Hierarchies in Dictionary Definition Space

Olivier Picard[1], **Alexandre Blondin Massé**[2], **Stevan Harnad**[1],
Odile Marcotte[3], **Guillaume Chicoisne**[1], **Yassine Gargouri**[1]

[1] Institut des sciences cognitives
Université du Québec à Montréal (UQÀM)
Montréal (Québec) Canada H3C 3P8
`picard.olivier.2@courrier.uqam.ca, harnad@uqam.ca,`
`chicoisne.guillaume@uqam.ca, yassinegargouri@hotmail.com`

[2] Laboratoire de combinatoire et d'informatique mathématique
Université du Québec à Montréal (UQÀM)
Montréal (Québec) Canada H3C 3P8
`alexandre.blondin.masse@gmail.com`

[3] Groupe d'études et de recherche en analyse des décisions (GERAD) and UQÀM
HEC Montréal
Montréal (Québec) Canada H3T 2A7
`Odile.Marcotte@gerad.ca`

Abstract

A dictionary defines words in terms of other words. Definitions can tell you the meanings of words you don't know, but only if you know the meanings of the defining words. How many words do you need to know (and which ones) in order to be able to learn all the rest from definitions? We reduced dictionaries to their "grounding kernels" (GKs), about 10% of the dictionary, from which all the other words could be defined. The GK words turned out to have psycholinguistic correlates: they were learned at an earlier age and more concrete than the rest of the dictionary. But one can compress still more: the GK turns out to have internal structure, with a strongly connected "kernel core" (KC) and a surrounding layer, from which a hierarchy of definitional distances can be derived, all the way out to the periphery of the full dictionary. These definitional distances, too, are correlated with psycholinguistic variables (age of acquisition, concreteness, imageability, oral and written frequency) and hence perhaps with the "mental lexicon" in each of our heads.

Keywords: categories, definition, dictionary, feedback vertex set, graph theory, language learning, lexicography, mental lexicon, semantics, symbol grounding, vocabulary, word meaning

1 Introduction

A category is a *kind* of thing (object, event, action, trait or state). To categorize is to do the right thing (eat, fight, flee, mate, etc.) with the right kind of thing. All species can acquire categories through trial and error *sensorimotor induction*. We are the only species that can also acquire and transmit categories through *verbal instruction*, by naming and defining them. The words in our dictionaries are almost all the names of categories, followed by their definitions. In principle, all

1

categories can be acquired through verbal definition, but we cannot acquire all of them that way: we have to know the meanings of some of the defining words already, by some other means. This is the "symbol grounding problem" [4] and presumably that other means of acquiring categories is sensorimotor induction. But how many words – and which ones – need to be grounded directly through sensorimotor induction in order to allow all the rest to be acquired through verbal definition?

We have been analyzing dictionaries in order to answer this question. By eliminating all the words that can be reached from other words through definition alone, we have been able to reduce the dictionary to its "grounding kernel" (GK) – a set of words (about 10%) – out of which all the rest of the words can be reached through definition alone [1]. The GK has some striking properties: The words in it are learned at a significantly younger age than the rest of the dictionary and are also more concrete [2], but if the variance correlated with age is removed, the residual GK words are more abstract than the rest of the dictionary. What is the cause of this polarity shift?

The GK is unique, and sufficient to ground all the rest of the dictionary, but it is not *minimal* – it is not the smallest set of words from which all the rest can be reached via definition alone. That would be a "minimum grounding set" (MGS), which is not in general unique; we have not yet been able to compute a MGS, because this problem (equivalent to finding a "minimum cardinality feedback vertex set" for a general graph) is NP-complete (i.e. too hard to compute in general). We hope to be able to compute MGSs for our special cases, but meanwhile the GKs of our dictionaries – Cambridge International Dictionary of English (CIDE) [8] and Longman Dictionary of Contemporary English (LDOCE) [7] – already turn out to have more differentiated internal substructure which we begin analyzing further in this article.

In particular two substructures play important roles: the GK itself and a strongly connected subset of the GK that we call the "Kernel Core" (KC). The GK words that are acquired earlier, and are more concrete than the rest of the dictionary, tend to be in the KC, whereas the GK words uncorrelated with age of acquisition tend to be in the outer layer surrounding the KC and are more abstract. These correlations between the KC and the rest of the GK, and between the GK and the rest of the dictionary as a whole, are binary (0/1), but one can make more graded comparisons by considering definitional chains of increasing lengths. We have accordingly extracted two hierarchies based on degrees of definitional distance, one based on the GK and one based on strongly connected components, to analyze how definitional distance correlates with age of acquisition, concreteness/abstractness and other psycholinguistic variables.

2 Definitions and Notations

This section introduces all necessary definitions of graph-theoretical objects studied in this article. The reader is referred to [9] for complete graph theory and discrete mathematics introductions.

2.1 Graphs

A *directed graph* is a couple $G = (V, E)$, where V is a finite set of elements called *vertices* and $E \subseteq V \times V$ is a finite set of couples of vertices called *arcs*. Given some graph G, we denote its set of vertices and edges by $V(G)$ and $E(G)$ respectively. The *density* of a directed graph is $d(G) = |E(G)|/|V(G)|^2$.

A vertex u is a *predecessor* (or *successor*) of vertex v if $(u, v) \in E$ (or $(v, u) \in E$). The sets of predecessors and successors of u are denoted respectively by $N^-(u)$ and $N^+(v)$. The *in-degree* and *out-degree* of u are defined respectively by $\delta^-(u) = |N^-(u)|$ and $\delta^+(u) = |N^+(u)|$. A vertex of null in-degree (or out-degree) is called a *source* (or *sink*).

A *finite path of length* n in a graph is a sequence $(v_0, v_1, \ldots, v_n)$, where $n \geq 0$ is an integer and $(v_{i-1}, v_i) \in E$ for $i = 1, 2, \ldots, n$. A *uv-path* is a path starting with u and ending with v. A *uv*-path is a *cycle* if $u = v$. A graph is *acyclic* if it contains no cycles.
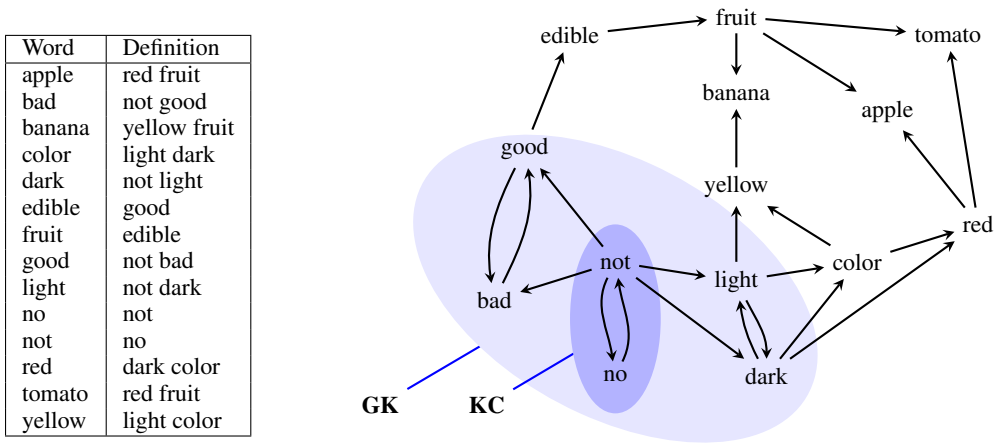
Given a graph $G = (V, E)$, we say that $G' = (V', E')$ is a *subgraph of* G if $V' \subseteq V$ and $E' \subseteq E$. Moreover, if $V' \subseteq V$, the *subgraph of G induced by* V', denoted by $G[V']$, is the subgraph $G' = (V', E')$ such that $E' = (V' \times V') \cap E$.

2.2 Dictionaries

Let W be a finite set whose elements are called *words*, and let 2^W denote the collection of all subsets of W. A *dictionary* is a subset D of $W \times 2^W$ such that for every $(w, d_w) \in D$: (i) $d_w \neq \emptyset$ (there is no empty definition) and (ii) $w \notin d_w$ (a word cannot be used to define itself).

Elements of D are called *entries*. An entry is therefore a couple (w, d_w), where w is a word and d_w is a set of words. The element w, the set d_w and the elements of d_w are called respectively the *definiendum* (the defined word), the *definition of* w and the *definientes* (the defining words).

There is a very natural way to derive a graph from a dictionary. The *associated graph* of a dictionary $D \subseteq W \times 2^W$ is the directed graph $G = (V, E)$ where $V = W$ and $(u, v) \in E$ if and only if there exists an entry $(v, d_v) \in D$ such that $u \in d_v$. In fact, associated graphs are exactly directed graphs without loops and without sources. An artificially constructed "toy" dictionary is illustrated in Figure 1.

Word	Definition
apple	red fruit
bad	not good
banana	yellow fruit
color	light dark
dark	not light
edible	good
fruit	edible
good	not bad
light	not dark
no	not
not	no
red	dark color
tomato	red fruit
yellow	light color

Figure 1: Graph of an artificially contrived (toy) dictionary. First, vertices "apple", "banana", "tomato" are removed, then "fruit", "red", "yellow" and, finally, "color" and "edible," leaving the dictionary's Grounding Kernel (GK, paler blue), the subgraph induced by {bad, dark, good, light, no, not}. The strongly-connected Kernel Core (KC, darker blue) is {no, not}. The dictionary has eight distinct minimum grounding sets (MGSs), each containing one and only one word from each of the subsets: {bad, good}, {dark, light} and {no, not}.

2.3 Grounding Kernel (GK)

Let $G = (V, E)$ be a directed graph. We say that $U \subseteq V$ is a *grounding set* (also called *feedback vertex set*) of G if $G[V - U]$ is acyclic, i.e. if U covers every cycle of G.

The problem of finding grounding sets of minimum size (MGSs) is NP-complete; hence it is unlikely that one will find an efficient algorithm for solving all instances of this problem. We hope to be able to exploit the particular structure of our graphs to get around this difficulty, and will report on our efforts in a forthcoming paper. Here, we are more interested in extracting hierarchies of definitional distance from dictionary-like graphs.

For this purpose, let $G = (V, E)$ be a directed graph and $\text{SINKS}(G)$ the set of its sinks. We define the operator OUT0 by $\text{OUT0}(G) = G[V - \text{SINKS}(G)]$. We define $\text{OUT0}^n(G)$ as $\text{OUT0}^{n-1}(\text{OUT0}(G))$ for $n \geq 2$, and it is easily verified that there exists some ℓ such that $\text{OUT0}^n(G) = \text{OUT0}^\ell(G)$ for any $n \geq \ell$. Then we define $\text{OUT0}^\infty(G)$ as $\text{OUT0}^\ell(G)$.

Definition 1. *The* grounding kernel *of G is given by $GK = \text{OUT0}^\infty(G)$.*

Note that the grounding kernel (GK) of G is well defined for any graph, even if it is acyclic, since the process of removing sources recursively must stop after a finite number of steps. Moreover, the GK is unique for every graph G. It is also easy to show that every MGS is included in the GK of G. Hence it is of some interest to study the linguistic and cognitive properties of this special set

of words as well as its internal graph structure. For instance, we can decompose it according to its strongly connected components (Subsection 2.4).

2.4 Kernel Core (KC)

We recall two classical relations on vertices. Given two vertices u and v of a graph G, we write $u \to v$ if there exists a uv-path in G and we write $u \leftrightarrow v$ if $u \to v$ and $v \to u$. Note that $\leftrightarrow$ is reflexive, symmetric and transitive so that it is an equivalence relation. Therefore it yields a natural partition of the vertices of G. The equivalence classes of this relation are called the *strongly connected components* of G.

Let G be a graph and $V_1, V_2, \ldots, V_k$ the strongly connected components of G. We construct a graph $G' = (V', E')$ as follows : $V' = \{V_1, V_2, \ldots, V_k\}$ and $(V_i, V_j) \in E'$ if and only if $V_i \neq V_j$ and there exist $u \in V_i$ and $v \in V_j$ such that $(u, v) \in E$. We call this graph the *SCC-quotient-graph*. The next proposition states a well-known fact about this kind of graph.

Proposition 1. *Let G be a graph and G' be the SCC-quotient-graph of G. Then G' is acyclic.* $\qquad\square$

Each acyclic graph induces a partial order on its vertices. In particular, the minimum elements are exactly the sources. They are of particular interest: we define the *kernel core* (KC) as the set of vertices of G belonging to the sources of the SCC-quotient graph G'. In the dictionaries we have been studying, there turns out to be only one source.

3 Hierarchies

The GK of a graph allows us to divide words into two categories: being in the GK or not. However, we would like to refine this division by introducing hierarchies on the vertices, generating more levels. In this article, we consider two hierarchies. The first is induced by the GK and the second is obtained from the strongly connected components.

Let $G = (V, E)$ be a directed graph associated with some dictionary. Let K be its GK. The GK-level of a vertex $v \in V$ with respect to K is defined by

$$L_{\text{GK}}(v) = \begin{cases} 0 & \text{if } v \in K, \\ \max\{L_{\text{GK}}(u) \mid u \in N^-(v)\} + 1 & \text{otherwise.} \end{cases} \tag{1}$$

We will call *GK-hierarchy* the categorization of the vertices of G induced by this level function.

The next hierarchy is based on the strongly connected components. Let G be a graph and G' be its SCC-quotient graph. We define the level $L_{\text{scc}}(v')$ of a vertex v' of G' as follows

$$L_{\text{scc}}(v') = \begin{cases} 0 & \text{if } v' \text{ is a source of } G', \\ \max\{L_{\text{scc}}(u') \mid u' \in N^-(v')\} + 1 & \text{otherwise.} \end{cases} \tag{2}$$

The level $L_{\text{scc}}(v)$ of a vertex v of G is the level $L_{\text{scc}}(v')$ of the strongly connected component v' to which v belongs. This level function is well defined since the graph is acyclic.

The *SCC-hierarchy* is induced by the L_{scc} level function on the vertices of G. In particular, all elements belonging to the same SCC have the same level.

In the next sections, we study three sets of orderings obtained from those two hierarchies: (1) the ordering induced by the GK-hierarchy on the dictionary as a whole; (2) the ordering induced by the SCC-hierarchy on the dictionary as a whole; (3) the ordering induced by the SCC-hierarchy within the GK alone. We include the third hierarchy for a better understanding of the internal structure of the GK.

Example. *Continuing with the graph of Figure 1, Table 1 contains the levels of each word according to both hierarchies. Notice that the words "no" and "not" have level 0 while all the other words u satisfy $L_{\text{scc}}(u) = L_{\text{GK}}(u) + 1$. This is not always the case: this is explained by the fact that the GK of this very small example consists of the KC (corresponding to one connected component) and two other components that are the successors of the GK. In natural language dictionaries, the two level functions may be quite different.*

Word	L_{GK}	L_{SCC}	Word	L_{GK}	L_{SCC}
apple	3	4	good	0	1
bad	0	1	light	0	1
banana	3	4	no	0	0
color	1	2	not	0	0
dark	0	1	red	2	3
edible	1	2	tomato	3	4
fruit	2	3	yellow	2	3

Table 1: Levels corresponding respectively to the GK-hierarchy and to the SCC-hierarchy.

4 Natural Language Dictionaries

We have applied the above hierarchies to the study of two dictionaries: CIDE [8] and LDOCE [7]. As in many natural language processing problems, we are confronted with variation in words' morphosyntactic form and with polysemy (multiple meanings for the same word-form). Morphosyntactic variation was removed using Porter's algorithm [6]. To reduce polysemy we applied a common approximation: we kept only the first definition for each word. Finally, we removed loops and words with empty definitions. The number of vertices, the number of edges and the density of the two dictionaries, along with those of their GK and their "kernel core" (KC, the subgraph induced by their larger strongly connected component, defined below), are represented in Table 2.

Dictionary	CIDE	LDOCE	GK_{CIDE}	GK_{LDOCE}	KC_{CIDE}	KC_{LDOCE}
Number of vertices	19053	23998	1725	2001	1453	1371
Number of edges	221374	239149	20718	20794	16871	14062
Density	0.000610	0.000415	0.00696	0.00519	0.00799	0.00748

Table 2: Main features of the complete dictionaries CIDE and LDOCE, their grounding kernels (GKs) and their kernel cores (KCs).

Like many networks derived from natural models, dictionary graphs satisfy almost all criteria of small-world graphs [11]. The two graphs are sparse and have density lower than 1%. Both dictionaries turn out to contain one single strongly connected subset. Moreover, their in-degree and out-degree distribution seem to follow a normal law and power-law, respectively. However, they also have a large number of small strongly connected components. The size of CIDE's biggest SCC is 1453 and LDOCE's is 1371 (see Table 2), while the remaining SCCs are very small.

A remarkable observation is that the KC of both dictionaries is obtained from only a single source, and that source corresponds to the biggest strongly connected component. Hence, the heart of the cyclic structure of CIDE and LDOCE is found in the KC.

5 Psycholinguistic Correlates of Words in the GK, KC, and Definitional Hierarchies

We analyzed how the structure of dictionary definition space – in terms of GK, the KC, and the two hierarchies of definitional distance that they induce – is related to five psycholinguistic variables:

**AOA: Age of Acquisition** - the age at which a word is learned
C: Concreteness - degree of concreteness/abstractness of word's referent
I: Imageability - how readily one could generate a mental (visual) image of word's referent
**BF: Brown Frequency** - spoken frequency of word
TLF: Thorndike-Lodge Frequency - written frequency of word

The psycholinguistic variables (AOA, C, and I) came from the MRC psycholinguistic database [12]. Because MRC only covered 10% of the words in CIDE (and 8% of LDOCE), we merged each with further databases that were highly correlated with the MRC psycholinguistic database and to increase coverage to 33% for CIDE and 26% for LDOCE. [3] [10]
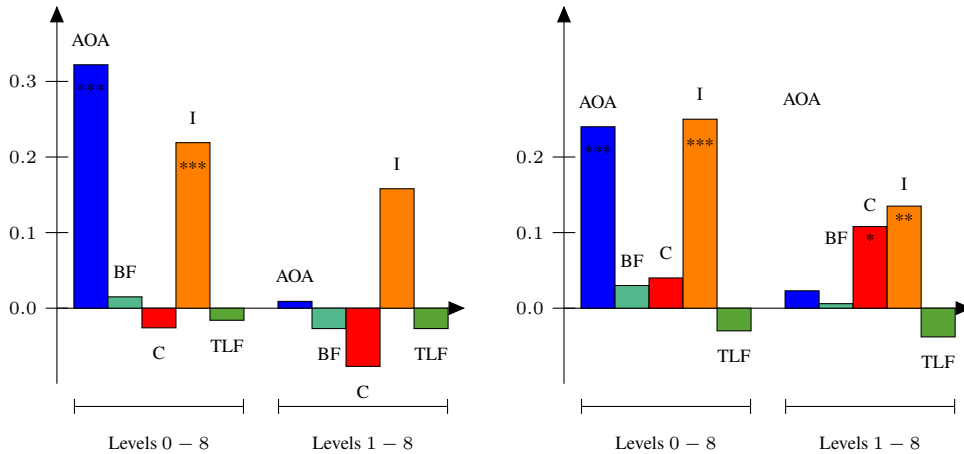
Figure 2: Beta values from multiple regression of the 5 psycholinguistic variables (AOA, C, I, BF, TLF) against the multiple levels of the hierarchies in definition space induced by GK (left) and SCCs (right) (for CIDE; LDOCE pattern was the same). **Left, GK Hierarchy:** For the GK hierarchy, significant effects occur only in the transition from the level 0 – GK itself – to level 1 (Levels 0-8, left): GK words are acquired younger and are more imageable; but if the 0 level (GK) is excluded from the analysis (Levels 1-8, right), there is no longer any significant correlation. **Right, SCC Hierarchy:** For the SSC hierarchy, KC words are acquired younger, less concrete and less imageable; again, the locus of the AOA (age) correlation is just in the transition from level 0 (KC) to level 1, whereas the C and I correlation continues through all the levels. To visualize the levels, see Figures 3 and 4.)

We did statistical analyses of the three hierarchies that were induced (1) by the GK across the entire dictionary, (2) by the SCCs (Strongly Connected Components) across the entire dictionary and (3) by the SCCs within the GK only.
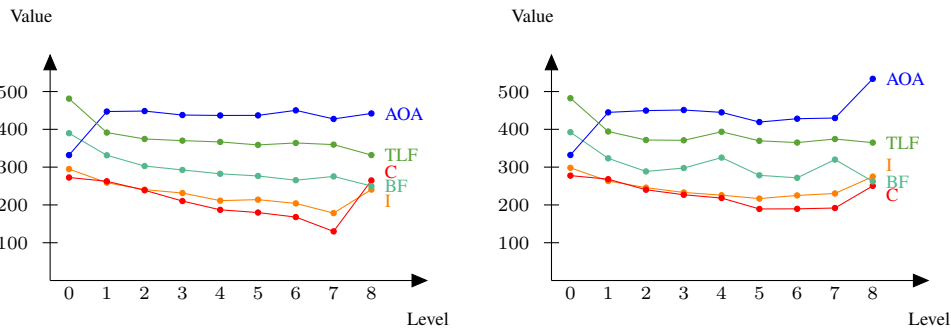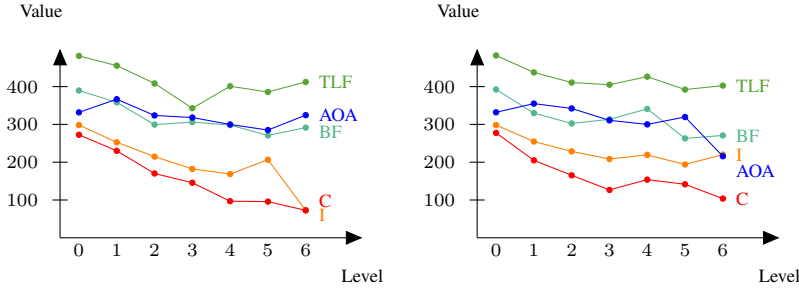

Figure 3: Means for AOA, C, I, BF, and TLF for each level of CIDE (left) and LDOCE (right) with respect to the *SCC-induced hierarchical levels across the entire dictionary*. In the SCC hierarchy, concreteness, image-ability, and both oral and written word frequency decline, whereas after its initial increase from the KC to level 1, age of acquisition stays flat. See text for ANOVA and Factor Analysis

With the GK-induced hierarchy, we did two linear regression analyses. Level 0 words (i.e., those within the GK) were included in the first analysis and excluded from the second. Figure 2 shows that in the first analysis AOA and I are significantly correlated with definitional distance. In the second, no correlation is significant. Hence the only significant effect here is the difference between the GK and the rest of the dictionary: GK words are learned earlier and are more imageable; this effect does not carry over to the higher levels in the induced hierarchy.

In contrast, in the analyses for the SCC-induced hierarchy across the entire dictionary, an ANOVA showed that differences in C and in I at higher levels of the hierarchy are significant too; in post-hoc tests, levels 0, 2, 3 and 4 differed from level 1. The means for each level and each variable are shown in Figure 3.

6

Finally, for the SCC-induced hierarchy within GK alone, AOA and I are significant in the multiple regression (see Figure 4). Moreover, all variables had significant effects in the ANOVA, with the more specific post-hoc tests showing that C and I change from level to level, and AOA differs for most levels.

Figure 4: Means for AOA, C, I, BF and TLF for each level of CIDE (left) and LDOCE (right) with respect to the *SCC-induced hierarchical levels within the GK*. If we induce the hierarchy within the GK alone, the pattern is similar to the external hierarchy: C, I and both frequencies decline with increasing definitional distance from the KC, while AOA (age), after the initial transition from KC to level 1 again remains flat.

Age of acquisition is earlier for level 0 words (the KC) compared to the other levels in the hierarchy induced by the SCCs for the dictionary as a whole. Words in the KC are also more abstract and less imageable than those in the other SCCs. KC words are also used more frequently, both orally and in writing. These results are similar to those reported by [11].

Within the GK alone, words become less concrete, imageable and frequent along the SCC-induced levels starting from KC. Age of acquisition is also significantly older than KC, but only for the first level, after which age remains flat. The correlation between concreteness and age of acquisition is also significantly higher for GK words outside than inside the KC.

These results cast further light on the prior finding of [2] that GK words are learned earlier and more concrete, but that when AOA's covariance with C is partialled out, C's polarity shifts: GK words that are not learned earlier are significantly more abstract than the rest of the dictionary. Our newer finding that the correlation between AOA and C is higher within the GK layer outside the KC suggests that the outer layer may be the locus of this difference. The GK's KC is more concrete and learned earlier, whereas its outer layer is more abstract, and unrelated to age of acquisition.

In the GK-induced hierarchy, only age and imageability were significant. This seemed to contradict our previous findings [2] comparing the GK with the rest of the dictionary, so we reanalyzed this hierarchy excluding its bottom level, the GK. This eliminated all correlations (Figure 2).

6   Concluding Remarks

The GK is a subset of the dictionary with features that differ substantially from the rest of the dictionary. The factors underlying the polarity change in concreteness observed in previous work [2] – with the GK being more concrete and learned earlier than the rest of the dictionary, but more abstract when the covariance with age is partialled out – has now been further refined: It turns out that the GK consists of a large, strongly connected KC plus a smaller, less interconnected outer layer. The KC, like the GK, is more concrete and learned earlier than the rest of the dictionary, and it is also more concrete and learned earlier than the outer layer. But with the KC, when the effects of age of acquisition are partialled out, there is no polarity reversal: The KC remains more concrete than its outer layer and also remains more concrete than the rest of the dictionary. So it is the outer layer that is more abstract than the KC , and hence the polarity reversal is related to the difference between the KC and the outer layer of the GK.

To further refine the differences between the GK's KC, the GK's outer layer and the rest of the dictionary, we induced three orderings in definitional space to produce a graded series of hierarchical levels at an increasing definitional distance from its bottom level or source, to test whether the dichotomous effects based on the GK or the KC extended beyond, along a graded series of definitional

distances. One hierarchy was induced from the GK; the second was induced from the entire dictionary's strongly connected components (SCCs) and the third was induced from the SCCs within the GK alone. The bottom level or "source" of the GK-based hierarchy of definitional distances was the GK itself; the bottom level of both the SCC hierarchies turned out to be the GK's large strongly connected KC. The successive levels of the SCC-induced hierarchy beyond the KC but within the GK and the levels on beyond the GK into the definitional space of the rest of the dictionary both turned out to be significantly correlated with the psycholinguistic variables (age, concreteness, imageability, oral and written frequency): Concreteness and imageability continue to decrease all the way out to the periphery in definition space; oral and written word frequency likewise continue to decrease; but the correlation with age of acquisition is present only in the contrast between the KC and the next level, which is the outer layer of the GK. The effects of SCC-induction for the entire dictionary were also similar to those of [11]: In their small-world analyses too, the KC was acquired earlier then the rest of the corpus and more frequently used (see Summary Figure 5).
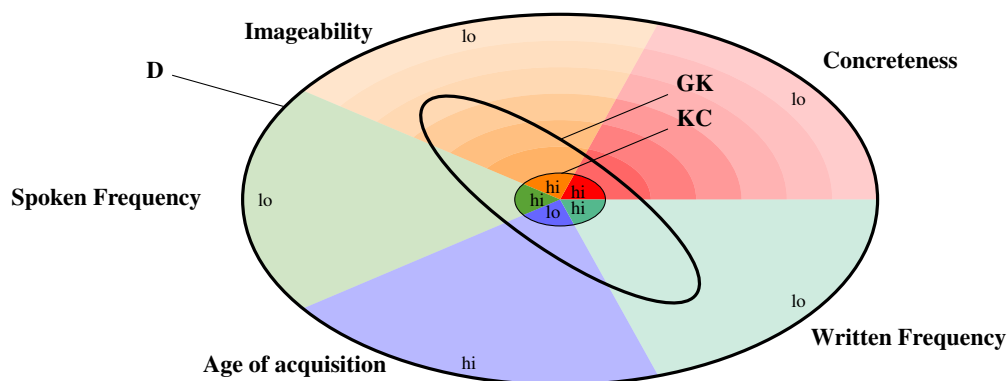
Figure 5: Summary of findings: Words in the Kernel Core (KC) are more concrete, imageable, and frequent, and learned younger. The effect is graded by definitional distance for concreteness and imageability but dichotomous for age and word frequency.

All categories, even the most "concrete" are in fact abstractions, because we must abstract from particular cases, even concrete sensorimotor ones, in order to find the invariant features that distinguish category members from nonmembers and allow us to do the right thing with the right kind of thing. But the more that categories are based on other categories, the more abstract they become, and this is reflected by the distances in our induced definitional space. It is in the nature of words to be amenable to combination and recombination in such a way as to define or describe ever more categories. Defining, like eating, is something we *do*. Our more concrete categories are answerable to the constraints of the sensorimotor world in which they are grounded, but our more abstract categories are increasingly answerable only to combinations of other categories, as we describe and define them. In abstract mathematics, that constraint, though only formal, is still a rigorous one. In more hermeneutic discourse (e.g., constitutional law or theology) the main constraint on words increasingly becomes just other words. Our mental lexicon must encode the meaning of all the words we use in our thought and discourse. Hierarchies in dictionary space may turn out to have counterparts in cognitive space.

References

[1] Blondin Massé, A., Chicoisne, G., Gargouri, Y., Harnad, S., Marcotte, O., Picard, O. *How is meaning grounded in dictionary definitions?*, In TextGraphs '08: Proceedings of the 3rd Textgraphs Workshop on Graph-Based Algorithms for Natural Language Processing, Manchester, United Kingdom, 2008.

[2] Chicoisne, G., Blondin Massé, A., Picard, O., Harnad, S. *Grounding abstract word definitions in prior concrete experience*, 6th Int. Conf. on the Mental Lexicon, Banff, Alberta, 2008.

[3] Cortese, M. J., Khanna, M. M. *Age of acquisition ratings for 3,000 monosyllabic words*, Behavior Research Methods, 40, 791–794, 2008.

[4] Harnad, S. *The symbol grounding problem*, Physica D 42:335-346, 1990

[5] Karp, R.M., *Reducibility among combinatorial problems*. In R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum Press, New York, 1972, pp. 85–103.

[6] Porter M.F., *An algorithm for suffix stripping*, Program, 14(3) 130–137, 1980.

[7] Procter, P., *Longman Dictionary of Contemporary English (LDOCE)*. Longman Group Ltd., Essex, UK, 1978.

[8] Procter, P., *Cambridge International Dictionary of English (CIDE)*. Cambridge University Press, 1995.

[9] Rosen, K.H., *Discrete mathematics and its applications*. 6th ed. McGraw-Hill, 2007.

[10] Stadthagen-Gonzalez, H., and Davis, C. J. *The Bristol norms for age of acquisition, imageability, and familiarity*, Behavior Research Methods, 38, 598–605, 2006.

[11] Steyvers, M. & Tenenbaum, J.B., *The large-scale structure of semantic networks: statistical analyses and a model of semantic growth*, Cognitive Science, 29(1), 41–78, 2005.

[12] Wilson, M., Qx, O.O., Quinlan, P. *MRC Psycholinguistic Database: Machine Usable Dictionary, version 2.00*, 1987.