Loops and Self-Reference in the Construction of Dictionaries

David Levary,[1,2] Jean-Pierre Eckmann,[3] Elisha Moses,[2] and Tsvi Tlusty[2,4]

[1]*Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA*
[2]*Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel*
[3]*Département de Physique Théorique and Section de Mathématiques, Université de Genève, CH-1211, Geneva 4, Switzerland*
[4]*Simons Center for Systems Biology, Institute for Advanced Study, Princeton, New Jersey 08540, USA*

Dictionaries link a given word to a set of alternative words (the definition) which in turn point to further descendants. Iterating through definitions in this way, one typically finds that definitions loop back upon themselves. We demonstrate that such definitional loops are created in order to introduce new concepts into a language. In contrast to the expectations for a random lexical network, in graphs of the dictionary, meaningful loops are quite short, although they are often linked to form larger, strongly connected components. These components are found to represent distinct semantic ideas. This observation can be quantified by a singular value decomposition, which uncovers a set of conceptual relationships arising in the global structure of the dictionary. Finally, we use etymological data to show that elements of loops tend to be added to the English lexicon simultaneously and incorporate our results into a simple model for language evolution that falls within the "rich-get-richer" class of network growth.

Subject Areas: Complex Systems, Statistical Physics

I. INTRODUCTION

Words are the building blocks of language. By stringing together chains of these simple lexical units, people can convey complex thoughts and ideas. As language evolves to meet society's changing communication needs, new words are constantly added to the lexicon. These additions generally serve one of two purposes: The first is to increase the rate at which a particular topic can be communicated by introducing a new word to label a concept previously represented by a string of existing words (the definition of the concept). The second is to introduce a new, previously incommunicable concept to the language.

Because of language's need to be both efficient and conceptually deep, the human lexicon is not a simple one-to-one mapping of concepts onto words but rather a complex web of semantically related parts. In studying language evolution, it is therefore convenient to represent the lexicon as a network. With this approach, words are considered to be the nodes of a graph with edges drawn based on a variety of possible relationships such as word co-occurrence in texts, thesauri, or word-association experiments on human users. Such language networks tend to be scale-free and exhibit the small-world effect (in which nodes are separated from one another by a relatively small number of edges), characteristics shared by many other complex, empirically observed networks [1,2].

The lexicon is a natural object that encompasses all of the relations between words and meanings that exist in a

language, making it extremely difficult to work with in its purest form. Dictionaries provide snapshot representations of the lexicon and as such provide an extremely useful model for studying the lexicon, and, in particular, the relationship between words and concepts. Although dictionaries link a given word to a single set of words (the definition) that can express the same meaning, this set is in fact not unique and differs between dictionaries. One might just as well replace some subset of the words in the definition of the original word in question with their respective definitions. Dictionary graphs, in which directed links are drawn from a word to the words in its definition, thus allow one to identify sets of words with equivalent meanings simply by selectively iterating through the descendants of a given node.

Although the importance of using graph theory and statistical mechanics to study dictionary graphs was recognized as early as the 1970s [3], the full structure of the graph was analyzed only recently [4]. It was found that dictionaries consist of a set of words, roughly 10% the size of the original dictionary, from which all other words can be defined. This subgraph was observed to be highly interconnected, with a central, strongly connected component, dubbed the core. The authors then studied the connection of this finding to the acquisition of language in children.

The strongly connected nature of the core suggests that definitional loops play an important role in shaping the underlying topology of the dictionary graph. Although treelike graphs are usually more amenable to analytic exploration and as such are often used to model lexical and other real-world networks [5], the existence of loops changes both the structure and the dynamics on a graph. Indeed, such loops indicate that paths are no longer unique,

and that dynamics imposed on a site may often be dependent on that site's own history, introducing a form of memory in the system. In the context of human language, loops are particularly intriguing, as they represent a form of self-reference, a condition that has been used in classical statements of logic (e.g., Russell's paradox and Gödel's theorem [6]).

Current research into the statistical physical properties of networks follows along two complementary paths. The first focuses on a variety of global properties that characterize the network, such as the shortest distance between two nodes or probability-distribution functions of the number of links pointing into the nodes ("in-degree") or out of them ("out-degree"). This direction led to concepts such as small world and scale-free or power-law networks [7]. This approach often uses graph models in which loops are neglected and the network is approximated to be a treelike graph. In random graphs, short loops are known to be extremely rare [8], and so the presence of many short loops is a clear indicator of nonrandom structure in any graph. (See also Ref. [9].)

Our analysis goes along the second path to investigating networks, which includes a growing body of research examining *local* properties that characterize networks. One example of this approach is the clustering coefficient [10], which counts the number of triangles in a particular location. A second example is the notion of network motifs, which identify the types of small loops that are functionally important in a network, and, in particular, in biological ones [11,12]. In previous work [13,14], we described the curvature induced in a network by the clustering of triangles and showed the importance of two-loops in identifying nodes that contribute highly to this curvature. Using this approach, we were then able to analyze local structures in both the World Wide Web and Email networks [13,14].

Here we investigate the role of loops in a large network, looking at the construction of a dictionary. Language, as represented by the dictionary, is an essential element of human communication that has undergone evolution under strict constraints. Within the semantic network created by the dictionary, we have identified a new role for loops. We introduce the idea of self-reference in a network, i.e., that there are physical networks that must rely on a bootstrap process for defining new elements. For such a network, the loop is an essential element of the growth process, and in a semantic context we show that new concepts can be introduced only by the insertion of a loop into the graph. Empirically, we observed that these loops occur in the dictionary graph as short units that occasionally coalesce to form larger connected components. Importantly, these components remain semantically coherent, the giant strongly connected core reported in Ref. [4] having been found to be a by-product of semantic misinterpretation in the construction of the dictionary graph. Finally, using etymological data, we demonstrate that words within the same loop tend to have been introduced into the English language at similar times, and we incorporate our results into a simple model for language evolution that falls within the "rich-get-richer" class of network growth.

II. THEORETICAL MOTIVATION

Formally defining a *concept* is both difficult and controversial [15,16]. In the context of a dictionary, one has the intuition of a highly connected set of words that are semantically linked, but formalizing this is far from trivial. Kant [17] suggested that concepts are generated by performing three types of logical operations on a set of mental images (*Vorstellungen*): comparison, reflection, and abstraction. This suggestion implies that the emergence of a concept requires the existence of a certain minimal set of "images," and the temporal order at which images are acquired therefore determines when a concept will emerge (i.e., when this minimal set becomes available). Our study goes in a similar direction, but to avoid the need for a precise definition of a concept itself, we consider the structure of the graph and the dynamic process by which the lexicon grows over time and new concepts are introduced into language.

Our intuition is that, if during the growth of the lexicon a group of words shows up, which is self-consistent and closed in itself, then a quantal increase has occurred in the capacity for representations available to the user of the lexicon. We thus associate the introduction of a new concept into language at a given time with the appearance of at least one word at that time *that was not definable at earlier times*.

Our first finding is that new concepts are introduced into language by the formation of loop structures in our graph—which are definitional loops in the dictionary. This relationship between concepts and loops reflects our basic intuition that new concepts must be self-contained and as such the collection of words used to represent them must be self-referential.

To formally prove our claim, we consider a discrete growth model of the language, letting $\mathbb{W}_t$ be the set of words added to the language at time t. When a word w is added to the dictionary, we assume for now only that its definition $D(w)$ is nonempty and allow existing words to add w to their definitions with no other changes occurring to the network. These rules allow both synchronous loops, which consist exclusively of words added to the lexicon at a particular time, and diachronic loops, which combine words added at different times. We assume that the word and its definition are equivalent and therefore interchangeable. It follows that we can produce multiple equivalent definitions for a word by iteratively substituting $w \rightarrow D_0(w)$ for words in its definition where $D_0(w)$ represents the original definition of a word w (the set of all nodes within a directed path of length 1).

In the context of an expanding lexical network, the addition of a concept can be observed by the appearance of an associated lexical structure (i.e., a group of connected words) that increases the breadth of communicable ideas. Stated more precisely, a concept is created at time t if and only if there exists a word at time t that was not definable before t (i.e., a word whose meaning was incommunicable before t). Formally, we consider a word w to be definable at time t if and only if there exists a definition $D(w)$ of w for which all elements are independent of the words added after time t, or, graphically, a definition for which no elements have a descendant (i.e., a node to which it can be connected by a directed path) introduced after time t.

With these definitions (applied to two simple toy graphs in Fig. 1), it becomes clear that a new concept can be created at time t if and only if a loop was formed at t. If we consider a loop created at time t (i.e., a loop whose youngest element was introduced at t), the meanings of all its elements can no longer have been conveyable before t for they now depend inseparably on a word or words added to the lexicon at time t. Formally, if $w_t \in \mathbb{W}_t$ is a member of a loop, we know that for at least one word of $D(w_t)$ there exists a directed path from an element of $D(w_t)$ to w_t. Thus w_t is not definable before time t, and a new concept was indeed created.

On the other hand, if no loop was created at t, we can simply replace all elements of $\mathbb{W}_t$ by their definitions and then remove them from the dictionary, as we are guaranteed that these definitions are independent of the words being defined. Stated more formally, for all words w for which $D_0(w) \cap \mathbb{W}_t$ is nonempty, we can make the substitution $w_t \rightarrow D_0(w_t)$ for all $w_t \in \mathbb{W}_t$ in their definition.
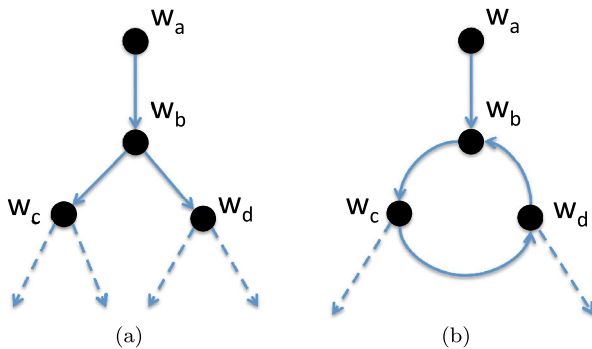

FIG. 1. Graph structure examples. Let $\mathbb{W}_{t+1} = \{w_a\}$, $\mathbb{W}_t = \{w_b\}$, and $\mathbb{W}_{t-1} = \{w_c, w_d\}$, and consider the definability of w_a. In (a), we can substitute for w_b its definition $w_b \rightarrow D_0(w_b)$, and use it in the definition of w_a so that $D(w_a) = \{w_c, w_d\}$. Since there is no directed path from w_c or w_d to w_b, we see that w_a was actually definable before time t; no new concept was created by w_b. In (b), any definition of w_a must include w_b, w_c, or w_d. Since w_b is a descendant of all of these words, w_a is not definable before t, implying that a new concept was formed. In fact, despite having originally been introduced at $t-1$, w_c and w_d are now undefinable before t, their meanings having changed with the addition of w_b.

Since we are assuming that no elements of $\mathbb{W}_t$ are in loops, we are guaranteed that, for all such w_t, no directed path exists from an element of $D_0(w_t)$ back to w_t. It follows that all words in the dictionary can be defined such that no directed path exists from them to an element of $\mathbb{W}_t$, and that they were therefore definable before time t. Thus no new concept was created at time t.

III. DICTIONARY CONSTRUCTION AND TOPOLOGY

In order to search for definitional loops in an actual dictionary, one must be able to link all of the words in a given definition to their respective definitions. This requires both the reduction of inflected words to their stems and the resolution of polysemous words to their proper sense. We therefore use as our primary dictionary the eXtended WordNet, which provides semantically parsed definitions for each WordNet 2.0 *synset* (set of synonymous words) [18–20]. To reduce complexity, we have chosen to restrict our attention to nouns as they are the part of speech generally most directly related to the main concepts within a text [21]. We have verified our basic results using a lower-resolution graph constructed from an online dictionary [22].

We treat the dictionary as a directed graph in which WordNet synsets are designated as nodes, with a directed link drawn from a node to all of the synset nodes that appear in its definition. With this construction, each sense of a word is represented by a separate node, corresponding to the fact that nodes in our graph represent unique meanings labeled by one or more words in the original dictionary. The resulting graph consists of 79 689 nodes and 285 773 edges.

Decomposition of the graph into strongly connected components (subgraphs in which every node is reachable along a directed path from every other node) using Tarjan's algorithm [23], yielding a set of 1123 strongly connected components (SCCs) with median size of 2. In keeping with the results of Ref. [4], we found a single large component, the "core," consisting of 6296 nodes reachable along a directed path from over 99% of nodes. Interestingly, as illustrated in Fig. 2, definitional paths converge on the core very quickly, accumulating only a very small number of origin-specific synsets. After only 12 steps, most paths have already encompassed half the core (inset to Fig. 2), and by 30 steps, all descendants have already been reached, indicating that the core is structured around a large number of overlapping loops.

Before we analyze these loops in greater detail, it is helpful to characterize the types of words that appear in the dictionary's strongly connected components. Theoretically, the set of all words involved in loops should be sufficient to define all words in the dictionary, albeit with extensive paraphrasing, and thus can be thought of as a simple vocabulary. Having identified these words by purely computational means, it is interesting to compare
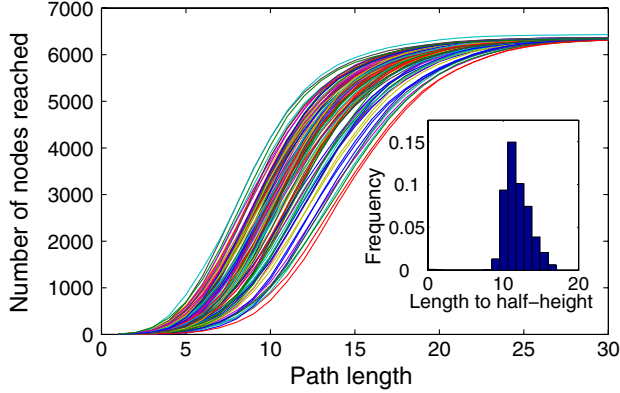
FIG. 2. Definitional iteration of words in the dictionary. Using a random sample of 100 words, the number of unique nodes that could be reached within the given directed distance of each node is recorded. Nearly all starting points lead to a strongly connected component of 6296 words labeled as the core.

them to those found in other such vocabularies. We have compared our core to Basic English [24], a set of 850 words that British linguist Charles Ogden claimed sufficient for daily discourse, as well as to the English translations of the words in Jōyō Kanji, the Japanese Education Ministry's list of 2136 characters required to be learned by Japanese secondary school students (accessed from Ref. [25]). As a control, we have also compared these lists to the top 1000 most frequently used words in all books found on Project Gutenberg (accessed from Ref. [26]). As these lists were of course not sense-disambiguated, we have temporarily reduced the resolution of our graph by making the nodes words (instead of synsets) and by using only the first sense of the definition. Again, we have considered only nouns in all comparisons.

Our low-resolution set of 1595 core words does share great overlap with all three lists. (See Table I.) Notably, however, the overlap never exceeds 50% of any word list. A survey of those words in Basic English not found in the

TABLE I. Intersection of the core with other simple word lists. Table entries represent the number of words in the intersection of the sets, with percent overlap given in parentheses. The core is obtained using a simplified WordNet dictionary graph, in which nodes were words (not synsets), with only the first sense of the definition considered. This method yields a graph with lower resolution than the one obtained using the sense-disambiguated data from eXtended WordNet. We use it because the lists we compare to are not sense-disambiguated. Only nouns in each word list are considered. Descriptions of the word lists are found in the main text.

	Core	Basic English	Jōyō Kanji	Gutenberg
Core	1595	314 (52%)	403 (29%)	265 (39%)
Basic English		600	328 (24%)	213 (32%)
Jōyō Kanji			1376	319 (47%)
Gutenberg				673

core reveals a trend of potentially useful but perhaps definitionally "overspecific" words such as *apple, brick, chalk, hammer*, and *glove*. While these words might come in handy in daily life, as Ogden had intended, it is easy to see how these words would be reduced in our dictionary into more general words which in combination can communicate those more specific words. (For example, in the case of *apple*, both *fruit* and *red* appear in our core.)

IV. THE LOOPS

In the theoretical discussion, we showed that the appearance of loops in the dictionary can be associated with the creation of new concepts. While we have made no assumptions about the interconnectivity of these loops, it seems unlikely that the majority of concepts in language would be interdependent as the existence of the core seems to suggest. To better understand how definitional loops form in the dictionary, we search for cycles in the dictionary graph.

A total of 9085 nodes are identified to be elements of loops. Given the observed high degree of overlap among loops, we have found it useful to classify these nodes and the links that connect them according to the shortest loop in which they appear. The corresponding distribution of loop lengths, shown in Fig. 3, turns out to be particularly illuminating. It appears that cycles in the dictionary fall into two classes: short (≤ 5) and long (> 5). While the appearance of long loops can be predicted solely based on the in- and out-degree distributions of our graph (the randomization in the figure), the short loops appear to be a unique feature arising from *meaningful* connections between nodes. Inspection of individual loops confirms this assessment. Whereas small cycles follow a very clear conceptual path, large cycles are for the most part characterized by one or more conceptual leaps, typically caused by a misinterpretation of word sense, as the following example illustrates:
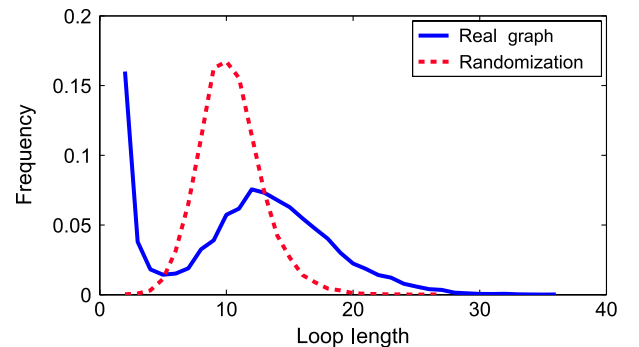

FIG. 3. Distribution of definitional loops in the dictionary. The data represent counts of links in the core indexed by the shortest loop in which they appear. For the randomization, links are randomly redrawn between nodes while keeping the in-degree and out-degree distributions of the graph constant.
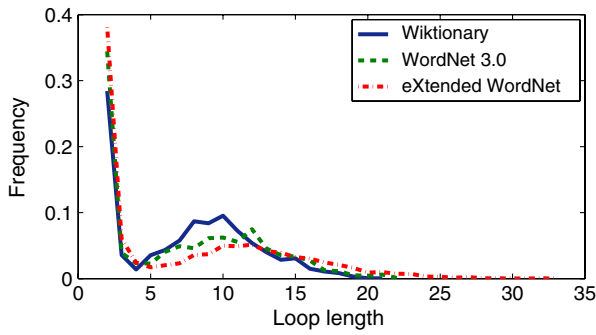
FIG. 4. Distribution of definitional loops for several dictionaries. Both the Wiktionary and WordNet 3.0 graphs are constructed by considering only the first sense of a word in the event of polysemy. However, in WordNet, the ordering of senses is determined empirically according to usage frequencies in written texts, while in Wiktionary, the ordering of senses is determined somewhat arbitrarily, with the definition page as a whole representing a general consensus of users.

$$\text{railcar} \rightarrow \text{rails} \rightarrow \text{bar} \rightarrow \text{weapon} \rightarrow \text{instrument} \rightarrow \text{skill}$$
$$\nrightarrow \text{train} \rightarrow \text{railcar}.$$

Although the link between *bar* and *weapon* is perhaps questionable, the link between *skill* and *train* clearly is a case of mistaken sense, in this case, between *train* the verb and *train* the noun. Such errors reflect the fact that the semantic tagging in eXtended WordNet was done largely computationally and is therefore subject to mistakes.

Figure 3 also shows a slight overabundance of links involved in large loops in the dictionary as compared to the randomization. This longer tail appears to result from the fact that not all connections within a long loop are false, as illustrated in the example loop above. It therefore takes

more connections for a false loop to form in the real data than in the randomization, where *every* link is likely wrong.

Given our finding that long loops are generally formed from semantic misinterpretations, one might expect that, the better the system for assigning links in cases of polysemy, the lower the ratio of large loops to small loops. To test this intuition, we construct two additional graphs based on the definitions in the English Wiktionary [22] and WordNet 3.0 [27]. While both the Wiktionary and WordNet 3.0 graphs have been constructed by considering only the first sense of a word in the event of polysemy, in WordNet, the ordering of senses was determined empirically according to usage frequencies in written texts while, in Wiktionary, the ordering of senses is determined somewhat arbitrarily, with the definition page as a whole representing a general consensus of users. As illustrated in Fig. 4, although the distribution of loop length for all three graphs is similar in shape, the ratio of small to large loops increases with the sophistication of the system used for link assignment, suggesting that long loops would be essentially nonexistent in a dictionary with completely manual semantic tagging.

It is important to note that, given the high-degree of connectivity between loops, meaningful longer loops do exist. The links within these large loops, however, are simultaneously involved in small loops, and as a result the loops generally follow a logical progression of ideas. Figure 5 provides a graphical view of the overlap among loops, depicting a strongly connected component formed by considering only links involved in small loops.

The finding that the dictionary graph contains many false loops suggests that the core does not reflect large-scale conceptual interdependence in the lexicon but rather
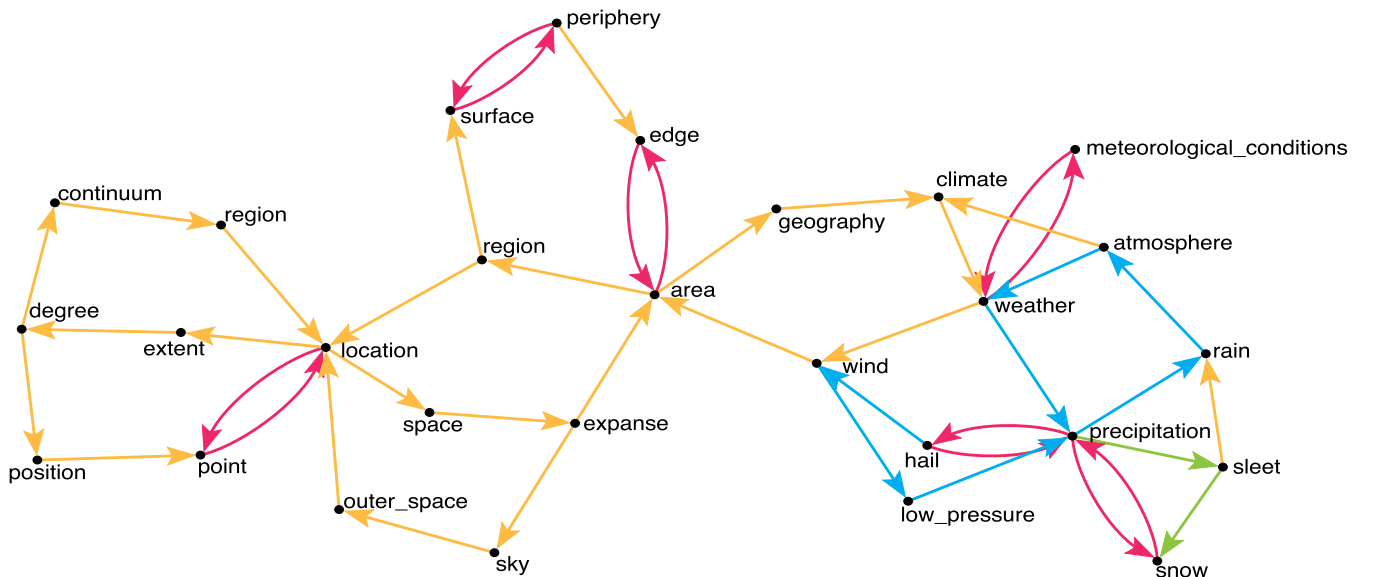

FIG. 5. An example of a large, strongly connected component in the decomposition. Arrows are drawn from a node to words in its definition. Red links appear first in two-loops, green in three-loops, blue in four-loops, and orange in five-loops.

TABLE II.   Each column lists examples of strongly connected components in the dictionary graph consisting of links involved in short (≤ 5) cycles.

emotion	height	bark	injury	winner
spirit	end	trunk	violence	contestant
dejection	dimension	tree	accident	competition
melancholy	length	lumber		
feeling				

exists as an artifact of imperfect dictionary construction. Indeed, when we consider only the links in the core involved in short loops (length ≤ 5), we find that the core decomposes into several hundred SCCs. Inspection of these and other SCCs outside of the core reveals a high degree of intracomponent semantic coherence (Table II), not surprising given that SCCs by definition are sets of words whose meanings are completely interdependent.

Although connected components in our decomposed graph represent unique semantic ideas or concepts, they are not completely independent of one another. Meaningful connections between the connected components do of course exist, our results suggesting simply that these connections are generally acyclic in nature. In order to better characterize the interactions among components and their role in the lexicon as a whole, we wish to "define" each word in the dictionary in terms of these semantic units. After ignoring SCCs outside of the core which we found consist almost exclusively of highly technical scientific words, we have identified 386 connected components consisting exclusively of links involved in short loops to serve as our conceptual vocabulary. To quantify the importance of each component in a given word's definition, we count the number of paths in our original graph leading from the word to a given cluster. In an attempt to increase the definitional weight of clusters located close to the word in question, we allow vertices and edges to be repeated

when counting paths so that the number of paths to a closer cluster continues to grow in the time taken to reach a farther one. This choice requires us to impose a bound on the length of path we consider. We choose this upper limit in path length as 5, in keeping with our finding that loops of size greater than 5 usually emerge from semantic misinterpretations. Each node in the original graph can now be associated with a vector whose elements are the number of paths from that node to each of the 386 components. Concatenating these vectors yields a sparse $79\,689 \times 386$ matrix.

In analyzing this matrix, we have found that five components appeared in over 80% of the vectors. Not surprisingly, these components consist of very general words (e.g., *entity* and *group*) and are thus ignored in further analysis and removed from the matrix. In an attempt to identify cohesive groups of connected components, we perform singular-value decomposition on our matrix. The resulting singular vectors (examples of which can be found in Table III) show a striking ability to capture major themes within the dictionary including geography, life, and religion. It is, however, the connections between the elements in these singular vectors that are most significant. Although normally obscured by noisy connections in the dictionary, links among topics such as the body, water, energy, and disease in our singular vectors reflect powerful semantic chains underlying the *conceptual* lexicon.

V. LOOP ETYMOLOGY

As we have seen, definitional loops form the conceptual basis of language. When one considers the evolution in time of the lexicon, the question arises how these loops came to exist. Using the Online Etymology Dictionary [28], we have manually looked up the dates of origin for words in definitional loops. Dates have been recorded only

TABLE III.   Examples of the highest singular components for the dictionary. The elements in the singular components are the semantically cohesive clusters of words obtained from decomposing the core. For table entries, word(s) representing the main theme of each cluster are chosen. Clusters are listed in order of the absolute value of their coefficient in the singular component. Only clusters whose coefficients have absolute values that are greater than 0.1 have been listed. Plain text and italics indicate the clusters that have coefficients with positive and negative values, respectively. The third, fourth, and seventh highest singular components are very similar to vectors already shown and are therefore not displayed.

Vector 1	Vector 2	Vector 5	Vector 6	Vector 8
Old World, oceans	*spine, brain*	grains	*body parts*	Jesus, Christianity
bodies of water	*body parts*	*bodies of water*	*flower, seed*	man, woman
The Americas	Old World, oceans	herbage	*tree, bark*	*nucleus, DNA*
influence, power	*energy*	flower, seed	grains	Roman Empire
	pathology	*land*	nucleus, DNA	student, teacher
	narrative	*water*	spine, brain	speaker, speech
	cognition		*Gymnosperms*	Vatican, absolution
	organic process		bodies of water	Old Testament
	respiration		pathology	book

when our manual inspection shows that the definitions in the etymology dictionary indeed match the same sense of a word as the one that appears in the loop. In the case of synsets with multiple words, only the first word in the synset is used, as we did for the data in Table I. Given the considerable vagueness surrounding dates of emergence in Old English, for the purposes of our analysis all Old English words are recorded as having emerged in the year 1150.

After eliminating proper nouns and compound words, we have found dates for 971 words representing 310 non-overlapping loops. As shown in Fig. 6(a), the distance among dates of origin of words in the loops is for the most part considerably smaller than that obtained by randomly clustering these dates. Figure 6(a) clearly shows that, in the real graph, the majority of the loops have dates of origin of words that differ by no more than 150 years, while, in the randomized data, the majority have more than 150 years between them.

While several loops do contain words with somewhat disparate dates of origin, we have found that such exceptions often reflected fundamental changes in the understanding of a word after its introduction. For instance, the word *atom* was first introduced to English in the late 15th
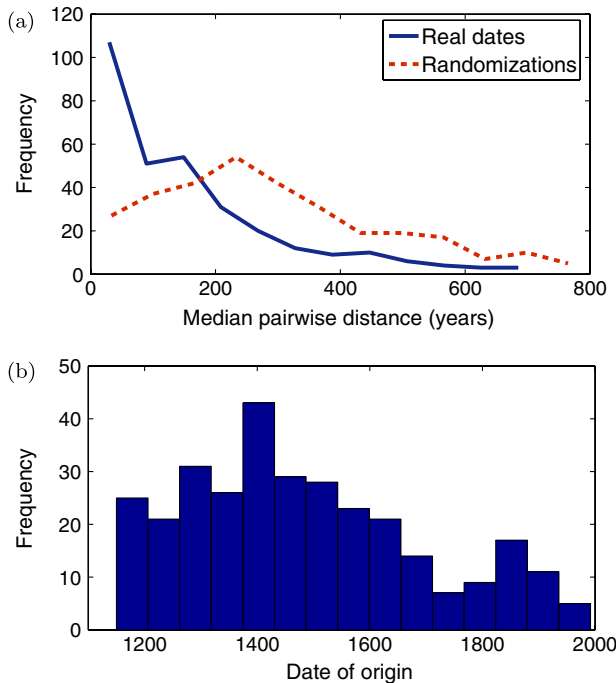

(a)

(b)

FIG. 6. Dates of origin of words in loops. For each loop, the dates of origin of its element words (in the desired sense) were looked up in the Etymology Dictionary. Compound words and proper nouns were ignored, as well as polysemous words. The median pairwise distance of elements (a) and the mean date of origin (b) were calculated for each of the 310 distinct loops in our analysis.

century to denote a hypothetical indivisible particle from which all matter in the universe was built. With the discovery of the atomic nucleus by Rutherford in the early 1900s, however, the concept of the indivisible atom was fundamentally changed and with it a new sense of the word *nucleus* was created.

The apparent coevolution of words in loops is quite striking. While words in a loop are of course semantically related, there is no *a priori* reason to assume that semantically related words in general emerge around the same time period. For instance, the word *sneaker* is clearly closely related to the word *shoe*, yet it is not surprising that the two words emerged at very different epochs. (The Online Etymology Dictionary places *sneaker* in 1895 and *shoe* in Old English.) The finding that words in loops are typically introduced into language at the same time thus appears to reflect the unique type of semantic relationship they share and bolsters our claim that loops necessarily appear in the lexical network since they allow for the communication of new concepts.

Given this relationship between loops and concepts, it is also interesting to consider the distribution of mean dates of origin for the loops [Fig. 6(b)], as it is likely indicative of major periods of conceptual expansion within the English language. The distribution peaks between the 14th and 16th centuries, which corresponds to the transition from Middle to Early Modern English, a period marked by the development of the printing press and the establishment of enduring English literature. A second, smaller influx of largely scientific words has occurred in the past two centuries, reflecting the prolific progress of modern science and technology.

VI. MODEL FOR LEXICAL GROWTH

Using our results from the Online Etymology Dictionary, we can now model the process by which new words are added to the lexicon. The basic attachment rules are manifested in $P(k_{in})$ and $P(k_{out})$, the probability distributions of the in-degree and out-degree, respectively. We observe that, for the dictionary network, $P(k_{out})$ is described by a Poisson distribution [Fig. 7(a)], in keeping with the intuition that most definitions have roughly the same number of words, while $P(k_{in})$ decays as a power following $P(k) \sim k^{-\gamma}$, where $\gamma = 2.1$ [Fig. 7(b)]. The fact that $P(k_{in})$ follows a power law suggests that new words preferentially link to existing words with high in-degrees [7]. Indeed, preferential attachment is logical in the case of the dictionary network if one treats the dictionary as a corpus, for it is reasonable to assume that in defining a new word we tend to use common words (i.e., high in-degree) to ensure that the new word is easily understood.

A simple model to explain the growth of the lexical network adds a new word to the dictionary at each time step with K_{out} outgoing edges, where K_{out} (the number of words in the word's definition) is a Poisson-distributed
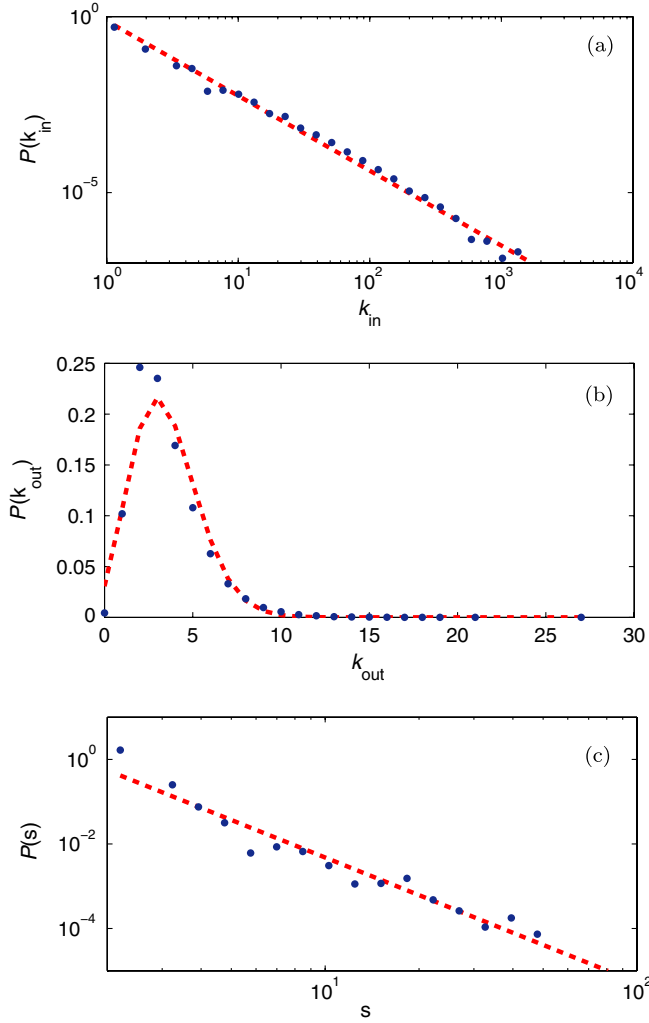
FIG. 7. Extraction of model parameters. Probability distributions were measured for (a) in-degree, (b) out-degree, and (c) strongly connected component sizes for the empirical dictionary graph. The dashed lines in (a) and (c) have slopes 2.1 and 2.9, respectively, while the Poisson fit in (b) has parameter 3.5. In the model, these statistics correspond to a new word attractiveness $a = 0.4$ and a probability of new SCC formation $p = 0.4$.

random variable ($K_{out} \sim \text{Poiss}(\lambda)$) [29]. We define the new word in terms of existing words by letting each new directed edge point to an existing word i with probability proportional to $a + k_{in}(i)$, where $a > 0$. Here, a is the so-called "attractiveness" of a new word [30], specifically what allows a newly added word (with in-degree zero) to be used in the definition of words at later time steps. It has been shown [29] that, given these growth rules, the in-degree distribution for this network decays as a power law with

$$P(k_{in}) \sim k_{in}^{-(2+\delta)}, \tag{1}$$

where $\delta = a/\langle k_{out} \rangle = a/\lambda$. For appropriate choice of the parameter a, this model is consistent with the observed in- and out-degree distributions in the dictionary network and generally explains how acyclic links are added to the network.

To incorporate the formation of loops into the model, we assume that the elements of loops are added to the lexicon simultaneously based on our results with the Online Etymology Dictionary. However, we must also explain the existence of the larger strongly connected components in the decomposed graph. It is unreasonable to assume that these components, and the concepts they represent, must always appear instantaneously in the network. Rather, existing concepts may be expanded by new discoveries as, for example, occurred to the notion of a *cell* with the advent of cellular and molecular biology. As such conceptual expansion can be driven only by the addition of new loops to an existing strongly connected component, we assume that, when a loop is introduced to the network, it may either start its own component (with fixed probability p) or join an existing SCC (with probability $1 - p$). Since the probability distribution $P(s)$ of SCC size follows a power law [Fig. 7(c)], we assume that loops preferentially attach to larger SCCs. Intuitively, this a reasonable assumption if we consider larger components to in general represent broader concepts, thereby presenting more possibilities for expansion. Specifically, we claim that the probability π_i that a loop attaches to an existing component i of size s_i is $\pi_i = (1 - p)\frac{s_i}{\sum_j s_j}$. Assuming the size of loops to be added is constant, this stochastic process can be shown to lead to a set of SCCs whose size distribution decays as a power law with

$$P(s) \sim s^{-[1+1/(1-p)]}, \tag{2}$$

which is consistent with the observed algebraic constant $\gamma = 2.8$ for $p = 0.4$ [31]. In reality, of course, the size of the loops to be added at each step varies. Numerical simulations, however, show that small variations in the loop size do not affect the overall power-law character of $P(s)$ (data not shown).

VII. CONCLUSIONS

Self-reference in dictionary definitions exists not as a trivial artifact of a dictionary's construction but rather as the mechanism by which concepts are created and stored in language. In contrast to the expectations for a random lexical network, meaningful definitional loops appear as short structures, typically consisting of from two to five elements. These loops are not strictly isolated but are often linked to form larger, yet still semantically coherent, strongly connected components. While these components have been observed to represent distinct semantic ideas, by analyzing the interactions among them, we have been able to reveal a set of conceptual

relationships upon which the lexicon appears to have been built.

Our finding that the words within loops tend to be added to the lexicon simultaneously underscores the unique relationship that these words share. Although in theory one need only know the meanings of some subset of the words in a loop in order to infer the definitions of the remaining words, at the conceptual level the meanings of these words remain completely intertwined. This analysis of course begs the question of how loops could have come to exist in the first place. In order for a word to be introduced into language, it must be understood by multiple individuals to mean the same thing. The necessary synchronization of word meaning among different individuals is particularly difficult when the meanings themselves exist as conceptual loops. A potential solution to this problem is for an individual to attempt to sequentially define all the elements of the loop. While the central concept of the loop cannot be directly communicated, the juxtaposition of partially defined elements of the loop may allow the receiver to infer the common link among the words, thereby completing the definition of all words in the loop. Such a system would be particularly effective with short loops and is consistent with our finding that words within a loop tend to enter the lexicon at the same time.

ACKNOWLEDGMENTS

[1] A. E. Motter, A. P. S. de Moura, Y. C. Lai, and P. Dasgupta, *Topology of the Conceptual Network of Language*, Phys. Rev. E **65**, 065102 (2002).

[2] R. F. Cancho and R. V. Solè, *The Small World of Human Language*, Proc. R. Soc. B **268**, 2261 (2001).

[3] K. C. Litkowski, *Models of the Semantic Structure of Dictionaries*, Am. J. Comput. Ling. **81**, 25 (1978) [http://www.clres.com/online-papers/ajcl78.html].

[4] O. Picard, A. Blondin Massé, S. Harnad, O. Marcotte, G. Chicoisne, and Y. Gargouri, in *Advances in Neural Information Processing Systems 22: 23rd Annual Conference on Neural Information Processing Systems 2009, Vancouver, British Columbia*, edited by Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta (Curran Associates, Inc., Red Hook, New York, 2009) [ISBN: 9781615679119; http://arxiv.org/pdf/0911.5703.pdf].

[5] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, *Introduction to WordNet: An On-line Lexical Database*, Int. J. Lexicography **3**, 235 (1990).

[6] Bertrand Russell, *Principles of Mathematics* (W. W. Norton, 1938), 2nd ed. [http://fair-use.org/bertrand-russell/the-principles-of-mathematics/]; K. Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*, Monatshefte für Mathematik **38-38**, 173 (1931) [http://www.springerlink.com/content/p03501kn35215860/].

[7] A.-L. Barabàsi and R. Albert, *Emergence of Scaling in Random Networks*, Science **286**, 509 (1999).

[8] B. Bollobás, *Random Graphs* (Cambridge University Press, Cambridge, England, 2001), 2nd ed.

[9] P. Collet and J.-P. Eckmann, *The Number of Large Graphs with a Positive Density of Triangles*, J. Stat. Phys. **109**, 923 (2002).

[10] D. J. Watts and S. H. Strogatz, *Collective Dynamics of 'Small-World' Networks*, Nature (London) **393**, 440 (1998).

[11] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Network Motifs: Simple Building Blocks of Complex Networks*, Science **298**, 824 (2002).

[12] A. Vazquez, R. Dobrin, D. Sergi, J.-P. Eckmann, Z. N. Oltvai, and A.-L. Barabàsi, *The Topological Relationship between the Large-Scale Attributes and Local Interaction Patterns of Complex Networks*, Proc. Natl. Acad. Sci. U.S.A. **101**, 17940 (2004).

[13] J.-P. Eckmann and E. Moses, *Curvature of Co-links Uncovers Hidden Thematic Layers in the World Wide Web*, Proc. Natl. Acad. Sci. U.S.A. **99**, 5825 (2002).

[14] J.-P. Eckmann, E. Moses, and D. Sergi, *Entropy of Dialogues Creates Coherent Structures in E-mail Traffic*, Proc. Natl. Acad. Sci. U.S.A. **101**, 14333 (2004).

[15] G. L. Murphy, *The Big Book of Concepts* (MIT Press, Cambridge, Massachusetts, 2002).

[16] E. Machery, *Doing without Concepts* (Oxford University Press, New York, 2009).

[17] I. Kant, *Logik, ein Handbuch zu Vorlesungen, Logische Actus der Komparation, Reflexion und Abstraktion, Section I, § 6* (Königsberg, 1800).

[18] R. Mihalcea and D. I. Moldovan, in *Proceedings of the NAACL Workshop on WordNet and Other Lexical Resources, Pittsburgh, 2001* (Association for Computational Linguistics, Stroudsburg, Pennsylvania, 2001) [http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.5938].

[19] G. A. Miller, *WordNet: A Lexical Database for English*, Commun. ACM **38**, 39 (1995).

[20] eXtended WordNet is a sense-disambiguated version of WordNet 2.0 [http://xwn.hlt.utdallas.edu/index.html].

[21] E. Alvarez-Lacalle, B. Dorow, J.-P. Eckmann, and E. Moses, *Hierarchical Structures Induce Long-Range Dynamical Correlations in Written Texts*, Proc. Natl. Acad. Sci. U.S.A. **103**, 7956 (2006).

[22] http://www.wiktionary.com

[23] R. E. Tarjan, *Depth-First Search and Linear Graph Algorithms*, SIAM J. Comput. **1**, 146 (1972).

[24] C. K. Ogden, *Basic English: A General Introduction with Rules and Grammar* (Kegan Paul, London, 1930).

[25] Wikipedia, *List of Jōyō Kanji*, http://en.wikipedia.org/wiki/List_of_joyo_kanji.

[26] Wiktionary, *Word Frequency Lists*, http://en.wiktionary.org/wiki/Wiktionary:Frequency_lists.

[27] WordNet 3.0: http:/wordnet.princeton.edu/.

[28] D. Harper, Online Etymology Dictionary [http://www.etymonline.com/].

[29] D. Fraiman, *Growing Directed Networks: Stationary In-Degree Probability for Arbitrary Out-Degree One*, Eur. Phys. J. B **61**, 377 (2008).

[30] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Structure of Growing Networks with Preferential Linking*, Phys. Rev. Lett. **85**, 4633 (2000).

[31] H. A. Simon, *On a Class of Skew Distribution Functions*, Biometrika **42**, 425 (1955) [http://www.uvm.edu/~pdodds/teaching/courses/2009-08UVM-300/docs/others/1955/simon1955.pdf].