You Can't Wear a Coat Rack: A Binding Framework to Avoid Illusory Feature Migrations in Perceptually Grounded Semantic Models

Michael N. Jones (jonesmn@indiana.edu)
Department of Psychological and Brain Sciences
Indiana University, Bloomington, Indiana USA

Gabriel Recchia (grecchia@indiana.edu)
Cognitive Science Program, 1910 E 10th St.
Indiana University, Bloomington, Indiana USA

Abstract

Recent Semantic Space Models (SSMs) are now integrating perceptual information with linguistic statistics into a unified mental space, offering a solution to the criticism that SSMs are disembodied. However, these new models introduce the problem of illusory feature migrations. When the word *dog* is perceived, its perceptual features should migrate to *hyena*, so the system can infer the perceptual features for a non-perceived word (hyenas have fur). In doing so, however, the models are unable to avoid migrating the features for *dog* to syntagmatically related words, such as *bone.* As a result, the models incorrectly infer that bones have fur. We argue that the problems of perceptual grounding and word order are not independent—a model of word order information is needed to correctly infer how features should migrate in mental space. We introduce a multiplicative binding framework that allows all information sources to be stored in a composite mental space, but features will only migrate to words that share sufficient order information with directly perceived words.

Keywords: semantic space models, symbol grounding problem, perceptual integration, embodied cognition.

Introduction

Semantic Space Models (SSMs) have seen remarkable success in recent years as models of how humans learn the meanings of words from repeated episodic experience, and for how lexical semantics are represented in mental space. Many types of SSMs now exist, with several modifications to better approximate human semantic cognition.[1] In general, these models all create semantic representations from statistical regularities in large linguistic corpora, building on Harris' (1970) *distributional hypothesis* of lexical semantics: the more similar the contexts in which words are experienced, the more similar their meanings. SSMs have successfully accounted for a wide variety of human semantic data, ranging from semantic priming and free association, up to high-level discourse processing by applying compositional algorithms to SSM representations.

Despite their successes, SSMs have been heavily criticized as implausible psychological models on a number of grounds. Firstly, most of these models have been criticized as "bag-of-words" models, in that they simply consider the context in which the word occurs, but ignore the statistical information inherent in word transitions. Recent solutions to the word-order problem use binding operations to learn a blended semantic space in which a word's representations reflects its history of co-occurrence with, and position relative to, other words (e.g., Jones & Mewhort, 2007). Further, these models are able to retrieve plausible n-gram information (coarse grammaticality) directly from the blended space, without the need for explicit rules of grammaticality. The integration of word order information has been shown to give a much better fit to human data in a variety of semantic tasks.

Secondly, SSMs have been criticized as "disembodied" in that they learn from only linguistic information but are not grounded in perception and action (see de Vega, Graesser, & Glenberg, 2008 for a workshop on the issue). The lack of grounding in SSMs is in direct contrast to the recent literature on embodied cognition, demonstrating that a word's meaning is grounded in sensorimotor experience. Sensorimotor information is an inherent part of the semantic organization of the human lexicon, but much of this information cannot be learned from statistics in a text corpus—it must be learned from multisensory experience (but see Riordan & Jones, in press). In addition, current models have a symbol-reference problem: there is no way to link a word's internal representation back to its referent in the real world.

We are now seeing the emergence of the first perceptually grounded SSMs. As a proxy for sensorimotor experience, these models use norms of human-generated features (such as the norms of McRae et al., 2005). These norms represent aggregate human productions of the physical properties, appearance, sounds, smells, functional properties, etc. for concrete nouns and event verbs based on multisensory experience. For example, the feature <has_4_legs> will have a high probability for *dog* and *cow*, but a low probability for *centipede*, and a zero probability for *strawberry*. However <is_red> is a highly salient feature of *strawberry* and not for *dog*.

Most of the new grounded SSMs simultaneously consider the distribution of words across contexts in a text corpus and the distribution of words across perceptual features, allowing them to extract joint information between the two data sources. This allows the models to make implicit inferences across the two information sources: if the model learns from perceptual experience that *sparrows* have beaks, and from linguistic experience that *sparrows* and *mockingbirds* are used in a similar distributional fashion, it naturally makes the inference that *mockingbirds* have beaks. The inference chain works in the opposite direction as well. Most impressive, given a novel word most of these models can retrieve an accurate representation of the perceptual features of the novel word's referent. Simulations have

[1] For recent advances in SSMs, see the upcoming issue of *Topics in Cognitive Science* edited by Danielle McNamara.

demonstrated that the blended linguistic/perceptual mental space may yield a superior approximation of human data.

However, a major issue common to all of these new grounded models is that they have no way to discriminate between syntagmatic relationships (e.g., the relationship between *bee* and *honey*) and category-based paradigmatic relationships (e.g., *bee* and *wasp*). The linguistic abstraction phase of these models will learn to position the vectors for *car, automobile,* and *road* close in semantic space. This produces the problem that the model cannot distinguish which regions of space may adopt features that migrate from a perceptually grounded word during the feature inference phase. The result is that the model correctly infers that *automobile* <has_wheels>, but it also incorrectly infers that *road* <has_wheels>. We refer to these errors as *illusory feature migrations*, and argue that errors of migration are much more common in semantic space than are correct migrations, which can severely pollute the resulting semantic space relative to a human representation that would not contain this type of error.

One reason these models fail to discriminate between context-based syntagmatic vs. category-based paradigmatic relationships is that they ignore word order information, which is a powerful cue for category membership (Jones & Mewhort, 2007). That is, words that are flanked by similar n-grams tend to belong to the same conceptual categories. Extensive study in the field of category-based inference has investigated the ways in which category structure constrains feature generalization (for reviews, see Heit, 2000; Rips, 2001). To ignore word information is to ignore a very salient cue to category membership at an SSM's disposal.

To be clear at the outset, we strongly commend the authors of these perceptually grounded models for taking a huge step in the right direction towards our understanding of human semantic representation. However, a plausible model must also be able to filter components of this representation so that perceptual information may generalize to paradigmatically but not syntagmatically similar words (i.e., from *car* to *automobile* but not *road*). Here we explore the utility of a formal binding framework based on ideas from signal processing and Jones and Mewhort's (2007) BEAGLE model that has these desiderata.

Grounding Semantics in Perception and Action

Recent attempts to ground SSMs in perception and action can be placed into one of two classes: post-hoc inference models, and ad-hoc inference models. Both types can be trained on the same text corpus and feature representations (e.g., TASA and McRae et al., 2005).

Post-hoc inference models begin with the abstraction of a text corpus into a reduced vector space (a traditional SSM), and then attempt to bind these linguistic vector representations to the feature norms. For example, Durda, Buchanan, and Caron (2009) train a feedforward neural network to associate linguistic vectors with their corresponding activation of features. Given the linguistic representation for *dog*, the output feature <has_fur> should be activated but the output feature for <made_of_metal> should be inhibited. After iterative training with backprop, the model can infer the correct pattern of perceptual properties for words that did not have a perceptual feature vector. At its core, this technique simply maps similar linguistic vectors to similar output vectors, as with other pattern generalization applications of feedforward networks.

Ad-hoc inference models typically begin with a raw word-by-document matrix of a text corpus and a word-by-feature matrix of a feature database. During learning, the model attempts to learn a word's representation by simultaneously considering inference across documents and features. An excellent example of an ad-hoc model is presented in Andrews, Vigliocco, and Vinson (2009). Andrews et al., use a Bayesian framework to infer the joint distributional information for a word between linguistic and perceptual data. It is important to note that their technique is *joint inference*: it squeezes more information out of the data than simply adding perception to linguistic experience. Andrews et al. convincingly demonstrate that their joint model gives better fits to word association data than a model that considers only one data source, or the simple addition of the two sources.

Illusory Feature Migrations

A major problem with both post-hoc and ad-hoc inference models is that they must exhibit illusory feature migrations as a consequence of their architecture. An illusory feature migration occurs when a non-perceived word adopts erroneous features from a linguistically related word simply because they are proximal in semantic space. This is a common issue in the aforementioned models because they do not have order information to discern between syntagmatic and paradigmatic word relations. If the models are optimized on free-association data (which is strongly dominated by syntagmatic productions), then they must position syntagmatically related words like *bee* and *honey* close in space, as well as paradigmatically related words like *bee* and *wasp*. As a result, the inference mechanism simply sees both *honey* and *wasp* as similar patterns to *bee*, and naturally makes the inference that *honey* can fly and has wings.

Note that the "migration" described need not be a dichotomous on/off feature. It is simply the case that the inferred distribution over possible features for *honey* has some correlation with that of *bee* simply because their distributional structure in language has overlap. This overlap introduces error in the labeling of novel referents (e.g., a novel object that looks like an insect will activate words like *honey* as potential labels). Furthermore, this inference error will introduce noise to the overall semantic organization, which will lead to a poorer account of human semantic data compared to a human who will not make these inference errors. The aforementioned models demonstrated examples of correct feature generalizations in their papers; what was not illustrated is the larger number of incorrect feature generalizations.

Presumably, humans use word-order information to constrain the inference of features in mental space. This information allows a model to distinguish what types of words may adopt features given a perceived target word. Rather than making this a terse rule-based model, we choose to adopt a graded feature migration framework—words adopt the aggregate features of proximal words that have features, weighted by their similarity in order space. However, it is also important to keep the sources (context, order, perception) blended to account for the wide range of embodied semantic data. This requires a model that can create a blended semantic representation, but that can know what part of the semantic signal to use in computing similarity for feature migration. We next describe a simple framework towards this type of integrated model, test its behavior on an artificial language paradigm, and then scale it up to a real language corpus to see how the properties are maintained at a large scale.

A Feature-Binding Framework

Our goal was to build an SSM with two key properties. First, context, order, and feature information should be represented as patterns in high-dimensional vectors. Even though these three sources of information should be blended within a single vector, it should be possible to determine the degree of similarity between two words in context space, order space, or feature space alone. Because context, order, and feature information is distributed, computing a vector cosine between two vectors reflects their similarity when all three sources of information are taken into account.

Second, feature migration should occur, but features should only migrate to words with which they share order information (i.e., words that are commonly flanked by similar n-grams). For example, *food* and *table* will share primarily context information, whereas *table* and *countertop* will share primarily order information; therefore, features should migrate from *table* to *countertop*, but not from *table* to *food*.

Encoding. Our model is similar to other SSMs that represent both context and order information with fixed-length high-dimensional vectors (Jones & Mewhort, 2007; Sahlgren et al., 2008). When a word w is encountered in the input text for the first time, it is assigned an initial "environmental" vector e_w—a random vector whose elements are randomly selected from a Gaussian distribution of mean 0 and variance 1. Environmental vectors are intended to represent the static properties of a word's surface form, such as its orthography and phonology, and are not updated during processing. The new word is also assigned an initially empty memory vector m_w to represent its semantics. When the model encounters a new sentence in the input corpus, m_w is modified according to the update rule:

$$m_w = m_w + (C_I \odot context) + (O_I \odot order) + (F_I \odot features_w)$$

where the circumpunct "$\odot$" denotes elementwise vector multiplication, one of a class of multiplication-like operators that vector symbolic architectures employ to combine vectors in a neurally plausible manner (Levy & Gayler, 2009; Kanerva, 2009). C_I, O_I, and F_I are *indicator vectors*—unchanging vectors that are bound with vectors representing context, order, and feature information, respectively. They serve to "tag" the source of the information signal (context, order, or perception). They may be initialized either as random vectors, or as binary vectors of ones and zeros sharing little or no overlap with each other.

As in Jones & Mewhort (2007) and Sahlgren et al. (2008), the *context* vector represents co-occurrence information: it is the sum of all environmental vectors of words occurring in the same sentence as w. The *order* vector is the sum of all n-grams surrounding w up to some fixed window size, where an n-gram is represented by binding the environmental vectors of all the words comprising the n-gram via elementwise multiplication. In the experiments presented here, only bigrams directly to the left and right of w are considered. As in Sahlgren et al. (2008), words to the right and left are distinguished by rotating the environmental vectors by one unit in a positive or negative direction, respectively. Finally, the *features* vector represents information about sensorimotor features of words. Each of 2,526 features taken from the feature norms of McRae, et al. (2005) was assigned a unique random vector. If w is the word for one of the 541 concepts for which feature norms were collected, $features_w$ is the sum of the five vectors that correspond to the five features that were attributed to w by the greatest number of participants. If w is not among the concepts in the McRae et al. feature norms, $features_w$ is initialized as a vector of zeroes (and only acquires nonzero values during training, when vectors are added to m_w via the update rule). The fact that $features_w$ has a w subscript while *context* and *order* do not reflects the fact that $features_w$ is derived from information about w in the feature norms, while *context* and *order* represent information about the sentence currently being processed.

Retrieval. After training, the cosine between every pair of memory vectors is calculated to determine the model's estimate of the semantic similarity between words. These similarity scores can be thought of as distances between points in a high-dimensional space, which we refer to as the *composite space*. In addition to having a lower computational complexity than circular convolution, one benefit of using elementwise vector multiplication for binding the information source tag is that the operation serves as its own approximate inverse when vector elements are sampled from a z-distribution, hence:

$$X \approx (X \odot Y) \odot Y \qquad (1)$$

This allows vectors to be elementwise multiplied with the aforementioned context indicator vector C_I before calculating their cosines. The operation serves to 'unbind' the C_I * *context* binding, yielding a *context space* in which two words' distance from each other reflects the amount of context information they share (but is not heavily influenced by shared order or feature information). Similarly, unbinding via elementwise multiplication with O_I yields an *order space* in which cosine similarity reflects the amount of shared order information; unbinding with F_I yields a *feature space* where feature information is paramount.

Experiment 1

The objective of Experiment 1 was to determine whether the binding model we outlined does in fact possess the desired property of representing context, order, and feature information in a separable fashion, and whether it behaves appropriately with respect to feature migration. Demonstrating this required training the model on a corpus in which the amount of context, order, and feature information that words share is known, which is best accomplished using a corpus of an artificial language. Strictly controlling the input allows us to determine conclusively whether the model at least exhibits the desired properties in the simplest case and lets us more clearly observe how the inclusion or exclusion of different types of information affects the similarity space.

Method

Input corpus. The model was trained on a corpus of 1,000 sentences from a simple artificial language. This language was designed such that it would contain some word pairs that shared context information but not order information, some pairs that shared order information but not context information, and some words that shared context as well as order information. The language used is described by the following context-free grammar (symbols in bold are terminal symbols):

$S \rightarrow$ **A** *Aux* **B** *Num* **Cs** | **D** *Aux* **E** *Num* **Fs**
Aux $\rightarrow$ **can** | **should** | **would** | **could** | **does**
Num $\rightarrow$ **two** | **three** | **four** | **five** | **six**

Sentences of the corpus were generated randomly, with each possible transition of equal probability. Thus, it consisted of sentences such as “A can B three Cs”, “A would B four Cs”, “D should E three Fs”, and so forth. In this corpus, *A, B,* and *Cs* each share context information, as they always co-occur, but they do not share order information. If this were a real language, one could think of *A, B,* and *Cs* as fillers for three different grammatical roles. Similarly, *D, E,* and *Fs* share context, but not order, information. In contrast, the members of pairs {A, D}, {B, E}, and {Cs, Fs} each share order information, but significantly less context information. The auxiliary verbs {can, should, would, could, does} and numbers {two, three, four, five, six} share significant amounts of order information with each other. They also share context information: even though the grammar allows auxiliaries and numbers to co-occur with any of *A, B, Cs, D, E,* or *Fs*, each auxiliary always co-occurs with some number.

Procedure. Two simulations were conducted. In Simulation 1, no feature information was included. In Simulation 2, we retrained the model with the full update rule $m_w = m_w + (C_I \odot context) + (O_I \odot order) + (F_I \odot features_w)$, adding five vectors corresponding to five features for the word “strawberry” from the McRae et al. norms to the concept for the word *A* (*a_fruit, grows_on_plants, grows_in_fields, grows_on_bushes,* and *has_green_leaves*). We compared the model under three conditions: context, composite, and order. In each condition, feature migration proceeded by unbinding $m_w \odot F_I$ to retrieve an approximation $features'_w$ of $features_w$, and adding this approximation to every other memory vector m_i in proportion to the strength of their similarity in the relevant space (context space, composite space, or order space, depending on condition). That is, features tend to be more likely to migrate in the order condition between two words that share a large amount of order information than between two words that do not. Because we are interested in migrating features not merely to words that are “close” to the perceived word but rather to words that are similar to *w* in terms of their relationships to other words, the similarity between words w_1 and w_2 is obtained by correlating a vector of w_1’s cosine with each word in the lexicon with a vector of w_2*’s* cosine with each word in the lexicon. However, using just the cosine of w_1 and w_2 yields largely similar results.

Simulation 1.1. Tables 1 and 2 illustrate the most similar words to *A, B, Cs, D, E, Fs, can,* and *two* in context and order space, respectively, after training using the update rule $m_w = m_w + (C_I \odot context) + (O_I \odot order)$; no feature information was included in this simulation. In the absence of feature information, context and order information are separable in this model, despite the fact that both information sources are fully distributed across vector elements. Appropriately, the members of {A, B, Cs} cluster together in context space, as do the members of {D, E, Fs}. Additionally, pairs {A, D}, {B, E}, and {Cs, Fs} cluster together in order space. Although they do not appear in the tables, auxiliaries and numbers also cluster together.

Table 1. Z-scores of cosines of the most similar words to *A, B, Cs,* and *D* in context space, Simulation 1.

A		B		Cs		D	
A	3.6	B	3.6	Cs	3.6	D	3.6
B	.20	Cs	.20	B	.21	E	.18
Cs	.16	A	.16	A	.13	Fs	.15
five	-.08	two	.01	two	-.07	three	-.06
two	-.09	five	-.01	five	-.09	could	-.09

Table 2. Z-scores of cosines of the most similar words to *A, B, Cs,* and *D* in order space, Simulation 1.

A		B		Cs		D	
A	3.5	B	3.7	Cs	3.5	D	3.5
D	1.2	E	.32	Fs	1.1	A	1.2
B	-.10	A	-.03	B	-.15	Fs	-.14
Cs	-.13	Cs	-.04	A	-.17	E	-.17
can	-.24	can	-.22	can	-.24	can	-.30

Simulation 1.2. Table 3 illustrates the standardized correlations of vector cosines of the four most similar words to A under each migration condition. Because the migration rule transfers feature information in direct proportion to these values, the higher the value of a word, the more feature information that word receives from A. The important pattern in Table 3 is the reversal of B and D: in context space, the syntagmatic relation between A and B is much more salient, but in the order space the paradigmatic relation between A and D is emphasized. In the overall composite space, these relations are mixed (our desired blending in full lexical space), but the information required for correct feature migration is still implicitly represented.

Table 3. Standardized correlations of vector cosines of the four most similar words to *A* under the context, composite and order conditions, Simulation 2.

context		composite		order	
A	3.5	A	3.2	A	3.4
B	.63	B	.06	D	1.2
Cs	.55	D	.04	B	.05
does	-.17	Cs	.00	Cs	-.03

Thus, it appears that only the *order* condition minimizes opportunity for illusory feature migrations while preserving the appropriate migration to D, which is paradigmatically similar to A in this corpus. Furthermore, when feature information is added, the separability between context and order space is maintained, (allowing features to appropriately migrate from A to D) and individual features can be successfully retrieved.

Experiment 2

The objective of Experiment 2 was to explore whether the proposed binding framework continues to yield distributions that inhibit illusory feature migrations (i.e., migrations to syntagmatically similar words) while facilitating appropriate feature migrations to paradigmatically similar words when scaled up to a corpus of natural language. We therefore designed a version of Experiment 1 trained on a real corpus, the TASA corpus of high-school level English text. Two simulations were conducted: The first to examine the similarity of the decoded context and order spaces to paradigmatic and syntagmatic relations, and the second to demonstrate feature migrations to category co-ordinates vs. non-categorical associates of a target word. Both simulations were identical to Experiment 1's Simulation 2 in terms of the update rule, the conditions (context, order and composite), and the feature migration rule.

Simulation 2.1. For each word, its feature vector $features_w$ was generated by summing the five vectors corresponding to the five features from the McRae et al. norms attributed to *w* by the greatest number of participants. As test items, we extracted 1075 word pairs from the word association norms of Nelson, McEvoy, & Schreiber (1998) for which both the first word of the pair (the cue) and the second word of the pair (the target) were members of the McRae et al. feature norms[2]. For each pair, we determined the category membership of each word, using the categories employed by Cree & McRae (2003, Appendix B): weapons, vehicles, foods, and so forth. Cree & McRae explicitly list which normed words belong in which categories, allowing us to code whether the cue was a member of the same conceptual category as the target. The 690 pairs in which both words shared a category were interpreted as being paradigmatically related (e.g., *apple-pear*), while the 385 paired words not sharing a category were interpreted as being syntagmatically related (e.g., *apple-crab*). The fact that two words are associates and do not appear in the same category does not guarantee syntagmatic similarity nor does it preclude phrasal association, however, informal observation suggests that many word pairs in the latter condition tend to appear in collocations or other classic syntagmatic relationships for which feature migration would be inappropriate. Indeed, the cosine similarity scores from the McRae et al. (2005) feature vectors for the word pairs were significantly higher for our paradigmatically related words than for syntagmatically related ones, $t(1073) = 24.66, p < .001$.

Motivated by the results of Experiment 1, we predicted that words sharing paradigmatic relationships would be closer in order space than in context space. This pattern of results would suggest that attending to order information facilitates more feature migrations among paradigmatically related words than among syntagmatically related ones, while attending to context information does just the opposite. For paradigmatically related words, the model's cosine similarities were significantly higher in order space than in context space, $t(689) = 2.96, p < .01$. That is, words in paradigmatically related pairs were gauged to be more similar to each other in order space than in context space. In contrast, for syntagmatically related pairs, the model's cosine similarities were significantly higher in context space than in order space, $t(384) = 4.371, p < .001$.

Simulation 2.2. To briefly demonstrate how illusory feature migrations may be corrected by incorporating order information, we selected 25 "triples" from Simulation 1, each consisting of a target **T** that existed in the McRae et al. norms, a category coordinate **CC** of T, and a syntagmatically related word **R** that had an associative relationship with T but was not a member of the same category. An example triple is <T:freezer, CC: refrigerator, R:ice>. *Freezer* and *refrigerator* each share a common class (kitchen appliances); *freezer* and *ice* are certainly related as well, but not by virtue of a category relationship. Intuitively, one would like features to migrate more strongly from T to CC than from T to R, given that categories for concrete words are defined at least partly on the basis of feature overlap. For example, the most popular features of *freezer* are *used_for_storage,* and *has_an_inside*, features that are

[2] We excluded the 24 concept words that the McRae et al. norms explicitly identify as having ambiguous meanings, such as "mouse_(animal)" and "mouse_(computer)."

much more applicable to kitchen appliances than they are to related non-category members (ice, frozen waffles, etc.). If a particular feature migrated more strongly from T to R than from T to CC, this was coded as an illusory feature migration. Otherwise, it was coded as an appropriate feature migration.

The (incorrect) migration of feature information from T to R was much stronger in the context condition than the order condition, and the (correct) migration of feature information from T to CC was stronger in the order condition than the context condition. By our coding scheme, 56% of the triples exhibited at least one illusory feature migration in the context condition (recall that this means the migration was stronger from T to R than it was from T to CC). In contrast, only 40% of the triples exhibited at least one illusory feature migration in the order condition. Most notable is that *all* illusory feature migrations that took place in the order condition also took place in the composite condition, and *all* illusory feature migrations taking place in the composite condition also took place in the context condition. In other words, some illusory feature migrations that took place in the context and composite conditions were avoided in the order space. Hence, emphasizing order information by unbinding with O_I (order space) yielded equal or better results for every triple when compared with emphasizing context information by unbinding with C_I (context space) or not unbinding at all (composite space). Table 4 presents four triples that differed by condition as to whether CC or R was deemed a better candidate for feature migration from T by the model. In each case, a feature migration error was committed in the context condition, but was avoided in the order condition.

Table 4. Example feature migration errors in context space that were corrected in the order space. Cases in which the related word was the stronger attractor were considered illusory feature migrations. Target word is bold.

Triple	Features most strongly attributed to target by participants in McRae et al. (2005)	Competitor that *features*$_w$' Migrated More Strongly To, By Condition		
		context	comp	order
bottle CC: jar R: fill	used_for_holding_things made_of_glass used_for_holding_liquids made_of_plastic has_a_lid	fill	jar	jar
cat CC: mouse R: tom	has_fur an_animal a_pet eats has_whiskers	tom	tom	mouse
horse CC: cow R: saddle	used_by_riding is_large an_animal has_a_mane has_legs	saddle	cow	cow
motorcycle CC: car R: wheels	has_wheels has_2_wheels is_dangerous has_an_engine is_fast	wheels	car	car

General Discussion

Integration of sensorimotor information is an important next step in the development of SSMs. While human-generated feature norms are admittedly an intermediary step, it is important to understand the cognitive mechanisms that humans might use to integrate perception/action and linguistic structure to organize meaning in memory for when perceptual models (e.g., computer vision) are sophisticated enough to directly represent environmental information to integrate with linguistic distributional structure (see Roy, 2008 for a discussion).

While early attempts at integrating perception and language in SSMs have shown much promise, our work here indicates that a model must have a mechanism to encode temporal linguistic information to know how perceptual information may be generalized in the mental space. The binding framework presented here shows the basic property of storing all information sources in a blended composite space (as is suggested by the literature in embodied cognition). However, the model is able to identify which components of the composite signal perceptual information should be allowed to migrate to. While this scheme needs more testing at a large scale, we believe it has promise for accounting for a wide range of semantic and embodied data, and is a step toward addressing criticisms of SSMs being ungrounded.

Acknowledgements

This research was supported in part by grants from Google Inc. and IBM to MNJ.

References

Andrews, M., Vigliocco, G., & Vinson, D. P. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116*(3), 463-498.

Cree, G. S. & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132*(2), 163-201.

de Vega, M., Graesser, A., & Glenberg, A. (2008). *Symbols and Embodiment: Debates on Meaning and Cognition*. New York: Oxford.

Durda, K., Buchanan, L., & Caron, R. (2009). Grounding co-occurrence: Identifying features in a lexical co-occurrence model of semantic memory. *Behavior Research Methods, 41, 1210-1223.*

Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review, 7,* 569-592.

Jones, M. N. & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114,* 1-37.

Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representations with high-dimensional random vectors. *Cognitive Computation, 1,* 139-159.

Levy, S. D., & Gayler, R. W. (2009). A distributed basis for analogical mapping. In B. Kokinov, K. Holyoak, & D. Gentner (Eds.), *New frontiers in analogy research.* New Bulgarian University Press.

McRae, K., Cree, G., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37,* 547-559.

Riordan, B., & Jones, M. N. (in press). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science.*

Roy, D. (2008). A mechanistic model of three facets of meaning. In de Vega, Glenberg, and Graesser (Eds.) *Symbols and Embodiment.*

Sahlgren, M., Holst, A., & Kanerva, P. (2008). Permutations as a means to encode order in word space. *Proceedings of Cognitive Science Society.*