NRC Publications Archive Archives des publications du CNRC

How language enables abstraction : a study in computational cultural psychology

Neuman, Yair; Turney, Peter; Cohen, Yohai

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1007/s12124-011-9165-8>

Integrative Psychological and Behavioral Science, 46, 2, pp. 129-145, 2012-06-01

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=1844c890-6798-4c19-9cf5-d626b88bc56c>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=1844c890-6798-4c19-9cf5-d626b88bc56c>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

**How Language Enables Abstraction:
A Study in Computational Cultural Psychology**

Integrative Psychological and Behavioral Science (in press)

Yair Neuman, PhD¹, Peter Turney, PhD², & Yohai Cohen³

¹Department of Education

Ben-Gurion University of the Negev

P.O. Box 653, Beer-Sheva 84105, Israel

Tel.: (972-8) 646-1844

E-mail: yneuman@bgu.ac.il

²National Research Council Canada

Ottawa, Ontario, Canada, K1A 0R6

E-mail: Peter.Turney@nrc-cnrc.gc.ca

³Gilasio Coding. Ltd.

Tel-Aviv, Israel

E-mail: yohai@gilasio.com

Abstract. The idea that language mediates our thoughts and enables abstract cognition has been a key idea in socio-cultural psychology. However, it is not clear what mechanisms support this process of abstraction. Peirce argued that one mechanism by which language enables abstract thought is *hypostatic abstraction*, the process through which a predicate (e.g., *dark*) turns into an object (e.g., *darkness*). By using novel computational tools we tested Peirce's idea. Analysis of the data provides empirical support for Peirce's mechanism and evidence of the way the use of signs enables abstraction. These conclusions are supported by the in-depth analysis of two case studies concerning the abstraction of *sweet* and *dark*. The paper concludes by discussing the findings from a broad and integrative theoretical perspective and by pointing to computational cultural psychology as a promising perspective for addressing long-lasting questions of the field.

Keywords: Thought and language; abstraction; hypostatic abstraction; computational cultural psychology

How Language Enables Abstraction: A Study in Computational Cultural Psychology

Introduction

In August 2010 the New York Times published an article titled: "Does your language shape how you think?" (Deutscher, 2010). The article, written by Guy Deutscher, a Professor of Linguistics, urges us to reexamine our "prejudice" against the notion that "language shapes the way we think". It starts with the Whorf hypothesis, according to which different languages impose on their speakers' thoughts in the sense that the lack of certain signs entails the lack of corresponding mental structures. Deutscher rejects this "hypothesis" but argues that on the positive side different languages oblige us to think differently. For instance, while the English language does not compel you to describe the gender of a third party (e.g., "I was speaking with a friend") German obliges you to do so. That is, different languages may oblige you to specify different types of information. The implication of these different specifications to the understanding of different minds is not clear. What does it mean that English and German are different in terms of specifying the gender of a third party or even an object? German may force you to convey information about the gender of your friend but Hebrew does too. Does it mean that the German and the Hebrew "Mind" share similar properties? And in what sense? As there are no clear answers to these questions the soft version of the Sapir-Whorf hypothesis seems to lead to amusing linguistic-anthropological anecdotes and no more.

The problem is deeper (as one may guess) and it relies on the general theoretical framework we choose in order to define and examine the meaning of "thought" and "language", and their relation. It was Vygotsky (1962) who considered the fact that despite their different roots thought and speech interact in their ontogenetic development as two "intersecting circles". Vygotsky located this point of intersection/interaction in "word meaning":

"A word without meaning is an empty sound; meaning, therefore, is a criterion of "word," its indispensable component. ... But from the point of view of psychology, the meaning of every word is a generalisation or a concept. And since generalisations and concepts are undeniably acts of thought, we may regard meaning as a phenomenon of thinking. It does not follow, however, that meaning formally belongs in two different spheres of psychic life. Word meaning is a phenomenon of thought only in so far as thought is embodied in speech, and of speech only in so far as speech is connected with thought and illumined by it. It is a phenomenon of verbal thought, or meaningful speech – a union of word and thought" (Chapter 7. I).

The idea of *word meaning* as the unit of interaction is highly important for the thesis presented in this paper and it will be used as our own unit of analysis. While Vygotsky is the one whose name popped into our mind in all discussions of "Thought" and "Language" and while Sapir and Whorf are the two people who are usually associated with the linguistic-relativity "hypothesis", the most important essay written about the subject is probably Volosinov's "Marxism and the Philosophy of Language" (1986). This remarkable essay proposed a totally different approach to the study of language and thought. According to Volosinov: "Outside the material of signs there is no psyche" (1986, p. 26). In other words, the sharp differentiation between the realm of signs (i.e., Language) and the realm of "Psyche" (i.e., Thought) should be replaced by a semiotic perspective according to which "Psychic experience is the semiotic expression of the contact between the organism and the outside [and the Inside] environment" (1986, p. 26).

According to this perspective, the simple Newtonian causal relation between "Language" and "Thought" is meaningless and the mind should be studied as a "semiotic interface" (Neuman, 2003) that human beings, as well as other intelligent systems, use to

make sense of their world (Neuman, 2008). This perspective is in line with the Whorf theory, according to which:

"There is little point in arguing about whether language influences thought or thought influences language for the two are functionally entwined to such a degree in the course of individual development that they form a highly complex, but nevertheless systematically coherent, mode of cognitive activity which is not usefully described in conventionally dichotomizing terms as either 'thought' or 'language'." (Lee, 1996, p. xiv).

Along this line, one may ask how language, in the Saussurean sense of a sign system (Saussure, 1973; Tobin, 1990), may become more and more "abstract". In this context, the enigmatic notion of abstraction may be explained as follows. We know that certain signs have a concrete denotational sense that is grounded in our sensorimotor experience: the sign *dark* concerns a certain visual experience shared by people across cultures and so is "sweet" and "bitter". However, there are other signs that are more abstract in the sense that we cannot easily trace their embodied origin, "God" for instance or even the connotation of *dark* in the word-pair "dark thoughts".

While people may share the basic experience of being in the world due to their shared physiology, the more abstract our sign system becomes the more evident are cultural differences that load our abstract signs with meaning. In other words, a higher level of abstraction is evident when the concrete reference of the sign cannot be easily traced and when it is loaded with connotations or metaphors that increase its variance across speech communities. Abstract signs may expose cultural, historical, or national differences through which the abstract notion of our sign system is constituted. How does the mind, as a social interactive and dynamic semiotic system, support us in performing this transformation? One interesting suggestion comes from C. S. Peirce, who is seldom if ever mentioned in this context despite his clear relevance to "embodied cognitive science" (Clark, 2006).

Peirce on Hypostatic Abstraction

To understand Peirce's notion of hypostatic abstraction, we should be familiar with his theory of relations. According to Peirce there are three basic relational types that correspond to his three categories of being: Firstness, Secondness, and Thirdness. Peirce describes these categories as Modes of Being. In what sense are those "Categories" "Modes of Being"? Peirce answers this question by saying:

"Therefore, we do not ask what really is, but only what appears to everyone of us in every minute of our lives. I analyze experience, which is the cognitive resultant of our past lives, and find in it three elements. I call them *Categories*." ('Minute Logic', CP 2.84, c.1902)

In other words the Categories are forms of experience as it is evident to human beings. Firstness is the "mode of being of that which is such as it is, positively and without reference to anything else" ("A letter to Lady Welby," *Collected Papers of Charles Sanders Peirce* [CP] 8.328, 1904); it is the "qualities of feeling" ("A letter to Lady Welby," CP 8.329, 1904). For example, consider the basic experience of sweet taste. Firstness is the category of being that is the closest to our embodied experience. Secondness "consists in one thing acting upon another" ("A letter to Lady Welby," CP 8.330, 1904). It is "the mode of being of that which is such as it is, with respect to a second but regardless of any third" ("A letter to Lady Welby," CP 8.328, 1904). For example, the category of *causality* actually involves Secondness as one variable (e.g. infection) influences another variable (e.g. body temperature). Thirdness is "mental or quasi-mental influence of one subject on another relative to a third" ("Pragmatism," CP 5.469, 1907). It is the "mode of being of that which is such as it is, in bringing a second and third into relation to each other" ("A letter to Lady Welby," CP 8.328, 1904). For instance, sign-activity is the expression of Thirdness as it involves the sign, the signified (i.e. object) and the mind it affects, namely the "Interpretant".

In sum, Peirce's categories involve the level of basic qualities (Firstness), a dyadic

relation between two objects, and a triadic form in which each element cannot be considered separately from the triadic whole of which it is a part. With this typology in mind, we may turn to Peirce's theory of relations (*CP* 5.119). What is the difference between the Categories and the Relations? Peirce's typology of relations is isomorphic to his categories but seems to emphasize another aspect. While his categories, as influenced by Kant, aim to represent universal forms of "experience" his relations seem to represent the "Valence" of the categories. In Chemistry "Valence" is a term used to indicate the number of chemical bonds formed by the atoms of a given element. Peirce's typology of relations emphasizes the three general forms of "Bonding" that constitute our forms of experience. In other words, while the categories demarcate our experience according to three abstract forms, the typology of relations characterizes the nature of the forms through the "bond" they create between the basic elements of our experience.

A monadic relation is a fact about a single object. Saying that the "cat is black" or that "I'm afraid" (or, more accurately, the feeling of being frightened) is an expression of a monadic relation. A monadic relation is thus not really a relation but a basic quality of feeling that may be expressed linguistically through an adjective or adverb. It is a "relation" between the object and itself. A dyadic relation is a fact about two things or objects. "Honey possesses sweetness" is an example of a dyadic relation of "possession" between two objects: Honey and Sweetness. A triadic relation is a relation among three objects, such as "A gave B to C". A triadic relation cannot be reduced to two dyadic relations and still preserve its meaning as a whole. For instance, the meaning of the utterance: "Danny sold the book to Jerry" cannot be reduced to "Danny sold the book" and "Book to Jerry". At this point we may move forward to "Hypostatic abstraction".

Peirce proposed Hypostatic Abstraction to be a crucial mechanism in abstraction ("The simplest mathematics," *CP* 4.235, *CP* 4.227–4.323). The idea of hypostatic abstraction

is that there is a procedure that *converts a quality expressed as an adjective or predicate into an additional object*. For example, the expression “honey is sweet” may be converted to “honey possesses sweetness”. The transformation is actually a transformation from a monadic relation to a dyadic relation. This transformation results in a *reification* of the basic quality such as in the transformation from *sweet* to *sweetness*. This transformation involves abstraction as *sweet* turns from a *quality* into an *object-in-relation*. Due to the unique ability of human beings to manipulate *objects* through orchestrated and sign-mediated social activity (Vauclair, 2003), it is clear why the mechanism of hypostatic abstraction is necessary for transcending embodied experience. For instance, while the predicate *sweet* is firmly grounded and constrained by our experience, *sweetness* as an object may be linked with a variety of associated senses, connotations, and metaphors. Therefore, it becomes functionally independent of its embodied experience and a sign arbitrarily associated with its origin and material cloth, a property that uniquely characterizes the human language (Saussure, 1973; Vauclair, 2003). In sum, turning a quality into an object involves its abstraction and relocation in a relational network of signs. Whether this idea is empirically grounded is an open question. In other words, it is not clear whether Peirce's idea can be supported through empirical evidence as collected and analyzed by using well-established scientific methods. In this sense, Peirce notion of "Hypostatic abstraction" is an "abduction", an "open question" or hypothesis that should be tested empirically. The aim of this paper is to examine whether Peirce's idea is grounded in our sign-use and to initiate the first steps in elucidating its dynamics.

Methodology

If Peirce's idea is empirically grounded, then we should expect that when examining word-pairs of the type X-Xness (e.g., sweet-sweetness), we should find that the right term, the reified noun, is more "abstract" than the left term – the predicate/adjective. To test this

hypothesis, we used recent advancements in Computational Linguistics for measuring the abstractness/concreteness of a word. In the following section, "Measuring Abstractness and concreteness", we describe our algorithm for rating words according to their degree of abstractness. The algorithm draws on a methodology for measuring the semantic similarity of words (explained in detail in the subsection "measuring semantic similarity") and involves the comparison of a given word to paradigmatic abstract and concrete words (details in the subsection "Using semantic similarity to measure abstractness"). In the subsequent section, "Testing the Hypothesis", we show how this algorithm can be used to test Peirce's hypothesis of hypostatic abstraction.

Measuring Abstractness and Concreteness

Concrete words refer to things, events, and properties that we can perceive directly with our senses, such as *trees*, *walking*, and *red*. Abstract words refer to ideas and concepts that are distant from immediate perception, such as *economics*, *calculating*, and *disputable*. In this section, we describe an algorithm that can automatically calculate a numerical rating of the degree of abstractness of a word on a scale from 0 (highly concrete) to 1 (highly abstract). For example, the algorithm rates *purvey* as 1, *donut* as 0, and *immodestly* as 0.5.

The algorithm is a variation of Turney and Littman's (2003) algorithm that rates words according to their semantic orientation. Positive semantic orientation indicates praise (*honest*, *intrepid*) and negative semantic orientation indicates criticism (*disturbing*, *superfluous*). The algorithm calculates the semantic orientation of a given word by comparing it to seven positive words and seven negative words that are used as paradigms of positive and negative semantic orientation:

Positive paradigm words: *good*, *nice*, *excellent*, *positive*, *fortunate*, *correct*, and *superior*.

Negative paradigm words: *bad*, *nasty*, *poor*, *negative*, *unfortunate*, *wrong*, and

inferior.

Turney and Littman (2003) chose these paradigm words manually, using their personal intuition about which words best convey positive and negative semantic orientation. Here we calculate the abstractness of a given word by comparing it to twenty abstract words and twenty concrete words that are used as paradigms of abstractness and concreteness. Unlike Turney and Littman (2003), we selected these forty paradigm words automatically, as described in the subsection "Using Semantic Similarity to Measure Abstractness".

Turney and Littman (2003) experimented with two measures of semantic similarity, pointwise mutual information (PMI) (Church & Hanks, 1989) and latent semantic analysis (LSA) (Landauer & Dumais, 1997). These measures take a pair of words as input and generate a numerical similarity rating as output.

The semantic orientation of a given word is calculated as the sum of its similarity with the positive paradigm words minus the sum of its similarity with the negative paradigm words. Likewise, here we calculate the abstractness of a given word by the sum of its similarity with twenty abstract paradigm words minus the sum of its similarity with twenty concrete paradigm words. We then use a linear normalization to map the calculated abstractness value to range from 0 to 1.

Our algorithm for calculating abstractness uses a form of LSA to measure semantic similarity. This is described in detail in the section titled: "Measuring semantic similarity".

Measuring semantic similarity

The variation of LSA that we use here is similar to Rapp's (2003) work. We modeled our similarity measure on Rapp's due to the high score of 92.5% that he achieved on a set of 80 multiple-choice synonym questions from the Test of English as a Foreign Language (TOEFL). The core idea is to represent words with vectors and calculate the similarity of two words by the cosine of the angle between the two corresponding vectors. The values of the

elements in the vectors are derived from the frequencies of the words in a large corpus of text. This general approach is known as a Vector Space Model (VSM) of semantics (Turney & Pantel, 2010).

We began with a corpus of 5×10^{10} words (280 gigabytes of plain text) gathered from university websites by a webcrawler.¹ We then indexed this corpus with the Wumpus search engine (Büttcher & Clarke, 2005).² We selected our vocabulary from the terms (words and phrases) in the WordNet lexicon.³ By querying Wumpus, we obtained the frequency of each WordNet term in our corpus and selected all terms in our corpus with a frequency of 100 or more. This resulted in a set of 114,501 terms. Next we used Wumpus to search for up to 10,000 phrases per term, where a phrase consists of the given term plus four words to the left of the term and four words to the right of the term. These phrases were used to build a word–context frequency matrix $\mathbf{F}$ with 114,501 rows and 139,246 columns. A row vector in $\mathbf{F}$ corresponds to a term in WordNet and the columns in $\mathbf{F}$ correspond to contexts (the words to the left and right of a given term in a given phrase) in which the term appeared.

The columns in $\mathbf{F}$ are unigrams (single words) in WordNet with a frequency of 100 or more in the corpus. A given unigram is represented by two columns, one marked *left* and one marked *right*. Suppose r is the term corresponding to the i -th row in $\mathbf{F}$ and c is the term corresponding to the j -th column in $\mathbf{F}$. Let c be marked *left*. Let f_{ij} be the cell in the i -th row and j -th column of $\mathbf{F}$. The numerical value in the cell f_{ij} is the number of phrases found by Wumpus in which the center term was r and c was the unigram closest to r on the left side of r . That is, f_{ij} is the frequency with which r was found in the context c in our corpus.

A new matrix $\mathbf{X}$, with the same number of rows and columns as in $\mathbf{F}$, was formed by calculating the Positive Pointwise Mutual Information (PPMI) of each cell in $\mathbf{F}$ (Turney &

1 The corpus was collected by Charles Clarke at the University of Waterloo.

2 Wumpus is available at <http://www.wumpus-search.org/>.

3 WordNet is available at <http://wordnet.princeton.edu/>.

Pantel, 2010). The function of PPMI is to emphasize cells in which the frequency f_{ij} is statistically surprising, and hence particularly informative. This matrix was then smoothed with a truncated Singular Value Decomposition (SVD), which decomposes $\mathbf{X}$ into the product of three matrices $\mathbf{U}_k \Sigma_k \mathbf{V}_k^T$. Finally, the terms were represented by the matrix $\mathbf{U}_k \Sigma_k^p$, which has 114,501 rows (one for each term) and k columns (one for each latent contextual factor). The semantic similarity of two terms is given by the cosine of the two corresponding rows in $\mathbf{U}_k \Sigma_k^p$. For more detail, see Turney and Pantel (2010).

There is a complex many-to-many mapping between words and meanings: A word may have many meanings, depending on context, and many words may have the same meaning. Unfortunately, we do not have direct access to meanings; we must infer meanings from word usage. Latent Semantic Analysis uses statistical techniques to attempt to infer meaning from patterns of word usage in a large collection of texts (i.e. a corpus). In this context, there are two parameters in $\mathbf{U}_k \Sigma_k^p$ that need to be set. The parameter k controls the number of latent factors (hidden units of meaning) in the LSA model of a corpus. The parameter p determines the weight of each latent factor (the importance of each hidden unit of meaning in the model) by raising the corresponding singular values in Σ_k^p to the power p . The parameter k is well-known in the literature on LSA, but p is less familiar. The use of p was suggested by Caron (2001). Based on our past experience, we set k to 1000 and p to 0.5. We did not explore any alternative settings of these parameters for measuring abstractness.

Using semantic similarity to measure abstractness

Now that we have $\mathbf{U}_k \Sigma_k^p$, all we need in order to measure abstractness is some paradigm words. Although Turney and Littman (2003) manually selected their fourteen paradigm words, here we use a supervised learning algorithm to choose our forty paradigm words. We used the MRC Psycholinguistic Database Machine Usable Dictionary (Coltheart,

1981) to guide our search for paradigm words. The MRC Psycholinguistic Database Machine Usable Dictionary (Coltheart, 1981) contains 4,295 words rated with degrees of abstractness.⁴ We used half of these words to train our supervised learning algorithm and the other half to validate the algorithm. On the testing set, the algorithm attains a correlation of 0.81 with the dictionary ratings. This indicates that the algorithm agrees well with human judgments of the degrees of abstractness of words.

We split the 4,295 MRC words into 2,148 for training (searching for paradigm words) and 2,147 for testing (evaluation of the final set of paradigm words). We began with an empty set of paradigm words and added words from the 114,501 rows of $\mathbf{U}_k \Sigma_k^p$, one word at a time, alternating between adding a word to the concrete paradigm words and then adding a word to the abstract paradigm words. At each step, we added the paradigm word that resulted in the greatest correlation with the ratings of the training words. This is a form of greedy forward search without backtracking. We stopped the search after forty paradigm words were found, in order to prevent overfitting of the training data.

Table 1 shows the forty paradigm words and the order in which they were selected. At each step, the correlation increases on the training set, but eventually it must decrease on the testing set. After forty steps, the training set correlation was 0.8600. At this point, we stopped the search for paradigm words and calculated the testing set correlation, which was 0.8064. This shows a small amount of overfitting of the training data.

For another perspective on the performance of the algorithm, we measured its accuracy on the testing set by creating a binary classification task from the testing data. We calculated the median of the ratings of the 2,147 words in the test set. Every word with an abstractness above the median was assigned to class 1 and every word with an abstractness below the median was assigned to class 0. We then used the algorithm to guess the rating of

4 The dictionary is available at <http://ota.oucs.ox.ac.uk/headers/1054.xml>.

each word in the test set, calculated the median guess, and likewise assigned the guesses to classes 1 and 0. The guesses were 84.65% accurate. This procedure validates the abstraction rate.

After generating the paradigm words with the training set and evaluating them with the testing set, we then used them to assign abstractness ratings to every term in the matrix. The result of this is that we now have a set of 114,501 terms (words and phrases) with abstractness ratings ranging from 0 to 1.⁵ Based on the testing set performance, we estimate these 114,501 ratings would have a correlation of 0.81 with human ratings and an accuracy of 85% on binary (*abstract* or *concrete*) classification.

Table 1

The forty paradigm words and their correlation, on the training set.

Concrete Paradigm Words			Abstract Paradigm Words		
Order	Word	Correlation	Order	Word	Correlation
1	donut	0.4447	2	sense	0.6165
3	antlers	0.6582	4	indulgent	0.6973
5	aquarium	0.7150	6	bedevil	0.7383
7	nursemaid	0.7476	8	improbable	0.7590
9	pyrethrum	0.7658	10	purvey	0.7762
11	swallowwort	0.7815	12	pigheadedness	0.7884
13	strongbox	0.7920	14	ranging	0.7973
15	sixth-former	0.8009	16	quietus	0.8067
17	restharrow	0.8089	18	regularisation	0.8123

⁵ A copy of the 114,501 rated terms is available on request from Peter Turney.

19	recorder	0.8148	20	creditably	0.8188
21	sawmill	0.8212	22	arcella	0.8248
23	vulval	0.8270	24	nonproductive	0.8299
25	tenrecidae	0.8316	26	couth	0.8340
27	hairpiece	0.8363	28	repulsion	0.8400
29	sturnus	0.8414	30	palsgrave	0.8438
31	gadiformes	0.8451	32	goof-proof	0.8469
33	cobbler	0.8481	34	meshuga	0.8503
35	bullet	0.8521	36	dillydally	0.8538
37	dioxin	0.8550	38	reliance	0.8570
39	usa	0.8585	40	lumbus	0.8600

Testing the Hypothesis

For testing our specific hypothesis, we automatically identified in the above list of words that were rated according to their level of abstraction, 1078 word-pairs of the form X - X_{ness} , for instance sweet-sweetness". Let $\text{abs}(X)$ be the abstractness level of word X . Our hypothesis is that $\text{abs}(X_{ness})$ is significantly higher than $\text{abs}(X)$. Let $\text{delta}(X) = \text{abs}(X_{ness}) - \text{abs}(X)$. Our hypothesis is that $\text{delta}(X)$ is significantly higher than zero. To test this hypothesis we used t-test for paired samples. The delta was found statistically significant ($t = 4.914, p < .001$).

We calculated the difference between the abstraction level of X_{ness} and X . If the difference was 0 or less then we scored the word-pair "0", and "1" otherwise. Fifty-nine percent of the word pairs scored "1" and the rest "0". By using the Sign-Test this result was

found statistically significant ($p < .001$). Like the previous result, this result too supports Peirce's idea.

We noted above that the Sign of $\text{delta}(X)$ is positive in 59% of the 1078 pairs. This is statistically significant, but it raises the question of whether there is an explanation for why $\text{delta}(X)$ is negative in some cases. We hypothesize that $\text{delta}(X)$ is more likely to be positive when X is a high frequency word, because the more a word is used, the more opportunities there are for the word to become an abstract object, through associations, connotations, and metaphors. To test this hypothesis, we split our data into word pairs in which the difference between X_{ness} and X was $\geq .10$ (one Standard Deviation of the Abstraction List) and word pairs in which the difference was 0 or less. We titled the first set "High Abstraction" and the second "Low Abstraction". If our hypothesis is correct, then the average frequency of X on the High abstraction set should be significantly greater than the average frequency of X on the Low Abstraction set. On the other hand, *there should be no difference* between the groups with regard to the frequency of X_{ness} .

To test this hypothesis, we used One Way ANOVA and found a statistically significant difference between the groups ($F(1, 1065) = 66.681, p < .001$). The average frequency of X in the High Abstraction set was statistically higher than the frequency of X in the Low Abstraction set (Mean = 22011 with 95% Confidence Interval of ± 645 versus Mean = 15380 with 95% Confidence Interval of ± 503 respectively).

In line with our hypothesis, no statistically significant difference was found between the groups on the average frequency of X_{ness} .

These results suggest that, as hypothesized by Peirce, the transformation of a word from a predicate into an object, from a monadic embodied quality into a dyadic *object-in-relation*, is associated with an increase in the level of abstraction. Furthermore, it seems that this process is motivated by language-use, which is probably the driving force behind the

reification of the predicates. Predicates used more by the sign community may contribute to the emergence of more abstract objects. To better understand this dynamic, we move to the next phase of our study, which involves the in-depth analysis of two case studies.

The Lexical Context of the Words

In order to better understand the bag-of-words associated with each sign in our *Xness* list, we used the word-context frequency matrix F , described in the previous section titled "Measuring semantic similarity".

For each predicate in our word-pairs (e.g., *dark*), we identified 50 words that tend to co-occur with the given predicate by examining the predicate's corresponding row vector in F . The 50 largest values in this row vector select 50 columns in F . The corresponding words for these 50 columns are the 50 words that co-occur most frequently with the given predicate. These are words that are "primed" by the given predicate: When we hear the predicate, these other words come to mind. By reversing this process, we can also find 50 words that prime the given predicate: When we hear one of these words, the given predicate comes to mind. We call these 100 words the "context" for the given predicate and use them for further analysis. The unique bag-of-words produced by this procedure for each primed word was defined as its "*context*" and was used for further analysis.

For each of the 1078 *X-Xness* pairs, we identified the context of *X*, the context of *Xness*, and the context of their intersection (e.g., words shared by "Sweet" and "Sweetness"). Moreover, we identified the complement context of *X* (i.e., words that exist in *X* but not in the intersection with *Xness*) and the complement of *Xness* (i.e., words that exist in the context of *Xness* but not in the intersection with *X*).

Home sweet home: Analyzing the transformation of "Sweet" to "Sweetness"

In order to dig deeply into the dynamics of hypostatic abstraction, we use two case studies. The first case is the word pair *sweet-sweetness* and the second pair is *dark-darkness*.

We chose to focus our analysis on the pair *sweet-sweetness* as it was used by Peirce to illustrate his idea of hypostatic abstraction. The pair *dark-darkness* was chosen because of the negative associations of dark color in many cultures specifically the negative associations of *darker* skin color in cultures even in cultures in which the people are characterized as having dark skin. In order to understand the transformation and the abstraction of *sweet*, we first applied the Linguistic Inquiry and Word Count Software [www.liwc.net] to study the word categories characterizing the context of *sweet*, *sweetness*, and their intersection. LIWC is a text analysis software program that calculates the degree to which people use different categories of words. The software includes a dictionary of words sorted out by human judges into different categories. The software works by simply segmenting the text and categorizing its words into the predefined categories. It is a powerful text analysis program that has been validated and intensively used in psychological research (Pennebaker & Francis, 2001; Tausczik & Pennebaker, 2010). The results of the LIWC analysis appear in Table 2.

Table 2

LIWC analysis of the context of "Sweet" and "Sweetness".

Category	Sweet_Com	Sweet	Inter	Sweetness	Sweetness_Com
Social	2.47	2.93	16.67	5.45	2.38
Affect	12.9	12.77	8.33	34.55	42.86
Positive Emo.	9.07	9.04	8.33	34.55	35.71
Negative Emo.	3.57	3.46	0	5.45	7.14
Bio	7.14	8.24	41.67	14.55	7.14
Ingest	3.57	4.26	25	9	4.76
Friend	0.27	0	8.33	0	0

* Results are in percentages. Sweet_Com: Complement of sweet, Inter: Intersection of the contexts.

We may consider the intersecting context as a bridge between *sweet* and *sweetness*. Several interesting findings pop up from a first glance at the table. The first finding concerns Emotions. The Affect associated with *sweetness* is greater than the one associated with *sweet*. This is mainly due to the difference in positive emotion associated with *sweetness*. This increase in positive emotions is mediated at the intersection by words in the category of social interaction (*Social* and *Friend*) and by words in the categories of *Biology* and *Ingestion*. This categorization does not tell us exactly how *sweet* is transformed into *sweetness*. To better understand this process, we identified the most concrete and frequent words associated with *sweet*. These words are: bread, girl, cherry, tea, recipe, corn, honey ... Not surprisingly, *sweet* is mainly associated with sweet food. Eating and tasting sweet food is the denotational source of *sweet*. However, the most *abstract and frequent* words associated with *sweet* were: love, life, success, taste, and surprise. While examining the most abstract and frequent words at the intersection of *sweet* and *sweetness* we found *love* and *life*. This finding points to the link between the basic experience of sweet food, and the psychological emotional attachment to the nurturing mother or father in terms of love. Love is basically associated with those who gave us food and therefore life.

We may ask what are the most abstract and frequent words associated as *complementary* of *sweetness*—the words that appear in the context of *sweetness* but *do not exist* at the intersection of *sweet* and *sweetness*. Among these words we find in descending order: *good*, *give*, and *hope*. This finding allows us to trace the emergence of *sweetness* from the basic embodied experience of *sweet*. As we are nurtured by our mother or father the basic interaction is associated on the abstract level of analysis with *love* and *life*. These abstract concepts are carried and propagated to *sweetness* through the platform of *biological* and *ingestion* words, mainly associated with food, and through the field of personal interaction signaled by the *friendship* and *social* category. This trajectory clearly illustrates the way

emotion is deeply associated with sense making through semiotic mediation (Salvatore & Venuleo, 2008).

One should remember though that, as argued by Adorno (in a letter to Walter Benjamin), “all reification is a forgetting”. We can see that the words associated with our concrete embodied experience appear to a significantly lesser degree in the context of *sweetness*. In a sense, they are "forgotten" in favor of abstract connotations. To prove this hypothesis, we measured the average level of abstraction of words at the intersection of the word-pair. The average abstraction levels of words in the context of sweet, sweetness, and their interaction were 0.47, 0.58, & 0.54, respectively. These results support the above interpretation. The associations of sweet are "grafted" to the "semiotic field" (Neuman & Nave, 2009) of *sweetness* and load this object with a high level of positive emotion and with words that extend its denotational source such as *good* and *hope*. Indeed, the abstract notion of *hope* is deeply rooted in our basic attachment with significant others and the expectation that love and life will be materialized through the interaction of nurturing us. To recall, the ancient Greeks metaphorically described hope as the "Exiles' bread". Unconsciously, they have made the association between expecting food and expecting a better future – hope. The above interpretation, powerfully supported by our computational psychological cultural analysis, also explains why *sweet* and specifically sweet food is deeply associated with love. The fact that chocolate is one of the favorite gifts on Valentine’s Day, is just another indication of this deep rooted dynamic.

Analyzing the transformation of *Dark* to *Darkness*

The context of *dark* is quite simple to understand if we examine its affective context.

Table 3

LIWC analysis of Dark and Darkness

Category	Dark	Dark_Com	Inter	Drkness	Darkness_Com
----------	------	----------	-------	---------	--------------

Social	.39	.42	0	2.4	3.23
Affect	7.28	7.31	6.45	8.8	9.68
Positive Emotion	2.56	2.71	0	1.6	2.15
Negative Emotion	4.53	4.38	6.45	7.2	7.53
Anxiety	0.98	1.04	0	1.63	2.15
Anger	0.98	1.04	0	1.63	2.15
Sad	1.77	1.46	6.45	2.44	1.08

We can see that the shift from *dark* to *darkness* involves an increase in negative emotions, specifically anxiety, and in social words. This transformation is mediated at the intersection through *sadness*. While Affect is reduced at the shared context of *dark* and *darkness*, this is mainly due to the disappearance of positive emotions and the increase of negative emotions. This is a trivial finding, but the way it is created is less trivial. The most concrete and frequent words associated with *dark* are *hair*, *skin*, and *eyes*. At the intersection of *dark* and *darkness*, we find two keywords, *deep* and *despair*. The dark color associated with body parts and closed and deep places is associated with the horror of the abyss and death as evident by *morbid* – the most abstract word in the context of *dark*, and in the intersection of *dark* and *darkness*. *Darkness* is associated through the lack of light with death and morbidity: trembles, battle, weeping, evil, and sin are some of the words that appear as complementary of *darkness*. The basic horror of being in a place with no light (i.e., deep, room, abyss) is propagated to *darkness* and associated with *evil*. This propagation is evident in the increase of negative emotions and may be traced to a positive feedback loop involved in the grafting procedure that starts from our basic embodied experience (Neuman, 2009). Bad things are dark things associated with fear and propagated to the abstract realm of evil. In this context, the unfortunate negative associations of dark skin, hair, and eyes can be easily

traced to our fear of dark places and the hypostatic abstraction that transforms this concrete fear into abstract concepts such as *evil* and *sin* and back into their concrete origin – the dark skin, eyes, and hair.

Interestingly, the dynamics described above may help us to trace the movement from "denotation" to "connotation" and "metaphor". While *dark* may be used as a denotation to point at the color of certain objects, it is also used as a connotation in the context of "dark thoughts", for instance. This transformation is comprehensible if we identify the abstract terms associated with the complement of *darkness*. "Dark thoughts" mean thoughts of *evil* and *sin*, as identified by our analysis, and these connotative senses result from grafting our basic fears from the denotations toward *darkness* as an object that stands in certain relations with other objects. In the case of *darkness* it is possible to identify these relations by using Corpus Analysis.

For gaining more insight into the relations in which *darkness* is embedded we use the Corpus of Contemporary American English (COCA) [<http://corpus.byu.edu/coca/>] (Davies, 2008). Through COCA, we identified the verb form collocated in a window + four words to the right of *darkness*. The results were sorted according to frequency and minimum Mutual Information equal to 3. We find that *darkness*: *descends*, *envelops*, *engulfs*, *enfolded*, *gropes*, *abides*, *gorged*, *extinguishing*, *cloaking*, *dwelled*, and *shrouds*. In metaphorical terms, *darkness* is like the mythical predator or the abyss that might engulf us in death. Portraying darkness in this way is possible through the grafting procedure described above.

In sum, the abstraction of *dark* starts from simple association with the lack of light and the horror associated with this experience. This horror is the source of the metaphorical/abstract sense of *darkness* as an engulfing substance. *Darkness* is therefore loaded with our fear of being swallowed by a primordial form, a fear that echoes in our myths, psychological theories, and stereotypes. For instance, an ancient myth that appears in

the Veda tells us that Prajapati, the androgynous being and the primordial lord of creatures, felt an enormous emptiness when he gave birth to the world. This horrible emptiness caused Prajapati to re-unite with his offspring by swallowing the newborn. When the newborn saw the empty mouth of his father he shouted in horror. This horror illustrates the fear of being swallowed into the darkness of death, of losing one's individuality and autonomy. This idea clearly echoed in the writings of Melanie Klein where the infant aggressive fantasies of consuming from the mother's breast may symmetrically turn on him as the object consumed by the mother. This idea may be further developed but as it is clearly beyond the major scope of the paper, we may settle with the above examples and reference to the psychoanalytic school of object-relations.

Discussion

Due to the lack of appropriate computational tools, the complex relation between language and thought has been too easily dismissed and misrepresented in the psychological literature. By using state of the art tools in Computational Linguistics, we can dwell deeply into this complex phenomenon. The semiotic perspective presented in this paper resonates with the notion of "embodied cognitive science" (Clark, 2006) in which the mind is in a constant flux that is coordinated by the use of words. This approach, which defies "any simple logic of inner versus outer, or of tool versus user" (ibid., p. 374), allows us to see the continuity of the conceptual system across species (Barsalou, 2005) and to understand the way human beings may transcend their grounding in the perceptual realm (Pezzulo & Castelfranchi, 2007) through sign-mediated interactions with others. According to this perspective, our sign system, which is social by nature (Saussure, 1973), starts with referential denotational signs to the concrete reality/experience, whether external or internal reality (Danesi, 2003). In contrast with animal communication, human language "deals mainly with entities" (Vauclair, 2003, p. 12) rather than with holistic situations (e.g., the

presence of a predator). Therefore, our language is a language of "reification" (Neuman, 2003), as it allows us to "crystallize" the dynamic flux of sensorimotor experience into objects of reflection and contemplation (i.e., signs). Through experimentation with the sign system and mechanisms such as hypostatic abstraction, some of the signs may detach from their concrete embodiment. The fact that signs are primarily social tools of interaction explains the abstraction of the mind as a process of internalization (Lawrence & Valsiner, 1993). In human beings, words and gestures function not primarily to obtain a result in the physical world, but "to direct another individual's attention (its mental state) to an object or event as an end in itself." (Vauclair, 2003, p. 14). This object-oriented behavior develops differently among human beings and primates:

"A typical form of communication between the infant and a competent adult (e.g., the mother) arises in human during object manipulation. This form is characterized by the mutual exchange between mother and infant regarding a large variety of discrete, movable objects ... By contrast, the nonhuman primate mother does not appear to intervene directly in the infant's object manipulations" (Vauclair, 2003, p. 20).

In other words, we think through the internalization of symbolic interactions with others as we relate with them to objects in the world. These designated objects and the social-semiotic interaction through which they are enacted are internalized and used for autonomously reflecting on our sign system as an object in itself. Our thoughts become more abstract as the signs we use to reflect on the external and internal reality turn far from their embodied origin. This thesis may explain why the brain as a physical system needs language in order to become self-conscious (Neuman & Nave, 2010), as well as other quandaries left unanswered by the cognitive-mentalist approach.

In sum, adopting a semiotic perspective and state of the art computational tools, we have shown that Peirce's idea is empirically grounded. Moreover, through case studies, we

have portrayed a process that explains the trajectory of abstraction. This explanation is in line with our evolutionary context, modern conceptions of "embodied cognitive science", and the variance that exists at the cultural level of analysis. In this sense, it is fully in line with Luria and Vygotsky's (1930) vision that the human being should be studied as the synergetic product of three trajectories: the evolutionary, the developmental-individual, and the cultural. Moreover, the methodology we introduced and used in this study - "computational cultural psychology" – is clearly in line with Valsiner vision of the appropriate methodology cultural psychology should adopt (Valsiner, 2007, p. 367). First, we use signs as our unit of analysis. This choice naturally derives from the semiotic perspective we adopt in this paper. Our signs are forms of mediation rather than simple causes in mechanistic causal models. Second, our study is not a simple technical procedure for supporting a theory. As argue by Valsiner (ibid): "An empirical proof of a hypothesis is productive only if it leads to a new idea, rather than if it confirms an existing one". Our analysis opens new ways for inquiring cultural mediation as evident from our case studies. It is for future studies to identify the exact patterns through which an embodied experience is translated into a reified object. Moreover, our methodology actually consists of studying "multilevel causal systems" and "downward causation" because signs as forms of cultural mediations that emerge from collective and historically grounded interactions, clearly influence the mind of individuals who contribute to this collective activity. In sum, there are many points of intersection between the methodological vision presented by Valsiner and our study.

It goes without saying that the current study does not fully resolve the "Language" and "Thought" issue, but provides a different approach to address it and several interesting findings that point at the way our sign system is being transformed. The approach used in this paper, "Computational Cultural Psychology", may have great potential for addressing questions in soci-cultural psychology. For instance, by using the resources in our paper and

the Books Ngram Viewer recently developed by Google [<http://ngrams.googlelabs.com/datasets>] (Michel et al., 2010), we can trace the abstraction of words through history and whether there are identifiable patterns of abstraction. Moreover, we may trace the way words have acquired and changed their meaning through history and contexts. Therefore, the paper concludes by inviting researchers in the field of psychology to examine new tools for addressing old questions.

References

- Barsalou, L. W. (2005). Continuity of the conceptual system across species. *Trends in Cognitive Sciences*, 9, 309-311.
- Büttcher, S., & Clarke, C. (2005). Efficiency vs. effectiveness in terabyte-scale information retrieval. *Proceedings of the 14th Text REtrieval Conference (TREC 2005)*, Gaithersburg, MD.
- Caron, J. (2001). Experiments with LSA scoring: Optimal rank and basis. *Proceedings of the SIAM Computational Information Retrieval Workshop* (pp. 157-169). Raleigh, NC.
- Church, K. W., & Hanks, P. (1989). Word association norms, mutual information and lexicography. *Proceedings of the 27th Annual Conference of the Association of Computational Linguistics* (pp. 76-83). Association for Computational Linguistics, New Brunswick, NJ.
- Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences*, 10, 370-374.
- Collected Papers of Charles Sanders Peirce [CP]* (Vols. 1-8). Cambridge, MA: Harvard University Press (Vols. 1-6, C. Hartshorne & P. Weiss, Eds.; Vols. 7-8, A. W. Burks, Ed.).
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A(4), 497-505.
- Danesi, M. (2003). Metaphorical “networks” and verbal communication: A semiotic perspective on human discourse. *Sign Systems Studies*, 31, 341–363.
- Davies, M. (2008). The Corpus of Contemporary American English (COCA): 410+ million words, 1990-present. Retrieved from <http://www.americanacorporus.org>.
- Deutscher, G. (2010). Does your language shape how you think? *The New York Times*, 29 August 2010.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The latent semantic

- analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Lawrence, J. A., & Valsiner, J. (1993). Conceptual roots of internalization: From transmission to transformation. *Human Development*, 36, 150-167.
- Lee, P. (1996). *The Whorf theory complex*. Amsterdam: John Benjamins Publishing Company.
- Luria, R. A., & Vygotsky, L. S. (1930). *Ape, primitive man and child*. London: Harvester Wheatsheaf.
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Aiden, E. L. (2010). Quantitative analysis of culture using millions of digitized books. *Science* [10.1126/science.1199644].
- Neuman, Y. (2003). *Processes and boundaries of the mind: Extending the limit line*. New York: Springer.
- Neuman, Y. (2008). *Reviving the living: Meaning making in living systems*. Oxford: Elsevier.
- Neuman, Y. (2009). On love, hate and knowledge. *International Journal of Psychoanalysis*, 90, 697-712.
- Neuman, Y., & Nave, O. (2009). Metaphor-based meaning excavation. *Information Sciences*, 179, 2719-2728.
- Neuman, Y., & Nave, O. (2010). Why the brain needs language in order to be self-conscious. *New Ideas in Psychology*, 28, 37-48.
- Peirce, C. S. (year). *The Essential Peirce. Selected Philosophical Writings* (Vols. 1-2). Bloomington, IN: Indiana University Press. (Vol. 1[1867-1893], N. Houser & C. Kloesel, Eds., 1992; Vol. 2 [1893-1913], Peirce Edition Project, Eds., 1998).
- Pennebaker, J. W., & Francis, M. E. (2001). *Linguistic inquiry and word count*. USA: LEA.
- Pezzulo, G., & Castelfranchi, C. (2007). The symbol detachment problem. *Cognitive*

- Processes*, 8, 115-131.
- Rapp, R. (2003). Word sense discovery based on sense descriptor dissimilarity. *Proceedings of the Ninth Machine Translation Summit* (pp. 315-322). New Orleans.
- Salvatore, S., & Venuleo, C. (2008). Understanding the role of emotion in sense making: A semiotic psychoanalytic oriented perspective. *Integrative Psychological and Behavioral Science*, 42(1), 32-46.
- Saussure, F. de. (1973). *Course in General Linguistics* (R. Harris, Trans.). London: Duckworth.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis method. *Journal of Language Social Psychology*, 29, 24-54.
- Tobin, Y. (1990). *Semiotics and linguistics*. London: Longman Publishing Group.
- Turney, P. D., & Littman, M. L. (2003), Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4), 315-346.
- Turney, P. D., & Pantel, P. (2010), From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research (JAIR)*, 37, 141-188.
- Valsiner, J. (2007). *Culture in minds and societies*. London: Sage.
- Vauclair, J. (2003). Would humans without language be apes? In J. Valsiner & A. Toomela (Eds.), *Cultural guidance in the development of the human mind*. Vol 7. Advances in child development within culturally structured environment (pp. 9-26). Greenwich, CT: Ablex Publishing Corporation.
- Volosinov, V. (1986). *Marxism and the philosophy of language*. Cambridge, MA: Harvard UP.

Vygotsky, L. S. (1962). *Thought and language*. Harvard, MA: The MIT Press. Electronic source: Thinking and speaking <http://www.marxists.org/archive/vygotsky/works/words/index.htm>

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments.