within a rich typology for a given domain of analysis. We show this device at work by applying the dimensions and typology to different communication contexts and by discussing other possible domains where they can be applied.

Background

Rhetorical and social theorists have long recognized the capacity of language not simply to express thoughts but also to maintain and change levels of social cohesion between speakers and their audiences. Theorists who approach language as a vehicle of social cohesion refer to its "symbolic" function. The word "symbol" is notoriously vague—yet it is often used as a *precise* alternative to the word "word."[1] Ordinary words are building blocks of structures that have audible or visible manifestations—utterances, discourses, sentences, texts. Symbols on the other hand are building blocks of structures considered fundamentally cognitive and (potentially) social in nature—histories, experiences, beliefs, interests. Symbols are thus often referred to as the "tissue" of social life, connecting our social world—but typically in ways that leave the connector, the things connected, and the nature of the connections poorly specified.

The capacity of symbols to "connect" our social world is an old idea. Theorists of language since Cassirer (1953) have recognized the unusual properties of symbolic language in relation to ordinary signification (Boulding, 1956; Burke, 1966; Duncan, 1968; Durkheim, 1947; Edelman, 1977; Lippmann, 1922; McGee, 1980; Pocock, 1971; Richards, 1936). Any ordinary word can function symbolically just in case its contextual meaning triggers a set of inferences tied to its history of usage over and above its strict denotative meaning (Sperber, 1975). In his study of British political discourse, for example, Hudson (1978, p. 66) notes that in British politics, describing the land speculator as a *slick operator* working in a *plush office* is part of a signature code used by members of the British communist party. Here *slick operator* and *plush office* are phrases with an ordinary sense and reference. Yet in British politics they have been assigned a symbolic meaning (i.e., the meaning assigned by the Marxist platform) through historical patterns of use.

What makes symbolic language symbolic? A variety of responses have been put forward to this question, but they all, sooner or later, converge on the metaphor of contextual connectivity. Graber (1976), for example, has suggested that words function symbolically when they are highly "connected" in their topical contexts, when they compress, or condense, a network of meaning into common words and phrases. She calls words functioning symbolically "condensation symbols" because of their connective properties—because in topical contexts, the use of these words forms many strong associations with the internal elaborations of their audiences. McGee (1980), writing about ideographs, equivalents of condensation symbols, also notes their high connectivity in specific contexts of use:

> ideographs such as "rule of law" are meant to be taken together, as a working unit, with "public trust," "freedom of speech," "trial by jury," and any other slogan characteristic of the collective life. If all the ideographs used to justify a Whig/Liberal government were placed on a chart, they would form groups or clusters of words radiating from the slogans originally used to rationalize "popular sovereignty"—"religion", "liberty", and "property." (p. 13)

The analysis of the connectivity of symbols in specific communication situations has been a *modus operandi* of rhetorical criticism. Moreover, several theoretical traditions have explored variants of the idea that meanings are determined by the connectivity of semantic concepts within social contexts (Bakhtin, 1981; Cicourel, 1974, 1985; Foucault, 1977; Garfinkel, 1967; Richards, 1936; Vygotsky, 1962, 1978; Wittgenstein, 1968). More recently, researchers have explored techniques to extract connectivity relations between the concepts deployed in situated discourse (Alexander & Danowski, 1990; Carley, 1986a; Danowski, 1988; Palmquist, 1990). Indeed, there now exist many approaches in the literature for extracting semantic networks from a corpus of text (see for example, Carley, 1988; Carley & Palmquist, 1992; Danowski, 1982). These techniques generate a semantic network in which concepts are the nodes and the relationships between concepts, the links.[2] The same level of attention, however, has yet to be given to *analyzing* the resulting network (see Knoke & Kuklinsky, 1982). In general, researchers are content to display the resulting network. Analyses that proceed beyond visual inspection begin to consider the connectivity between concepts by examining the relative density of the different concepts in the network (Danowski, 1982, 1988). Sometimes a density analysis is augmented by contrasting networks in terms of their similarity in both concepts and links (Carley, 1986a, 1988; Palmquist, 1990; Saburi, 1991). Although density is an important dimension for analyzing connectivity within a semantic network, there are other important dimensions as well. The purpose of this article is (1) to introduce these other dimensions and show how to compute them on a semantic network of any complexity; (2) to introduce a typology formed by these computations; (3) to apply this typology to (a) the analysis of connectivity *within* specific discourse contexts and (b) the comparison of connectivity levels of the same concepts *across* discourse contexts; and (4) to enumerate multiple domains of application where we have applied these techniques. The dimensions and typology we present are based on those originally developed by Carley (1984, 1987) and further elaborated in Kaufer and Carley, 1993).

Historical and Theoretical Assumptions

The technical considerations on which this work is based are informed by historical and theoretical assumptions underlying the specification of connectivity relations across symbols. While the literature cited above has made much progress operationalizing the idea of symbol in terms of contextual connectivity, there has been less progress refining the very idea of connectivity itself. Previous theorists of symbolic expression have either missed the variable nature of the primitives underlying symbolic connectivity or have not sought to classify and incorporate this variability at the level of explicit theory. The literature on the symbol has mainly taken its direction from the analysis of the literary symbol, often the literary metaphor or allegory, which are well known to elicit multiple levels of rich, often imagistic, inference. This linkage has produced a heavy bias in favor of explaining the symbol in terms of interpretive density, the sheer number of continuous connections that the symbol makes available to the understanding (as, for example, in *the world is a stage*).

Density, however, is not the only primitive constitutive of symbolic connectivity. A second and independent primitive is consensus. Some symbols function as such only because they are connected to historical inferences that are widely shared. A symbol like *1492* has relatively low density for a grammar school student but nonetheless performs a symbolic function because it draws on beliefs that are almost universally shared across that population. An expression like *Mata Hari*, again low in density among a group of adults with a college education, nonetheless succeeds as a widely understood emblem for espionage. Rhetorical theorists, of course, have always understood consensus as a primitive underlying the functioning of some symbols. This understanding, however, has not led to an explicit recognition of the variable foundations of symbols or to a recognition that symbols can draw on the interaction of more than one primitive type. A symbol combining density and consensus, for example, instigates a high number of connective inferences converging on a shared system of belief. The coupling of the quantity of connections with their convergence (i.e., level of social agreement) across speakers makes such expressions efficient in rehearsing belief and behavior within groups. Elliptical expressions combining density with consensus were known in ancient rhetoric as enthymemes. Following print, in 1513, a certain class of these expressions came to signify a set of agreements about a person, commodity, or cause, intended to create solidarity and, often, to mobilize action through repetitive use. These expressions, according to the OED, came to be known as *slogans*. During the industrial revolution, other classes of symbols combining density and consensus were rechristened according to metaphors drawn from the print trade. In 1798, for example, techniques were developed in printing to cast and copy whole pages in simple and fixed form without having to retypeset individual letters. The English name for this cast was *stereotype*. The French called it *cliché*. Walter Lippmann (1922) coined the modern definition of the stereotype as a set of hardened, unchanging, and simplified beliefs held by the mass audience. By 1892, clichés in English became names for stereotypical forms of language that had frozen (i.e., resisted in-depth processing) into everyday terms and so had lost their symbolic function (see Kaufer & Carley, 1993, for an extended discussion of linguistic forms dealing with print).[3]

Density and consensus, alone and in combination, can explain a wide assortment of symbolic expressions. However they do not exhaust the supply. There is a third primitive constitutive of symbolic expression, one that can combine with the primitives of density and consensus but can also stand on its own. We call this primitive conductivity. Conductivity is the capacity of an expression in context to carry (or trigger) information in a two-directional flow. Information flows in two directions when it both triggers and is triggered by other available information in the context. The basic intuition about (purely) conductive expressions is that they are both "noninitiating" and "nonterminating," meaning that they are neither starting nor stopping points for ideas being discussed but only gateways *from* and *into* these ideas. The importance of a word known only for its conductivity (and so lacking in density or consensus) is not the expression itself but rather the flow of ideas it keeps stimulating.

The primary example of such a purely conductive symbol is the *buzzword*. An example of a buzzword, for some audiences at least, is *mips and megs*. Few can elaborate what this term means with the precision of a computer engineer. Nonetheless, the term offers a fast gateway for purchasers of computers who want to gauge how their machine will run under certain conditions. The "meaning" of a buzzword lies not in its direct or immediate denotation but in the elaborations that everyday users have come to give it. Like the other primitives, conductivity is highly interactive with density and consensus. A conductive symbol filled out with density still acts as a gateway in the manner of a buzzword but also becomes a more sustainable topic of discussion in its own right (a placeholder for vogue ideas; see below). A conductive symbol functioning as a gateway to a surrounding consensus can turn an erstwhile hot idea (e.g., *Gorbachev*) into an emblem of the latest conventional wisdom ("old news") (for more formal definitions of placeholders and emblems, see below).

To sum up, the theoretical literature on symbolic expression, dominated by the study of the literary text, has tended to collapse the connectivity at issue in symbolic expression into the single primitive of density. Insofar as previous theorists have recognized the variability at the foundation of symbolic connectivity, the fact of this variability has yet to be incorporated at the level of explicit theory or even the precursor of theory, a systematic conceptualization. That is our task here.

Preliminary Assumptions

Let us call the concept or symbol whose connectivity in the network we are interested in measuring the *focal concept*. We will assume that focal concepts are included in larger messages, be they oral or written. We will further assume that the enveloping contexts of a focal concept (the larger message, the larger external situation) enables the focal concept to be elaborated into larger networks of meaning, called semantic networks. We will finally assume that the semantic network associated with the focal concept provides a semantic representation of this larger message in its situational context. A semantic network thus represents a focal concept's "connectivity in context." Technically, the semantic networks of concern in this discussion are known as valued networks. In a valued network, each node is a concept, each link a relation between two concepts, and the value on any link connecting concepts A and B an indication of the level of social agreement; more specifically, it is an indication of the number of language users (who share the context) who agree that a link in fact exists between concepts A and B. Finally, let us think of a "dimension" as a systematic way in which focal concepts can vary within and across semantic networks.

Three Dimensions of Connectivity in a Semantic Network

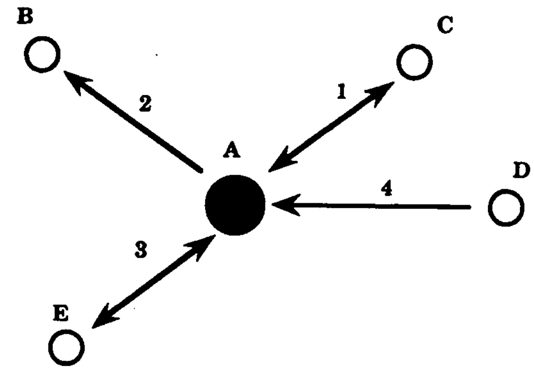
Let us now explore three dimensions along which we can analyze connectivity[4] relations in a semantic network: (1) density, (2) conductivity, and (3) consensus.

Density

One way—perhaps the most common way—we think of a focal concept as acquiring meaning and reach is through the number of links that connect it to other concepts. This is the familiar property of *density*.[5] In fact, one might argue that the greater the density of a concept the greater its breadth of meaning. The density of a focal concept is the number of concepts to which the focal concept is directly linked, regardless of the direction of the link. A direct link is a link that begins or ends at the focal concept—that is, the link does not pass through any intermediate concepts. In Figure 1, focal concept A is directly linked to four other concepts (B,C,D,E) and so has a density of four.

An analogy may further help to clarify density. Imagine a road map of the United States. The cities are the concepts, and the roads are the links between them. When one looks at a map, one sees that some cities, like Boston and New York, have many roads leading immediately in and out; whereas others, like Walsenburg, Colorado, have only one or two. The city whose intercity road

Figure 1
Illustration of focal concepts (A's) density, conductivity, and consensus. A's density is 4; A's conductivity is 9, and A's consensus is 2 when the threshold is 3.



density we would like to measure is the focal city (similar to the focal concept). The intercity road density of the focal city is measured by counting the sheer number of cities that are directly connected (i.e., without passing through another city) to the focal city by a road (regardless of whether the road is one-way, a freeway, or a dirt lane). This is the identical measure of density we shall be using when we say a focal concept is dense relative to others in a network.

Significantly, directionality is a not a factor in computing the density of a focal concept. Imagine now that the intercity roads on our roadmap are all one-way. That is, suppose that all roads either lead *out* from a city or *in* but not both (thus, these roads have a direction). Now suppose there is some city (A) with 10 direct routes out to 10 other cities, and 11 routes in from 11 other cities. Suppose also there is another city (B) with only 1 direct route out to another city and 20 direct routes in from 20 other cities. Cities A and B have the identical intercity road density (21) but nonetheless very different access patterns to and from different cities. A is almost equally easy to enter and leave, while B, like Rome, is easy to enter but hard to leave.

Were density the sole measure for indicating how a concept can stand out in a network as highly connected, we could stop here. Yet we can see from our intercity road map example that density is not sensitive to patterns of connectivity that depend on the direction of connections. Insofar as the direction of connections might be judged an important factor in determining connectivity overall, density offers no help at all. We need, it seems, in addition to density a measure of connectivity that is sensitive to the direction of connections. We have that measure in a second dimension, conductivity.

Conductivity

A second way to affect the meaning of a focal concept in a network is to alter the number of connected paths through it. A connected path exists just in case there are concepts that lead both *into* and *out from* the focal concept. These are two-way links. The number of paths leading through a focal concept increases as the number of "in-links" increases (assuming there are already some "out-links"), the number of out-links increases (assuming there are already some in-links), or both increase. With multiple links leading out and leading in, a highly conductive focal concept lies at the crossroads of many possible paths in the network. Such an increase in paths through the focal concept gives it an increasing "centrality"[6] in the network that density alone cannot guarantee. (As we will see more clearly below, a very dense focal concept can still be peripheral in the network if its links flow in only one direction.) Metaphorically, a focal concept lying on many paths acts as a highly conductive wire through the network—hence the name of this dimension.

The total conductivity of a focal concept is measured by multiplying the number of concepts directly linked into it by the number of concepts directly linked out from it. In Figure 1, for example, the focal concept A has a conductivity of 9 because 3 concepts (C,D,E) link into it and 3 (B,C,E) link out from it. Metaphorically, conductivity refers to the amount of information a focal concept can "carry"—like current over a wire. Notice that a focal concept has zero conductivity if it is a pure launching point (out-links but no in-links) or a pure sinkhole (in-links but no out-links). In such cases, a focal concept will be very dense but still, in a sense, a peripheral node in the network and not at all conductive. Note also that a focal concept acquires density and conductivity in a network at a very different rate. Density grows one concept at a time, additively. Conductivity grows faster than that, multiplicatively. Suppose, for example, to Figure 1 we add a bi-directional link from focal concept A to a new concept F. This simple change increases A's density by 1 (from 4 to 5), but it increases A's conductivity by 7, thereby almost doubling it (from 9 to 16).

Consensus

A third way to affect the meaning of a focal concept with a semantic network is to establish that many language users in the discourse context agree on the links and the concepts that flow in or out of it. The more individuals who share the links and concepts tied to the focal concept, the greater the social consensus. Consensus seems straightforward to measure when one can rely on a show of hands or answers to a survey question. But this typically occurs when one's interest is focused on a single belief or decision rather than an entire environment of language use. In the latter case, consensus is not limited to a single choice point but involves a large set of choices. As such it involves at least two factors—how broad the consensus is (how much is agreed to) and how strong the consensus is (how many people agree). Herein we define consensus relative to the meaning of the focal concept as the number of direct links that meet or exceed some threshold.[7]

To continue an earlier analogy, imagine now that the roads connecting cities differ in width. Some are dirt roads, others are multilane highways. The width is a function of the number of people who use the road. Cities with many multilane highways entering and leaving them (like Dallas) are in some sense more socially important than cities surrounded only by dirt roads, and the more multilane highways the more important the city. When the map in question is a semantic network (a network potentially connecting any two concepts), the width of the road (the link) is the number of people who agree that these two concepts are connected. We can compute consensus by surveying or sampling the agreements of language users about which concepts are connected to which on a pair-by-pair basis. The more highly consented to links to or from the focal concept, the higher its consensus.

In Figure 1, for example, consensus is understood as the number of links that meet or exceed some threshold of agreement for some population. Imagine, for example, a population of 5 language users. The numerical annotations on the links in Figure 1 represent the number of language users who agree to the link. Thus, Figure 1 says that of the 5 users, 2 share a directional link between A and B, 1 between A and C, 4 between A and D and 3 between A and E. If we assume a threshold of 3 (representing that the majority of the language users agree to the statement and hence agree that there is a relationship between the two concepts), we can say that focal concept A has a consensus score of 2 for this population, because A has 2 links that meet or exceed the threshold set at level 3. To continue our road analogy, consensus is measuring the number of cities to which the focal city is connected by highways with a number of lanes that meet or exceed the threshold.

There are many important issues surrounding consensus that go well beyond the scope of this article. We mention a few such issues here so as to clarify why the level of consensus surrounding a focal concept has been defined as the number of direct links that meet or exceed some threshold. First, let us consider the values on the semantic network. These values represent the extent of social agreement. Whether this is indicative of the number of individuals in the society who know that they agree, or the number who tacitly agree, or some other notion of agreement depends on the researcher and the research goals. For many research purposes a threshold can be located for which the researcher can argue that there is at least "this amount of" consensus to that link. While it is often difficult for individuals to judge whether and the extent to which a link (e.g., *smoking is harmful*) is agreed to in absolute terms, it is often easier for them to judge which links are and are not consented to relative to some threshold (e.g., smoking is agreed to be harmful if we set 60% of the American public as our threshold of agreement). A consequence is that the number of links can be counted for which there is a consensus that meets a threshold level.[8] The measure for consensus becomes comparable in this respect to the measures for density and conductivity.

The use of a threshold does not imply the absence of agreement for links that fall below it; it only implies the absence of "social knowledge." Social knowledge consists of that information that is more or less known by most individuals in the society. Social knowledge exists on a link-by-link basis and can be thought of as agreement at or above a certain threshold. Threshold-setting for social knowledge is important because the level of agreement required to achieve social knowledge may vary by context. The level of agreement required to attain social knowledge in physics may be far higher than that required to attain it in literary criticism. To be able to study social knowledge is either context, the researcher will want to be able to adjust the level of agreement that constitutes social knowledge, and this is exactly the flexibility that the use of a threshold accommodates. For additional discussion of social knowledge see Polanyi (1962) and Carley (1986b).

In the application to be discussed below, consensus is viewed as tacitly shared knowledge, and the values in the semantic networks are indicative of how many people think that a link exists between those two concepts. On this interpretation of consensus, individuals can agree (have a consensus) without being aware of it. As Polanyi (1962) notes, one way such tacit consensus can arise is if the individuals in question both went through the same experience (such as two women in different cities giving birth or a set of students sitting through the same lecture course though perhaps during different years). Whorf's shared culture (1956), Bar-Hillel's universal encyclopedia (1960), and Sowa's background knowledge (1984) are similar conceptions of consensus through tacitly shared knowledge. Alternatively, other researchers, using the same measure of consensus defined above, could measure the values on the links as the number of individuals who *realize* that they agree that there is a link between the two concepts. Such researchers would rely on a notion of an active, rather than a dormant, consensus. This difference in interpreting the nature of consensus would affect the interpretation of the results but not the definition of the consensus measure as the number of direct links that meet or exceed some threshold.

A Typology of Semantic Categories Formed by the Intersection of These Dimensions

Let us assume that focal concepts can be either "high" or "low" on the dimensions of density, conductivity, and consensus, depending upon whether their computed score for each dimension falls above or below the associated cutpoint value for that dimension. What value is used for the cutpoint depends on the researcher's intent. Possible values include, but are not limited to, the mean and median of all concepts' score for that dimension. The median might be used if the goal is to divide the sample in half. The mean might be used if the goal is to locate concepts that are "above average." Thus, for example, if the cutpoint is based on the mean and if the concepts in the network have a mean density score of 20, then a focal concept with a density score of 20 or more will be judged "high" in density (as it is higher than the average) and a density score lower than 20, "low" in density (as it is lower than the average). Regardless of how the cutpoint value is defined for the three dimensions, each focal concept can be characterized as either high or low on each dimension. Thus, we can define 8 ($=2^3$) semantic categories into which focal concepts can be classified. The cube in Figure 2 represents a typology of these categories.

Each category represents a different corner of the cube. Let us describe each of these categories in more detail. In doing so, we will assume that the cutpoint for density is 7, for conductivity 15, and for consensus 3. These numbers are chosen purely for the sake of illustration. In a real data set, as will be illustrated later in this article, the cutpoint might be chosen using a measure of central tendency such as the mean or median, or using a measure of extremes such as quartiles.

Ordinary Words

The easiest way to begin is with words that remain "ordinary" in their symbolic functioning, that have minimal connectivity in context and so have dictionary value but negligible value in structuring complex social inferences through their

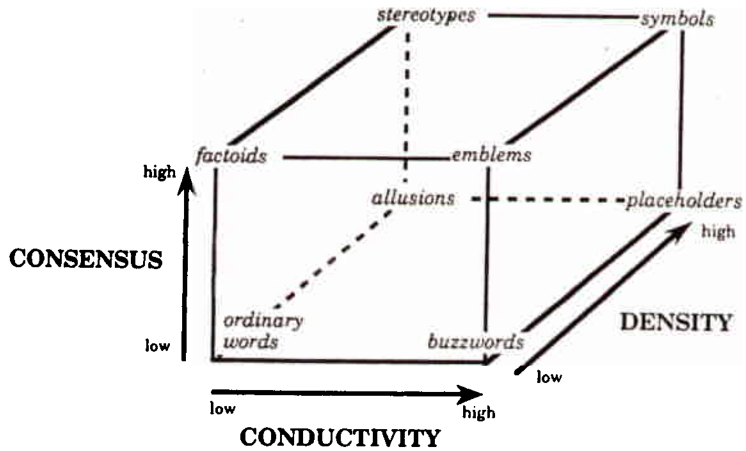


Figure 2
A typology of semantic categories into which concepts can be classified.

history of use. Imagine two people driving down a country road and one saying to the other, "There's a farm." This scenario is schematized in Figure 3. The contextual focus is the act of driving and pointing to the farm, and the focal concept is "farm." The arrows in Figure 3 indicate some internal elaborations going on in the speaker's mind about barn (e.g., an elaboration moving from the speaker's concept of self to *farm*—"I wish I had grown up on a farm"—and an elaboration moving from the *farm* to one of its properties—"that farm is large"). These elaborations, however, are few and unshared with the interlocutor. The moment of utterance will likely come and pass without symbolic marking.

Allusions

The likelihood of a focal concept becoming marked as such enhances as it increases in density. Imagine a communicative situation where the focal concept is very dense but still low in conductivity and consensus (schematized in Figure 4): an economic report that includes various references to the "deficit." The report includes many statements that are about the deficit (i.e., out-links from *deficit*; *deficit* as subject) as well as some couched in terms of the *deficit* (i.e., in-links into *deficit*; *deficit* as predicate). Recall that to be dense, a focal concept must have many direct links, regardless of their direction or bi-directionality. In the

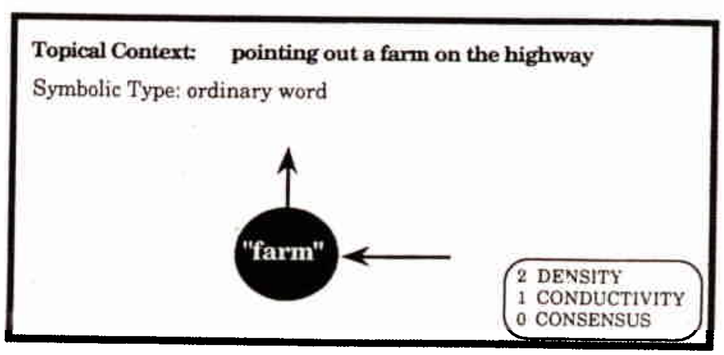


Figure 3
An example of a concept that is below average in density, conductivity, and consensus. Such words are "unremarkable" in their connectivity—hence "ordinary."

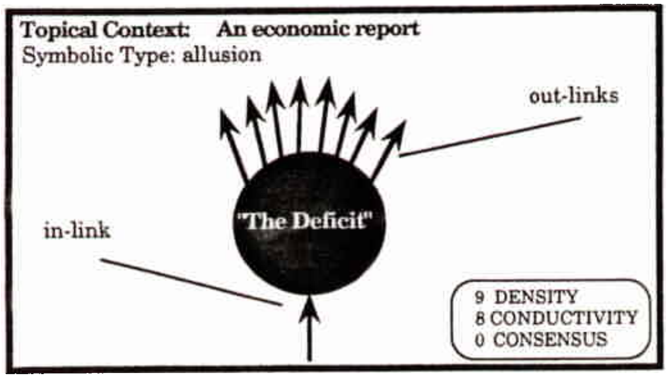


Figure 4
An example of a concept that is high in density but relatively low in conductivity and consensus.

report in question, we might say that the concept *deficit*, insofar as it is dense, functions as an allusion within the report. Allusions function as important preconditions for cultural communication. Lamenting the low cultural literacy of American students, E.D. Hirsch (1987) was in fact complaining that American students lack the requisite knowledge to "read" cultural concepts (e.g., *the Civil War, Waterloo, the New Deal*) allusively. Students, Hirsch observes, read them instead as ordinary, unconnected words, shorn of their sociohistorical meaning.

Buzzwords

Buzzwords emerge when ordinary words become above average in their conductivity but remain low in density and consensus. In the annual address the dean speaks of the need for the college to be *synergistic* (Figure 5). In the course of the address, the dean links four other themes (e.g., *user-friendly college, caring faculty, more computers* and *student aid*) to *synergy* and *synergy* to them. The audience agrees that there is a link between *user-friendly college* and *synergy*. Yet *synergy* remains underelaborated, only sparsely tied to other concepts. The other thematic words, by contrast, are all more dense than *synergy* and better

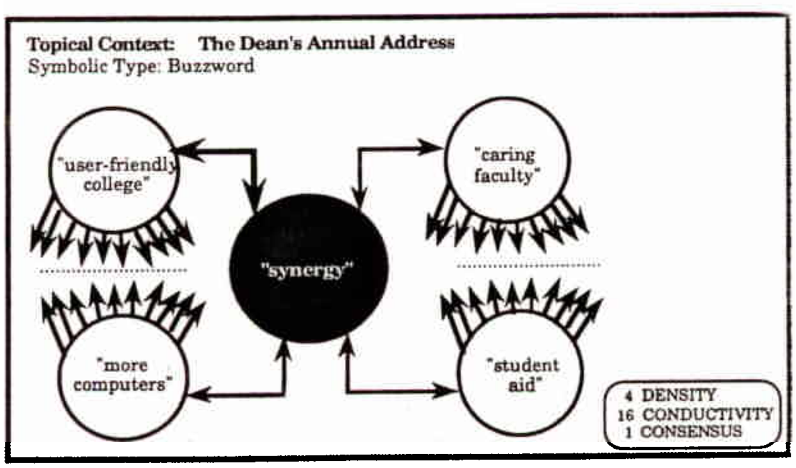


Figure 5
An example of a concept that is high in conductivity but relatively low in density and consensus—a characteristic of buzzwords.

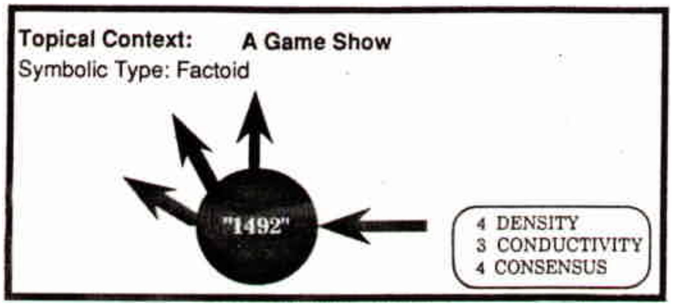


Figure 6
An example of a concept that is above average in consensus but relatively low in density and conductivity.

specified, each linked to 10 other concepts. Yet *synergy* is the only word falling on a high number of paths. The other thematic words have more out-links (10), but fewer in-links (1) than *synergy*, and so lie on only 10 paths. *Synergy* has 4 in-links and 4 out-links and so lies on 16 paths. Although less dense and defined than the other themes, *synergy* is the best concept to spark recognition of all the other themes—which is the kind of kindling spark a buzzword is supposed to ignite.

Factoids

Factoids emerge when ordinary words become above average in their consensus value while remaining low in density and conductivity (Figure 6). The focal concept *1492* in America seems to have elaborations that are highly consented to even though the concept is not above average in density or conductivity.

Placeholders

A focal concept that is high in both density and conductivity but low in consensus functions as a pregnant placeholder, or simply a placeholder (Figure 7). At a conference on communication, *communication* has the high conductivity of a buzzword and the high density of an allusion; yet there is, on average, little relative consensus about how to elaborate it.

Stereotypes

Focal concepts that are high on density and consensus but low in conductivity function as stereotypes (Figure 8). Stereotypes are typically allusions whose links to related concepts have become widely taken for granted over time, at least

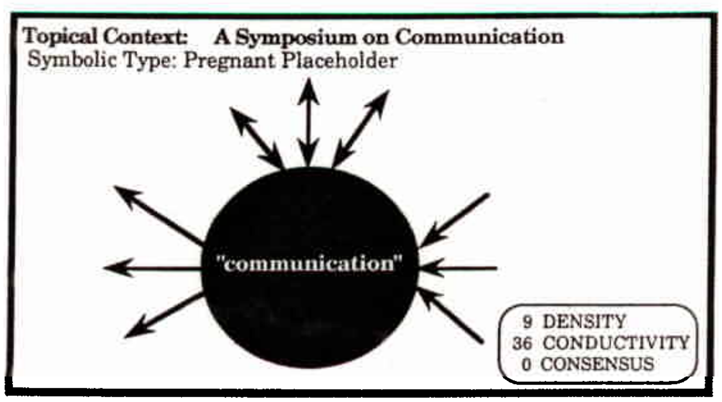


Figure 7
An example of a concept that is above average in density and conductivity and below average in consensus.

Table 1
Univariate Statistics for Each of the Three Networks

Dorm Semantic Network		
	Mean (std.dev.)	Range
Density	13.5 (13.1)	0–96
Conductivity	71.3 (214.3)	0–2268
Consensus	2.5 (4.2)	0–38
Thesaurus "Comedy" Semantic Network		
	Mean (std.dev.)	Range
Density	25.8 (24.8)	2–130
Conductivity	315.5 (1026.8)	0–1309
Consensus	4.2 (15.6)	0–5
Research Writing Classroom Semantic Network		
	Mean (std.dev.)	Range
Density	4.4 (7.3)	0–81
Conductivity	17.8 (63.6)	0–1640
Consensus	0.4 (1.4)	0–18

(Palmquist, 1990); (3) a thesaurus entry with *comedy* as the focal or index concept (original to this essay).

In the first environment, interviews from former and current residents of a dorm were used to extract a semantic network containing general social knowledge regarding life in that living group, particularly as it pertained to the resident tutor. The values of links were derived from estimates about the degree of agreement across the students pertaining to the existence of the link. This network was composed of 217 concepts.

In the second environment, the semantic network was extracted by interviewing students in a writing classroom on the topic "research writing" (the genre taught in the course) at the beginning, middle, and end of a semester. The concepts elicited from students and the relationships they reported between them were coded and combined into a single classroom network. The value of each link in the network represented the actual number of students (not an estimate) who associated the two concepts related by it. This network contained 107 concepts.

In the third environment, a semantic network was constructed from an electronic thesaurus and, specifically, from the thesaurus entry for the word "comedy."[9] The values of the links were the assumed level of social agreement about the *representativeness* of the second concept as a synonym for the first. (Judgments of representativeness were based on the ordering and visual layout of words in the thesaurus entry; see Appendix.) This network contains 310 concepts.

It is beyond the scope of this article to discuss how these semantic networks were extracted. Details of the extraction process followed for the dorm data can be found in Carley (1988), for the thesaurus data in the Appendix, and for the

Table 2
Connectivity for Selected Concepts in Each of the Three Networks

Dorm Semantic Network

Symbol Type	Density	Consensus	Conductivity	Concept
Ordinary	0	0	0	capable of doing the job
Ordinary	12	1	11	studious, scholarly, academic
Ordinary	8	1	12	attends hall meetings
Emblem	11	3	30	afraid to try, doubts own abilities
Placeholder	29	1	78	inspiring, unusual, interesting
Symbol	79	37	1554	personality
Symbol	70	16	78	friendly, gets along

Thesaurus "Comedy" Semantic Network

Symbol Type	Density	Consensus	Conductivity	Concept
Ordinary	0	0	0	thriller
Allusion	5	0	0	caricature
Factoid	4	1	3	theatrics
Emblem	3	1	2	esprit
Allusion	12	0	10	small talk
Stereotype	7	1	22	cut
Symbol	56	5	588	joke

Research Writing Classroom Semantic Network

Symbol Type	Density	Consensus	Conductivity	Concept
Ordinary	0	0	0	writer
Ordinary	4	0	4	communication
Factoid	4	2	0	final copy
Factoid	7	2	12	revise
Placeholder	9	0	20	card catalog
Symbol	28	6	195	insider
Symbol	81	18	1640	research writing

classroom data in Palmquist (1990) (see also Carley & Palmquist, 1992). The approach to semantic connectivity elaborated here is meaningful across different extraction procedures. Further, the measures of semantic connectivity discussed above can be calculated using the computer program CUBE (Carley, 1990) on any semantic network—regardless of the extraction method.[10]

The univariate statistics for each of the three networks are reported in Table 1. The values on each dimension for selected concepts in each network are reported in Table 2. These tables illustrate how dramatically these dimensions can vary across environments.

The three dimensions, density, consensus, and conductivity are logically distinct. Each captures a different aspect of connectivity. For each of the environments we examined we calculated the correlation between each pair of the three dimensions. The results are reported in Table 3. For the environments examined, these three dimensions are not co-linear and so are not only logically but empirically distinct.

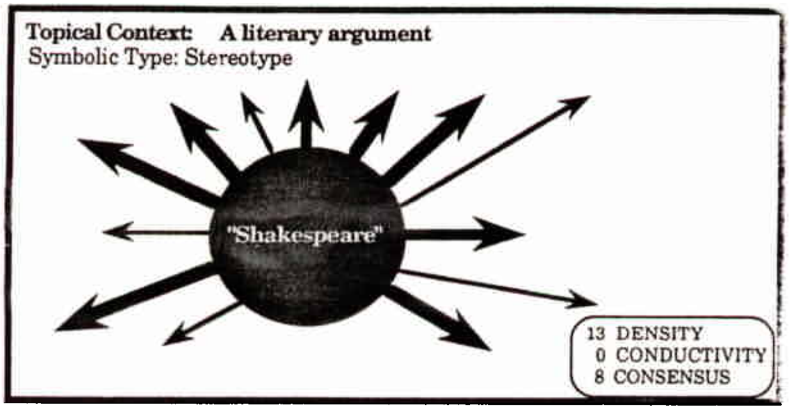


Figure 8
An example of a concept that is highly dense, highly consensual, and below average in conductivity.

within a group. The concept *Shakespeare*, arguably has turned into a stereotype among Renaissance literary scholars interested in building new arguments while relying on a wealth of taken-for-granted knowledge about the Bard.

Emblems

Focal concepts that are high in conductivity and consensus but low in density function as emblems. Emblems are islands of consensus that fall on paths sparsely but centrally linked to the topical focus. In issue contexts, antagonists and protagonists who are known for their argumentative stands (e.g., George Will for conservatism; Sam Donaldson for liberalism) function as emblems. The name of the late Senator Pepper has become an emblem for elderly rights. It may be that academic citations function as emblems (Figure 9). Employing a citation often implies that the meaning of the previous work is both agreed upon and a central conduit to the work being reported—though not the major preoccupation of the current work and so relatively low in density.

Symbols

Focal concepts that are high in density, conductivity, and consensus are expected to have the most impact of all rhetorical symbols. We label these standard symbols or simply symbols. Such concepts are also called terms of art or the

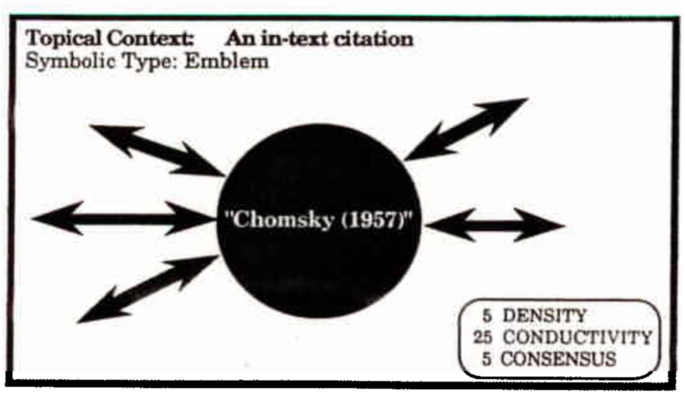


Figure 9
An example of a concept that is above average in conductivity and consensus but relatively low in density.

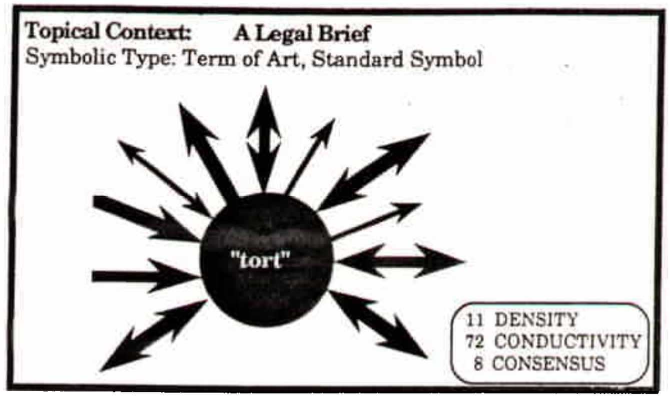


Figure 10
An example of a concept that is relatively high in consensus, density, and conductivity.

Cassirerian or Burkean symbols. Standard symbols should be more common among individuals who have enough shared history to be able to invoke a relatively well-defined "community" through their language. Standard symbols are most likely, in other words, to "induce cooperation in beings that by nature respond to symbols" (Burke, 1969, p. 43). In professional communities, standard symbols are most commonly called "terms of art" (Figure 10). In civic communities, these symbols are what most rhetorical theorists and critics have in mind when they refer to the "symbol" as term of communal invocation, a shibboleth or verbal logo for a group. Leach (1976) calls such symbols "standard symbols" and Henry Small (1978), in his study of citation in scientific specialties, calls them "concept symbols."

Properties of connectivity vary across contexts. The same concept may be below average in its connectivity in one context (e.g., *farm* in the example above) and above average in another (e.g., a politician recalling *life on the farm* to a group of farmers). Further, in analyzing the relative connectivity of the available concepts in a contextually primed network, one is observing a time-slice of an often changing sociohistorical and cultural context (Danowski, 1982, 1988). For example, the concept *Nixon* might have very low semantic connectivity for students coming to consciousness in the 1980s; might mark an emblem (of corruption) for students coming to consciousness in the late 1970s; a negative stereotype for students coming to consciousness in the early 1970s; and a negative symbol for students who protested the war from the start of the Nixon administration in the late 1960s.

Computing Connectivity Relations in Different Semantic Environments

Let us now consider how this typology can be applied to computing connectivity scores in different semantic environments. These environments (which were tracked by one of us or our associates) include (1) a college residence hall where students had to elect a new resident tutor (Carley, 1984, 1986a, b); (2) a writing classroom where students were being taught how to write a research paper

Table 3
R^2 of Dimensions for Each of the Three Networks

Semantic Network	Density-Consensus	Density-Conductivity	Consensus-Conductivity
Dorm	.692	.739	.870
Thesaurus "Comedy"	.441	.634	.761
Research Writing Classroom	.604	.811	.573

Making General Inferences About Connectivity Across Semantic Environments

On occasion, we want to make comparative inferences about connectivity across topical contexts. For example, in comparing scientific disciplines one might wish to test the proposition that disciplines associated with physical as opposed to social reality produce shorter articles in part because the language of such disciplines is less amorphous and there is greater consensus as to what words mean. Physical scientists, as a result, will not need to restate first principles as much as social scientists (Bazerman, 1988). We could begin to examine such an hypothesis by contrasting the connectivity of a sample of concepts used by members of those disciplines. Assume, for purposes of illustration, that "amorphous" operationally means low in density as well as consensus. Were this hypothesis to have merit, we might then expect that the more physical the discipline, the higher the proportion of concepts of high consensus (factoids, emblems, stereotypes, and symbols) and density (allusions, placeholders) and the lower proportion of low consensus or dense concepts (ordinary words, buzzwords, factoids). One could test such a hypothesis by analyzing the published articles of such disciplines. To take another example, in comparing cultures, one might want to test the proposition that the greater the number of technologies available to individuals in a society to perform a particular task (e.g., contact a friend in another city), the greater the potential number of concepts associated with the task and so the higher the density but the lower the consensus on the meaning of the task. This proposition could be tested by measuring the connectivity of task-related words in two or more cultures that vary in technological sophistication and then determining whether in the less sophisticated cultures the density of the words is lower but the consensus higher.

These are examples of hypotheses relying on comparative judgments of symbolic behavior. We cannot make these comparative judgments from raw data alone. Notice that in the tutor selection environment, the concept *attends hall meetings* is low in density even though it has a higher raw density score than the high density concept *caricature* in the thesaurus environment. We cannot make comparative inferences from raw data without standardizing these scores and to standardize these scores, we must work with the maximum connectivity of symbols in their semantic environments.

Computing Maximum Connectivity

In this section we define measures of maximum connectivity for density, conductivity, and consensus respectively. Within a specific environment of N concepts, the maximum density of any concept is $N - 1$. The standardized density for each concept can be computed by dividing its observed density by the maximum density ($N - 1$) and multiplying by 100. This provides a measure of the percentage of maximum density within that environment that the concept in question attains. The maximum conductivity of a concept within a specific environment is achieved when that concept intersects a path between each pair of other concepts in that network. In a network with N concepts, a concept has a maximum conductivity of $(N - 1)^2$ (from all $N - 1$ other concepts to all $N - 1$ other concepts). The standardized conductivity for each concept can be computed by dividing its observed conductivity by the maximum conductivity and multiplying by 100. The maximum consensus of a concept within a specific environment is achieved when it attains a link to all other concepts in the environment ($N - 1$) that has met or exceeded some threshold of agreement among members of that discourse community. The standardized consensus for a concept is obtained by dividing its observed consensus by the maximum consensus and then multiplying by 100. Such standardized measures allow the researcher to make comparative inferences about density, conductivity, and consensus across topical contexts.[11]

Using these standardized measures we now examine and contrast the three topical contexts on each dimension. For each dimension we begin by suggesting a theoretical reason that one might expect differences in semantic connectivity relations across contexts. These are not the only hypotheses that could be offered but they serve to illustrate the kinds of propositions that one can explore by contrasting semantic connectivity across topical contexts. Our purpose is not to offer confirmations of the hypotheses we propose but only to illustrate the kinds of hypotheses that can evolve and be explored based on the notion of semantic connectivity. At the very least, our discussion should help the reader understand how differences across topical contexts can also reflect the way language itself is differently constituted across contexts.

Comparing Semantic Density Across Topical Contexts

Density is associated with the breadth and elaborateness of meaning associated with a concept, and one might expect the distribution of density to vary across communication environments. There is, for example, some reason to believe that density would be especially high in educational contexts (at the end of a term of study) because in these contexts teachers are trying to help students acquire terms of art. To acquire a term of art, a student must learn to share the breadth and elaborateness of the teacher's meanings. The purpose of education, of course, does not stop with density. The teacher also wants students to converge on the individual links of the network associated with the subject matter being learned. But much learning goes on even in the absence of such consensus building. And much educational theory stresses the importance of students being able to elaborate the subject matter on their own terms, making connections to ideas not previously thought through or anticipated by the teacher. This suggests that "putting ideas together" in dense relations is an invariant aspect of educational contexts, regardless of whether the connections produced are private (allu-

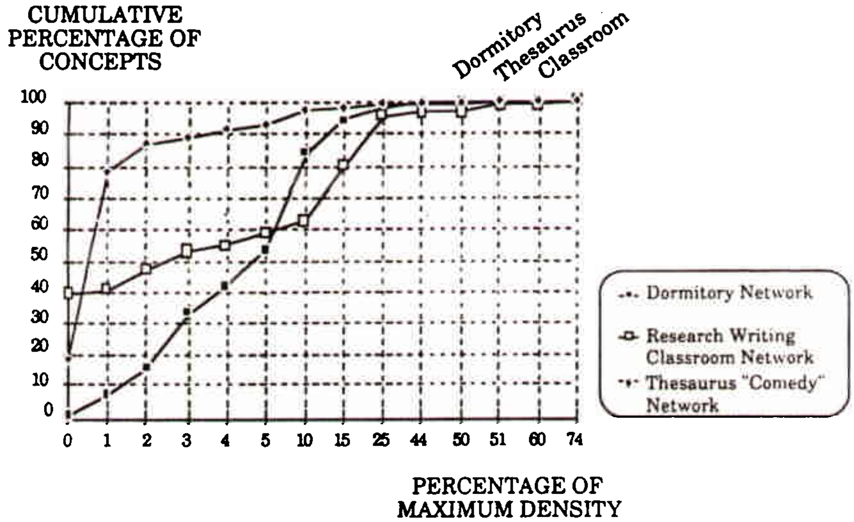
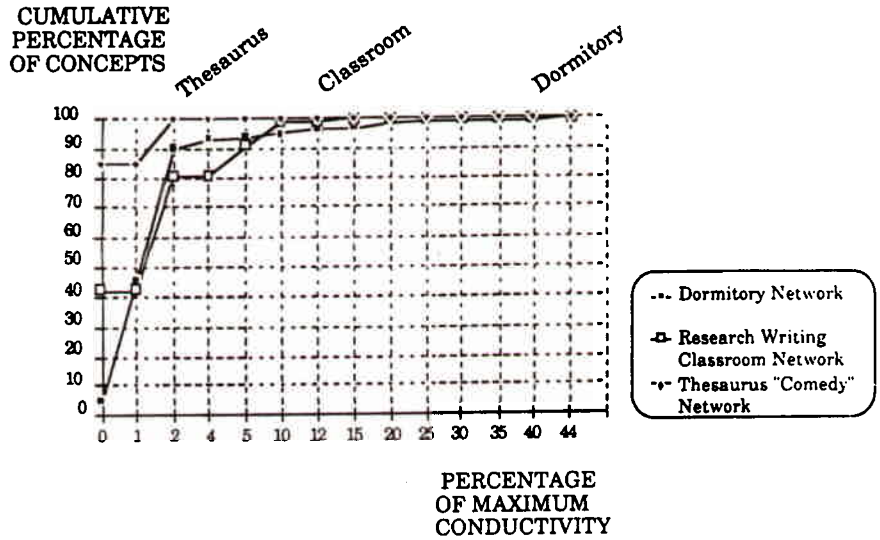


Figure 11 **Comparative density of language in dormitory, research writing classroom, and thesaurus.**

sions), drawn from the students' common experiences outside the classroom (stereotypes), or directly supervised by the teacher (terms of art).

It is perhaps then not altogether surprising that of our three contexts, we found that the concepts in the most purely educational of our contexts, the classroom, are, on average, the most dense. It is in this environment that a concept comes closest to reaching its maximum density. The concept *research writing* has links to 81 of a possible 106 other concepts in the classroom network, which gives it a standardized density of 76.4%. Nonetheless, we should keep in mind that, despite the greater density of concepts in the educational context, the majority of concepts *across all our contexts* are very low in density. Across all topical contexts, for example, well over 95% of the concepts in our corpi have less than 25% density. Furthermore, the concepts in the thesaurus corpus are noticeably less dense than the concepts in the other two contexts. For example, only 40% of the concepts in the classroom network and less than 10% of the concepts in the dorm network have a density less than 1%. Yet a full 80% of the concepts in the thesaurus network have a density of less than 1%. Conversely, approximately 50% of the concepts in the classroom network have a density greater than 3% and about 50% of the concepts in the dorm network have a density greater than 5%. Yet only 20% of the concepts in the thesaurus network have a density greater than 1%. Whether the lower density of the thesaurus environment reflects the fact that the thesaurus is not a "living" semantic environment or simply comes as an artifact of our coding decisions (see Appendix) remains an open question and one beyond our scope here. In Figure 11 we have plotted, for each topical context, the cumulative percentage of concepts by standardized density. To aid in the interpretation of Figure 11 we use the names "Dormitory," "Thesaurus," and "Classroom" to indicate the point at which 100% of the words in the associated context have been plotted. By examining these points we see that while the dorm and the thesaurus have no concepts with a density greater than 51%, the classroom has several.

Figure 12
Comparative conductivity of language in dormitory, research writing classroom, and thesaurus environment.



Comparing Semantic Conductivity Across Topical Contexts

Conductivity is associated with the capacity of a concept to link together multiple conceptual domains by providing gateways from one domain to others. In a classroom focused on concept learning and the rehearsal of these concepts in skilled practice (e.g., research writing), one would expect concepts to become increasingly elaborated, thickening the connections within them, rather than increasingly conductive, spiraling out like thin wires to other remote domains. Decision making, however, provides a context very different from that of focused concept- and skills-based education. In a decision making and particularly an election environment (e.g., electing a resident tutor), the voters need to consider a candidate across a spectrum of not entirely commensurable dimensions. They must consider social dimensions such as how well each candidate gets along with other students; they must consider intellectual dimensions such as the kind of study atmosphere a candidate can create for other students on the floor; they must consider moral dimensions such as the honesty and integrity of the candidate, particularly as he or she must deal with situations that raise moral dilemmas. The meanings evolving about a candidate during the course of an election are likely to remain highly conductive, and this conductivity should reflect itself in the language of the electorate talking about their decision making. In short, there should be a higher proportion of buzzwords, placeholders, emblems, and standard symbols in this context than the others.

In consonance with this hypothesis, we found that concepts in the dorm network, the only network of the three contexts associated with decision making and voting, are, on average, more conductive than the concepts in the other networks. In the dorm network, a concept comes closest to reaching its maximum value. That concept is *interacts with students*, which has a standardized conductivity of 44.4% and falls on 2268 of the possible 46,656 (i.e., 216^2

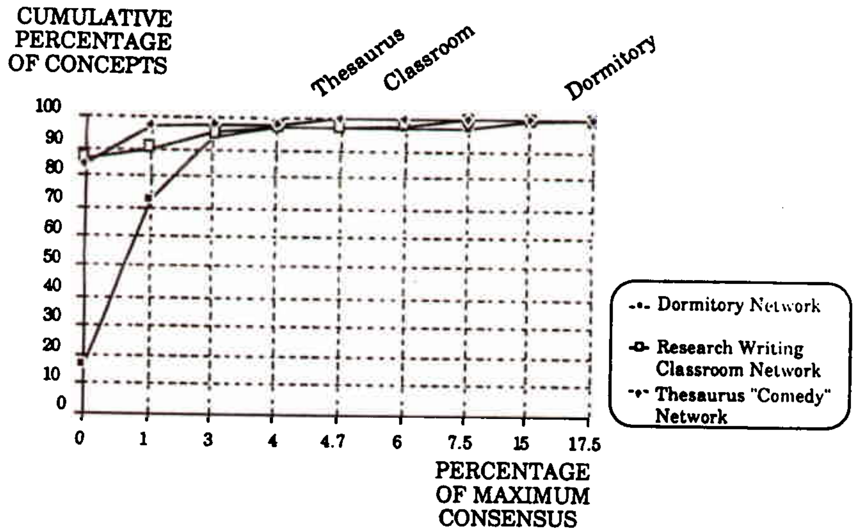


Figure 13
Comparative consensus of language in dormitory, research writing classroom, and thesaurus environment.

paths. In Figure 12, we can see that no concept in the classroom network has a conductivity exceeding 15% of the maximum; yet the dorm network has several such concepts. We need, nonetheless, to issue the same qualification for conductivity we issued for density: the majority of concepts across all three contexts are very low in conductivity. For example, 90% of the concepts have 5% conductivity or less. We can further note a second observation for conductivity that we noted before for density. The thesaurus environment seems less connected overall. Concepts in the thesaurus network are less conductive than either of the other "live" linguistic environments. Approximately 45% of the concepts in the classroom or dorm network fall within 1% of maximum conductivity or lower. Yet 85% of the concepts in the thesaurus context fall within this range. Approximately 20% of the concepts in the classroom network and approximately 10% of the concepts in the dorm network have a conductivity greater than 2%. Yet no concepts in the thesaurus network have a conductivity greater than 2%. Contrasting Figures 11 and 12, we see that although the classroom network contains concepts that are more dense than the dorm network, the dorm network contains the largest number of concepts with the highest maximum conductivity.

Comparing Semantic Consensus Across Topical Contexts

Consensus involves *how many* of the elaborations of specific concepts are shared, either tacitly or explicitly. One would expect this level of sharing to increase in small and informal social groups where individuals can interact face to face across a range of experiences. This expectation leads to the hypothesis that consensus should be highest in the dorm context. The plausibility of this hypothesis, moreover, is at least indicated by our comparative analysis of topical contexts. The dorm network has the highest average consensus and includes the concept that comes closest to its maximum consensus value. This concept is *interacts with students*, which has a standardized consensus of 17.5% and 38 highly consented to links out of a possible of 216. The concept *interacts with students*, we saw above, was also the most conductive term in the dorm network. Its high consensus and conductivity indicates that *interacts with students* had at the very least become, among the dorm residents, an emblem (i.e., very high in conductivity and consensus) of the role the successful candidate would have to play. As with density and conductivity, the majority of concepts in all three contexts have below average consensus. For all contexts, 95% of concepts attain at most 3% of their maximum consensus. In Figure 13 we have plotted the cumulative percentage of concepts against the standardized consensus.

As Figures 11 through 13 illustrate, distinct topical contexts tend to have distinct patterns of connectivity. These three studies of connectivity in topical contexts suggest that concept learning environments facilitate density relations and decision-making environments in small intense settings facilitate conductive and consensual relations. Moreover, concepts in "live" linguistic environments seem to be more dense, conductive, and consensual than are concepts in non-"live" environments, like that of the thesaurus. Further studies are needed to determine whether any of these hypotheses have merit, and we should be especially cautious to draw any inferences about data extracted from a thesaurus since the logic underlying the semantic organization of a thesaural environment is still very much an open issue (see below). Nevertheless, these three contexts present an intriguing picture of language in which we can glimpse the semantic connectivity of the language varying as a function of the sociohistorical environment in which it is couched.

Variations in Connectivity for the Same Concept in Different Topical Contexts

Since these three topical contexts contain very different words, the differences we have been able to discern in connectivity across them may arise in part from the differences in the words themselves and their interactive potential. As a counter to this bias, let us now restrict ourselves to the same set of words and see whether we can still see the effect of context influencing connectivity relations. Imagine a drama critic wanting to describe two plays, both concerning an emperor's life, but of different genres. One play is a drama; the other, a comedy. The critic will need to think about the audience for the critique as well as the type of play in order to know how to present the criticism. The same word (e.g., *misfortune*) intended to be elaborated in the context of one genre might be elaborated very differently when read in the context of the other. Let us now explore the extent to which different connectivity relations can be induced over the same set of words through shifts in genre expectations alone.

To examine this proposition we constructed the semantic network for the concept *drama*.[12] This network contains exactly the same 310 concepts that the comedy network contains. The difference between these networks is simply the focal concept at their center. The drama network spirals outward from the concept

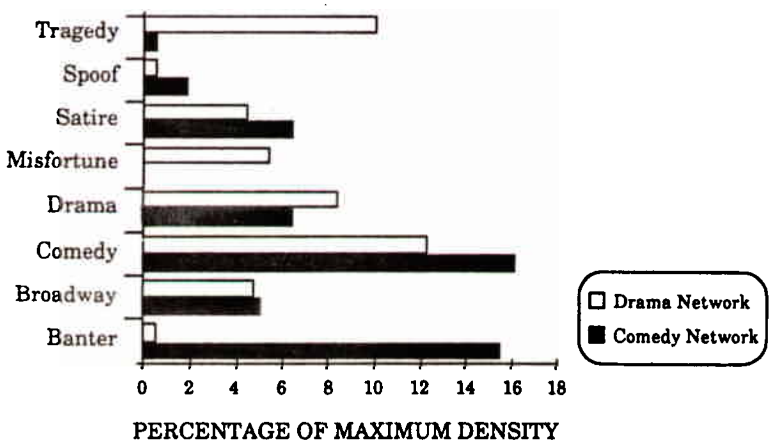


Figure 14a
Comparative density for selected concepts.

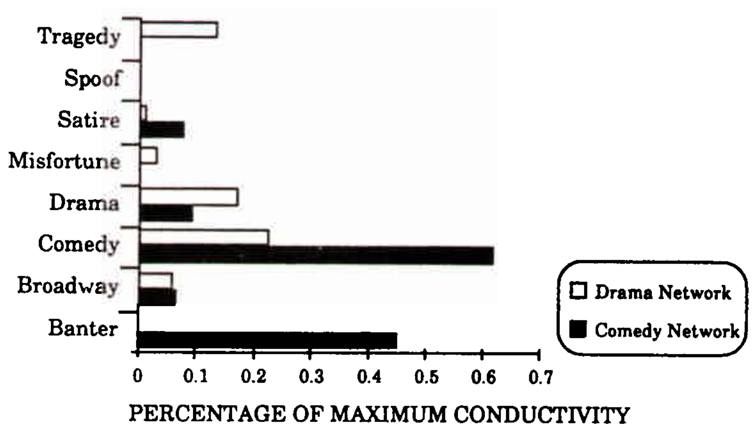


Figure 14b
Comparative conductivity for selected concepts.

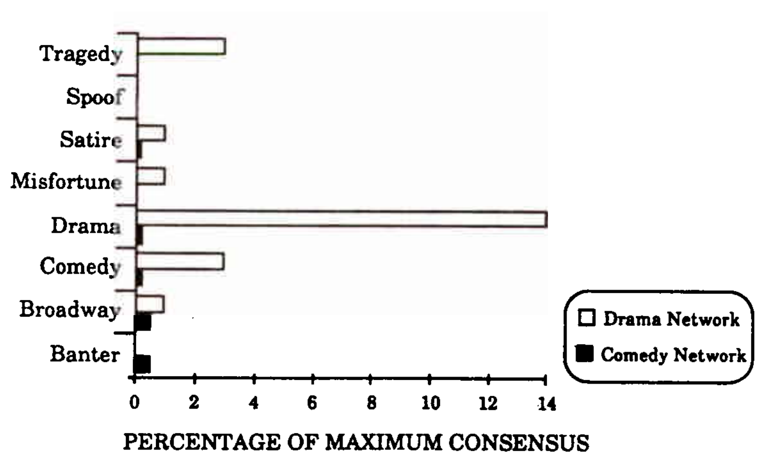


Figure 14c
Comparative consensus for selected concepts.

drama; the comedy network, from the concept, *comedy*. The networks contain the identical words but each network imposes, through a different projection into the network, an entirely different set of connectivity relationships. Let us now discuss the differences in connectivity relations that followed from the projection of each genre and the possible implications of these differences.

Overall, the *drama* network has lower connectivity than does the *comedy* network. For the *drama* network the average density is 11.1 (compared with 25.8 for *comedy*), the average consensus is 2.2 (compared with 4.2 for *comedy*), and the average conductivity is 80.4 (compared with 315.5 for *comedy*). One (by no means definitive) interpretation of these numbers is that, for the American language user represented by this thesaurus, comedy is a richer and more developed conceptual structure than drama. The American audience, and perhaps the critic, may have more elaborate mental models of comedies than dramas. Whether this observation is reflective of American audiences is not at all obvious and merits further investigation. But if it were the case, it would suggest that critics writing about comedy would already share with the American audience a great deal more connected knowledge about comedy than about drama. It would further suggest that critics writing about drama need to choose a smaller audience with a more elite understanding of the subject or need to educate the larger audience about the subject before proceeding to discuss it. Although the overall *drama* network is less connective than the *comedy* network there are still concepts within the *drama* network that have higher connectivity than they do within the *comedy* network. Figures 14a, b, and c illustrate the comparative density, conductivity, and consensus of 8 concepts for both the *comedy* and the *drama* networks.

The concepts *tragedy, misfortune*, and *drama* show consistently greater connectivity in the *drama* network than under *comedy. Comedy* and *banter*, by contrast, have consistently greater connectivity in the *comedy* network. *Broadway* and *satire* have a mixed profile—being slightly denser and more conductive within the *comedy* network but having a slightly greater consensus in the *drama* network. And *spoof* is not a good conduit or an object of consensus in either network, though *spoof*, as one might imagine, is denser in the *comedy* than the *drama* network. This suggests—an empirical hypothesis worth pursuing—that in writing about comedy, *spoof* is used mostly as an allusion, something that audiences of comedy readily recognize and can elaborate broadly, though in different ways. And *spoof* is almost never used as a gateway or word of transit through which different facets of comedy are discussed. Overall, as one might expect, concepts with greater comparative connectivity (density, conductivity, consensus) within comedy than within drama are *banter, buffoonery, burlesque, crack, farce, funniness, humor, jest, joke, lampoon, parody, quip, sarcasm, satire, spoof, tease, wisecrack*, and *wit*. Words with enhanced connectivity within drama are *dilemma, disaster, distress, enigma, lyric, melodrama, misfortune, mystery, quandary, riddle, secret, setback, soap opera, suffering, theatrics, thriller*, and *tragedy*. Those concepts that have basically the same connectivity in both networks include *playhouse* and *vaudeville*.

Illustrative Application Domains for Semantic Connectivity

We have illustrated that different sociolinguistic contexts vary in their semantic connectivity profiles. The use of these measures, however, goes well beyond demonstrating that concepts vary across contexts in meaning, likelihood of use, and social acceptance. Rather, by using these measures researchers can begin to address a wide variety of issues concerning the role of language in argument, persuasion, education, and other cultural contexts. To illustrate this point, we now describe some of the environments where we and our colleagues have put these measures of semantic connectivity to use.

Analyzing Argument Discourse

Many interesting problems of misunderstanding in political argumentation seem to arise from the failure or unwillingness of the participants to elaborate the implicit connections of highly connected words. This may be especially true for stereotypes, words whose connections are dense and agreed upon by a reference group but not highly conductive (and so relegated to the implicit background of argument). In many argumentative contexts, speakers often seem to talk past one another because the words that function for them (and their external reference groups) as stereotypes are received by their opponents as vague and underspecified allusions. This was demonstrated for the case of affirmative action and university hiring (Kaufer & Carley, in press). In that case, proponents and opponents proceeded from different historical stereotypes of the focal concept *university*. Opponents of affirmative action in university hiring proceeded from a medieval view of the university as an institution set apart from others, a sequestered haven from politics, tying it with symbols like *scholarship, merit, truth*, and *disinterest*. Proponents proceeded from the post-Marxian stereotype of the university as the site of political domination and power, elaborating it with symbols that make the university no different in kind from any other political institution. From each camp's perspective, the concept *university* has a dense, highly consented-to (i.e., stereotypical) meaning, but across the two camps, the social meanings remain unreconcilable networks of allusions. Kaufer and Carley discuss how different symbols from the typology discussed herein have predictable functions in political argument.

Decision Making and Voting

Carley (1984, 1986a) assessed the relationship between shared knowledge and the interaction among a group of college residents who were in the process of deciding on a dorm tutor. Listening to and reading histories of the hall and conducting interviews of former and current residents, she identified a corpus of 217 concepts that had some sociohistorical meaning for hall members relative to what they were currently looking for in a resident tutor. Using the typology above and some computational methods for connectivity (discussed below), she found that none of the concepts were exceptionally high on all three dimensions, and suggested that the only concepts that would take on such a characteristic would be concepts denoting social roles. Further, she found that the vast majority of the concepts analyzed were low on all three dimensions, were simply ordinary words. Two of the concepts that stood out, *friendly* and *fits in with the hall*, were high in density and conductivity but not consensus. These concepts served as placeholders and were used to bridge entrenched differences between two relatively self-contained cliques living on the hall. One group, referred to as the "heads," were very sociable, had long hair, liked loud music, and did drugs. The second group, known as "gnerds" (by that spelling), were conservative engineers who were into academics and quiet individual study. To win an election as tutor, a candidate had to build a coalition across these groups. Carley noted that, as a result of living together, the residents developed a high prohibition against any candidate who would "cause conflict" between the groups, who would not "fit in." She found that messages that linked a candidate positively with these two placeholders enhanced that candidate's chances of winning the election. She was also able to document the elimination of certain candidates whose behavior (e.g., taking drugs) and the negative talk pursuant to it destroyed their image as coalition builders. This work did not present a full theory of persuasion. Nevertheless it did suggest that messages that employed concepts that stood out on one of the three dimensions of connectivity did effect social change and the decision process.

Classroom Learning

In a classroom setting, a teacher utters many concepts that, from the teacher's knowledge scheme, are professional terms of art. They are highly conductive terms, densely linked to other symbols, and within the teacher's field have a highly consented to meaning. They are, in the taxonomy presented, symbols. While the teacher may utter such symbols as terms of art, that is not the form in which students typically receive them. Students are not part of the instructor's field. Concepts that are established factoids, emblems, stereotypes, and symbols in the teachers' field are, for students (at least at first), only ordinary words. Students typically pick up the meaning of such concepts in a piecemeal way and only after repeated interaction with them can they begin to fashion them into a complex internal network. Whether the networks students build resemble the teachers' is a central issue of educational evaluation. We can begin to address this issue by examining whether students, in a course of study, learn a set of concepts and relations that fall in the same section of the cube (Figure 2) as the concepts and relations understood by the teacher. Employing the approach to semantic connectivity described in this article, Palmquist (1990) tracked students in a class designed to teach the research paper. He was particularly interested in studying the network each student built over time around the topical term of art *research writing*.[13] He found that over the course of the semester many of the concepts that were established symbols or terms of art for the instructor increased in density and consensus for students. Concepts that began for students as ordinary words increased in their connectivity over the semester and moved into the realm of allusions, stereotypes, placeholders, and symbols. In addition, Palmquist found that when the teacher tried to teach research writing using a procedural model (i.e., "First you do this, then . . . ") students were more likely to learn, and hence share, elaborations about *research writing* (increasing the classroom consensus for *research writing*).[14] For example, almost all students learned to elaborate *research writing* with *topic*. Interviews revealed that they had proceduralized this elaboration (i.e., "First you find a *topic* . . . "). While this work clearly does not provide a full theory of education and lexical development, the measures and typology suggested here have promise as input to such theories.

Lexical Choice

Imagine a seasoned writer trying to express a situation in which Mary has said something to John for comedic effect. Assume for the moment that the writer is currently at a loss to translate the "comedy" idea into words. The writer, being experienced, can call upon a wealth of discrimination (i.e., dictionary) knowledge to reduce the space of hundreds of possible words to a few dozen. Let us assume Mary did not make an attempt to mimic John's persona in the act of speaking—which, the writer knows, eliminates *mimic, parody, takeoff, spoof, lampoon*, and *mock* from the choice set. The writer also knows that Mary's humor is pointed and sharp edged—intended to have a quick and short effect on John rather than a longer, sustained one. The intent thus cannot be a *roast*, which is drawn out and slyly respectful, or simple *scorn, derision*, or a *slur*, which are meant to have longer-term effects and betray (especially *slur*) the speaker's chronic lack of respect for the victim. This kind of "elimination by aspects" characterizes the work accomplished by discrimination knowledge. In doing this elimination by aspects the writer slides through the semantic network locating words through words with which they are connected. Concepts high in conductivity, buzzwords, emblems, placeholders, and symbols aid this search.

But notice that once the writer's discrimination knowledge has had a chance to kick in, the writer is still a long way from a single optimum choice. There usually comes a point when the meanings, sliced further, become finer than the granularity of the writer's intentions. For example, the choice between *barb, crack, swipe, knock, needle, cut, dig, jab*, and *rib* depends almost entirely on the visual metaphor the writer might want to attribute to Mary's comedic effort. The writer makes a choice between *jeer* and *sneer* by deciding whether the descriptive focus of Mary's verbal assault should center on the tone of her voice or the geometry of her mouth. But the writer may well be indifferent to either of these foci in describing Mary's actions—and so would have no preference between these choices.

From the standpoint of discrimination knowledge, selection should now be arbitrary. But it is not. Given a choice among *barb, dig, sneer, jab, swipe, needle, rib, jeer, gibe, tease*, and *quip*, informants show a preference for choices like *tease* and *quip* over choices like *sneer* and *swipe*. We can explain this preference by noting the very stratified positions of prominence that *tease* and *quip* and *sneer* and *swipe* play in a semantic network around *comedy*. Using a thesaurus network for *comedy*, we calculated that *tease* is denser than *sneer* and *swipe* by a factor of 6 and *quip* is denser than *sneer* and *swipe* by a factor of 12. Furthermore, *sneer* and *swipe* have no conductivity whatever within a thesaural network for *comedy*; while *tease* falls on 22 central paths and quip, 144. *Tease* and *quip* are more like placeholders while *sneer* and *swipe* are somewhere between ordinary words and allusions. The higher conductivity of placeholders than allusions makes it easier for the author to locate them in the first place. In addition, the higher density and higher conductivity of *tease* and *quip* make them better lexical choices as they are more likely to be understood by the readers and more easily tied to other ideas. Measures of semantic connectivity as developed herein, it would seem, promise to have explanatory power in a theory of lexical choice. Such a theory, although not developed herein, would be a next logical step.

A thesaurus is an organized system of language and an external aid designed to augment the internal system of lexical choice. Roget in the mid-19th century spent much of his lifetime cataloguing synonyms into semantic families and many alternative organizations (some alphabetical) have been widely used since. There has, however, yet to be a theory for constructing thesauri to optimize them for facilitating lexical choice. One would want the words in semantics families that are most highly connected (and so the easiest from which to search the whole network) to lie in the most typographically prominent positions. The electronic thesaurus we used for the *comedy* and *drama* network sometimes obeys this constraint; sometimes does not. Given the techniques described here, it is conceivable to think of more empirically derived thesauri, thesauri automatically constructed from the semantic networks extracted from the empirical judgments of language users.

Conclusion

We have tried to suggest in this article that while the literature is rich with techniques for generating semantic networks from linguistic data, conceptual devices for analyzing these networks have not been as forthcoming. We have described a more comprehensive perspective for analyzing such networks based, in part, on an extension of the theoretical literature on the symbol. We have applied these measures to semantic networks drawn from four different contexts. These applications illustrate that sociolinguistic contexts differ in their semantic connectivity profile and that the same words in different contexts may have very different profiles. Using such measures researchers can begin to address questions of language usage and change across discourse communities at the semantic level. Moreover, as we have suggested, these dimensions and the associated typology form a useful conceptual device that may be instrumental in developing and analyzing theories of consensus formation, argument, persuasion, lexical development, and lexical choice.

Appendix

In creating both the *comedy* and the *drama* semantic networks we used an electronic thesaurus (Word-Finder[R]). We now describe this process for *comedy*. The reader should keep in mind that the exact same process was used to generate the *drama* network. We constructed a set of concepts by locating all concepts that occurred immediately in the entry for *comedy* (i.e., level 1 concepts) as well as all the concepts contained in each entry for all level 1 concepts (level 2 concepts).

For example, for the concept *comedy* we first consulted the entry for *comedy* and made records of all the concepts (e.g., *musical, humor*, and *musical-comedy*) in that entry and placed a link between *comedy* and each concept. An entry is structured as follows:

comedy **musical musical-comedy theater;**
humor laughter banter

Or in more generic terms:

index **classifier associate associate;**
classifier associate associate

In this entry, *comedy* is the index concept, while *musical* and *humor* function as the primary classifier concepts. The concepts following each classifier are simple associates on the classifier concept. In a thesaurus entry there is an implied link between each index concept and each of its classifiers and associates. Further there is an implied link between each classifier and its associates. In order to estimate consensus in this network, we posited that the strongest (i.e., most agreed upon) associations exist between the index concept and the classifier (*comedy* → *musical*; *comedy* → *humor*). The next strongest associations obtain between the classifier and the associates (*musical* → *musical-comedy*; *musical* → *broadway*; *humor* → laughter; *humor* → *banter*). The least strongest, finally, hold between the index concept and any of the embedded associates (e.g., *comedy* → *theater*; *comedy* → *banter*). Second, we consulted each of the concepts appearing in the entry for *comedy*, and treating each of these concepts as an index concept also coded the entry for each of these concepts using the scheme outlined above.
Combined, the *comedy* and *drama* networks contain 310 concepts. For the *comedy* network there are a total of 750 links, and for the *drama* network there are only 359 links.

Authors

Kathleen M. Carley *is an associate professor in the Department of Social and Decision Sciences and* David S. Kaufer *is a professor in the Department of English, both at Carnegie-Mellon University, Pittsburgh, PA 15213. The order of authorship is alphabetical.*

Notes

[1] Below we distinguish symbols and ordinary words more precisely by making "ordinary word" a category in a typology of symbols. Throughout this article, when we make reference to a symbol, concept, or word generically, we are referring to an "ideational kernel"—a single idea, totally bereft of meaning except as it is connected to other concepts (Carley 1986a, b, 1988).

[2] These techniques vary in how they locate relationships and consequently what they admit as being a relationship (see Carley, in press, for more discussion).

[3] The effect of this hardening for cliché was to turn potentially symbolic words into ordinary, unremarkable ones. While stereotypes continue to function as enthymemes, clichés do not. In our classification below, clichés fall into the category of ordinary words.

[4] We use the term connectivity in order to emphasize the concern with the connections between concepts that provide meaning to it. The measures we describe are, from a social network perspective, measures of a node's centrality in a network. For additional discussions of meaning and a concept's position in a concept network see Carley (1988) and Carley and Palmquist (1992).

[5] Density, divided by maximum density, is equivalent to what has been termed the simplest measure of centrality (see Knoke & Kuklinski, 1982).

[6] A large number of measures of centrality has been forwarded by researchers in the social network community. Most such measures are defined only for nondirected graphs (binary symmetric data). As we indicated in the last note, our density measure, divided by the maximum possible density, is equivalent to the simplest centrality measure. Because the social network community is mainly interested in nondirected graphs, the conductivity measure we have proposed is a less common measure. Many alternate measures of centrality have been developed within this scientific community, and one could recast all such measures of centrality in terms of their meaning in a semantic network (see Freeman, 1979, for a review of such measures). Such a discussion, however, is beyond the scope of this essay.

[7] This measure of consensus is also a measure of node's centrality in a valued network. Because the social network community is mainly interested in valueless networks (graphs that do represent the strength of a particular link), our measure of consensus is not typically used. Nonetheless, both Danowski and Carley have noted the importance of measuring the strength of links in many contexts. Danowski uses a windowing approach so that the

strength of a link indicates the number of times the concepts in a network are proximal. At times Danowski uses something like a threshold for consensus (though he does not refer to it as such) to prune a network.

[8] It may be useful in addressing some hypotheses to consider how results vary as this threshold varies. If no threshold is used, then consensus equals density.

[9] Since the procedures for creating this network are not published elsewhere, we report them in the Appendix.

[10] CUBE generates for each concept in the semantic network its value on each of the three dimensions. It also generates univariate statistics for all concepts in the network. Other values not germane to this article are also calculated. CUBE is part of the MECA software package for coding and analyzing texts. For more information contact the first author.

[11] Henceforth, measures reported refer to standardized measures.

[12] The procedure for generating the drama network is identical to that described in Appendix for extracting the network for *comedy*. Additional details are provided in Appendix.

[13] In fact, Palmquist studied 13 people each in two classrooms. For our purposes in this article, we reanalyzed data from 10 students in a single class.

[14] Palmquist was most concerned with consensus even though he captured the full semantic network. We have recomputed his data to measure density and conductivity as well.

References

Alexander, M. L., & Danowski, J. A. (1990). Analysis of an ancient network: Personal communication and the study of social structure in a past society. *Social Networks, 12*, 313–336.

Bakhtin, M. (1981). *The dialogic imagination: Four essays*. (M. Holquist, Ed., and C. Emerson, Trans.). Austin: University of Texas Press.

Bar-Hillel, Y. (1960). The present status of automatic translation of languages. In F. L. Alt (Ed.), *Advances in computers* (Vol. 1, pp. 92–164). New York: Academic Press.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison: University of Wisconsin Press.

Boulding, K. E. (1956). *The image: Knowledge in life and society*. Ann Arbor: University of Michigan Press.

Burke, K. (1966). *Language as symbolic action*. Berkeley: University of California Press.

Burke, K. (1969). *A rhetoric of motives*. Berkeley: University of California Press.

Carley, K. (1984). *Constructing consensus*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Carley, K. (1986a). An approach for relating social structure to cognitive structure. *Journal of Mathematical Sociology, 12*, 137–189.

Carley, K. (1986b). Knowledge acquisition as a social phenomenon. *Instructional Science, 14*, 381–438.

Carley, K. (1987). *Language: Society's chronicle*. Paper presented at the American Sociological Association Meeting.

Carley, K. (1988). Formalizing the social expert's knowledge. *Sociological Methods and Research, 17*, 165–232.

Carley, K. (1990). *Computer analysis of qualitative data*. Paper presented at the American Sociological Association Meeting, Washington DC.

Carley, K. (in press). Content analysis. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics*. New York: Pergamon Press.

Carley, K., & Palmquist, M. (1992). Extracting, representing and analyzing mental models. *Social Forces, 70*, 601–636.

Cassirer, E. (1953). *The philosophy of symbolic forms* (R. Manheim, Trans.). New Haven: Yale University Press.

Cicourel, A. V. (1974). *Cognitive sociology: Language and meanings in sociolinguistics*. New York: Free Press.

Cicourel, A. V. (1985). Text and discourse. *Annual Review of Anthropology, 14*, 159–185.

Danowski, J. A. (1982). Computer-mediated communication: A network-based content analysis using a CBBS conference. In M. Burgoon (Ed.), *Communication yearbook 6* (pp. 905–924). Beverly Hills: Sage.

Danowski, J. A. (1988). Organizational infographics and automated auditing: Using computers to unobtrusively gather and analyze communication. In G. Goldhaber & G.

Barnett (Eds.), *Handbook of organizational communication* (pp. 385–433). Norwood, NJ: Ablex.

Duncan, H. (1968). *Symbols in society*. New York: Oxford University Press.

Durkheim, E. (1947). *The elementary forms of the religious life: A study in religious sociology* (2nd ed.). (J. W. Swain, Trans.) Glencoe, IL: Free Press.

Edelman, M. J. (1977). *Political language: Words that succeed and policies that fail*. New York: Academic Press.

Foucault, M. (1977). *Language, counter-memory, practice*. (D. F. Bouchard & S. Simon, Trans.). Ithaca, NY: Cornell University Press.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks, 1*, 215–239.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Cambridge, MA: Polity Press.

Graber, D. A. (1976). *Verbal behavior and politics*. Urbana, IL: University of Illinois Press.

Hirsch, E. D., Jr. (1987). *Cultural literacy: What every American needs to know*. Boston, MA: Houghton Mifflin.

Hudson, K. (1978). *The language of modern politics*. New York: Macmillan.

Kaufer, D., & Carley, K. (in press). Condensation symbols: Their variety and rhetorical function in political discourse. *Philosophy and Rhetoric*.

Kaufer, D., & Carley, K. (1993). *Communication at a distance: The influence of print on socio-cultural organization and change*. Hillsdale, NJ: Erlbaum.

Knoke, D., & Kuklinsky, J. H. (1982). *Network analysis*. Beverly Hills: Sage.

Leach, E. R. (1976). *Culture and communication: The logic by which symbols are connected*. Cambridge, England: Cambridge University Press.

Lippmann, W. (1922). *Public opinion*. New York: Macmillan.

McGee, M. C. (1980). The "ideograph": A link between rhetoric and ideology. *Quarterly Journal of Speech, 66*, 1–16.

Palmquist, M. (1990). *Studying classroom lexicons*. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, PA.

Pocock, J. G. A. (1971). *Politics, language and time: Essays on political thought and history*. New York: Atheneum.

Polanyi, M. P. (1962). *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.

Richards, I. A. (1936). *The philosophy of rhetoric*. New York: Oxford University Press.

Saburi, N. (1991). *World bank's adjustment lending policy 1980–1989: Towards understanding the socio-political dimensions of adjustment*. Unpublished doctoral dissertation, American University, Washington, DC.

Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science, 8*, 327–340.

Sowa, J. F. (1984). *Conceptual structures*. Reading, MA: Addison-Wesley.

Sperber, D. (1975). *Rethinking symbolism* (A. L. Morton, Trans.). New York: Cambridge University Press.

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society* (M. Cole, Trans.). Cambridge, MA: Harvard University Press.

Whorf, B. L. (1956). *Language, thought and reality*. Cambridge, MA: MIT Press.

Wittgenstein, L. (1968). *Philosophical investigations* (3rd ed.). New York: MacMillan.