Conceptual vs. Referential Affordance in Concept Composition[*]

Louise McNally
Universitat Pompeu Fabra

Gemma Boleda
Universitat Pompeu Fabra

January 5, 2016

1   Introduction

The goal of this paper is to confront and explore the larger implications of a problem that we have repeatedly observed in our ongoing work on the semantics of modification within noun phrases, which is one instantiation of concept combination. The problem is that, in the absence of context, sometimes the default interpretation for the modifier-noun combination is so strong as to make other possible interpretations seem impossible, whereas in context any interpretation – even the seemingly impossible – is possible. Here is just one example, involving so-called ethnic adjectives, which provide information about the ethnic origin, nationality or other locational origin of individuals.[1] Kayne (1984) and many others have claimed that when ethnic adjectives like *Canadian* combine with eventuality-denoting nouns, the adjective must contribute information about the most external argument of that eventuality, typically the agent. When it does not, a prepositional phrase expressing the corresponding participant role must be used. Thus, in (1), where the context does not previously mention Yeltsin visiting Canada, the PP *to Canada* rather than the adjective *Canadian* is what the author chose, and indeed the adjective sounds very odd.

(1)   Yeltsin met the prospective Democratic presidential candidate Bill Clinton on June 18. His itinerary also included **an official visit to Canada**/??**an official Canadian visit**. (*BNC*)

However, one does not have to go far to find counterexamples to Kayne's claim. When context or background knowledge make salient that some role other than agent is assigned to the location/ethnicity, the adjective is perfectly felicitous and attested, as in (2).

[1]The sources of examples taken from corpora or the internet via Google searches are indicated in parentheses after the example. 'BNC' refers to a local installation of the British National Corpus, though we have also consulted the corpora at http://corpus.byu.edu for supplementary information.

(2)   Prince Edward and wife begin **Canadian visit**
        (http://metronews.ca/news/canada/365325/prince-edward-and-wife-begin-canadian-visit/)

Confronted with the contrast between the strong default interpretation and the possibility of any interpretation in context, linguists have tended to follow one of two routes, both of which we will discuss and exemplify below. The first involves taking the default interpretation as the crucial fact to account for, leaving the non-default interpretations in context unexplained. The second involves providing an analysis that is weak enough to capture all interpretations, and saying little or nothing about the strength of the default interpretation. In this paper, we argue that, in effect, both routes must be taken because two fundamentally different interpretative processes can be appealed to in the composition of modified noun phrases or, to echo the title of this volume, in concept composition. Specifically, we take default interpretations to be the result of what we will call *conceptually afforded concept composition*, and non-default interpretations to be the result of *referentially afforded concept composition*. We borrow the term *affordance* loosely from the psychology literature, specifically the interpretation of the term in Chemero (2003), as we discuss in further detail in section 3.1.

    This distinction builds on the long-standing observation that language mediates between concepts in our mind and the things they refer to in the world (Ogden and Richards, 1923, among many others). We take these connections to concepts and to the world to be distinct aspects of language, each of which facilitates a different process of concept composition. Take for instance the phrase *red box* in the examples in (3). In the absence of any context, *red*, when modifying *box* (or indeed any noun denoting a physical object), refers to its color, and so we can usually paraphrase (3-a) as "Identify a box that is red in color and put the relevant scarf inside it". However, it may also refer to other properties of the box referent, such as the intended color of its contents, if the discourse context makes the relevant property clear (3-b).

(3)   a.   Put the scarf in the red box.
        b.   *(Context: For a fundraising sale, Adam and Barbara are sorting donated scarves according to color in different, identical, brown cardboard boxes. Barbara distractedly puts a red scarf in the box containing blue scarves.)*
              Adam: Hey, this one belongs in the red box!

We call cases like (3-a) *conceptually afforded*. In these cases, some component(s) of the concepts contributed by two expressions in a phrase match in a way that indicates how they should be composed, and interlocutors avail themselves of such a suggestion. This matching invites the hearer to identify red as the color of the box in (3-a). In contrast, in *referentially afforded* cases like (3-b), specific, independently available information about the referent described by the phrase is used to guide the way in which the concepts in question are composed.

    This paper has three goals. First we develop this distinction, which has a precedent in Asher (2011), in an explicit manner and support it with empirical evidence we gathered in previous work. Second, we suggest modeling conceptually-afforded concept composition via (compositional) *distributional semantics*, which represents meaning as a function of the contexts in which words and phrases appear in naturally occurring language data, usually a large text corpus (Landauer and Dumais, 1997, Turney and Pantel, 2010). A fundamental hypothesis of some work in distributional semantics (e.g. Lenci, 2008) is that the resulting semantic representations can be used to model the

concepts associated with words. For this reason, we will present a brief introduction to distributional methods in Section 4. Finally, we propose a way to formally distinguish the two kinds of concept composition and integrate them into a more general framework for semantic analysis.

2   Two approaches to analyzing modification

We start by discussing previous approaches to the problem outlined in the introduction. As mentioned, the existence of strong, but overridable, defaults in the interpretation of modifiers has led to two lines of analysis. The first involves the proposal of an inventory of primitive semantic relations to capture the defaults; the latter, the use of an underspecified modification relation which gets resolved in context or via an appeal to indexicality. We consider these in turn.

   The use of primitive semantic relations to mediate in modification has a long history. We cite two representative and well-known examples here. The first involves Levi's (1978) analysis of relational adjectives such as *microscopic* or *tropical* (ethnic adjectives like *Canadian* are also considered a subclass of the relational adjectives). Relational adjectives (as their name indicates) are morphologically adjectives, but they are also noun-like in several respects: They are synchronically or diachronically derived from nouns; they are generally defined as introducing a relation between an individual of the sort described by the adjective's nominal stem and that described by the modified noun (Bally, 1944); and they have a more restricted syntactic distribution than other types of adjectives, occupying in English essentially the same position in nominal syntax as do noun modifiers of nouns, very close to the head noun (e.g. *computer* in *computer store*). Some examples from Levi (1978, 27–28) are provided in (4), with typical paraphrases:

(4)   a.   microscopic analysis – analysis carried out using a microscope
      b.   tropical butterflies – butterflies found in the tropics
      c.   planetary mass – mass of a planet
      d.   editorial comment – comment by an editor
      e.   dramatic criticism – criticism of drama

Levi proposed that such examples are derived from an underlying structure that makes the relation in question explicit. She further proposed that an inventory of primitive relations could be specified: CAUSE, HAVE, MAKE, USE, BE, IN, FOR, FROM, and ABOUT. For the derivation of examples involving deverbal nominalizations, as in (4-d,e), she proposed somewhat more complex derivations that nonetheless also availed themselves of primitives, including in some cases AGENT and PATIENT.

   A second example of an appeal to primitive relations emerged in part from the strong tendencies in the interpretation of (non-relational) adjectives described in Pustejovsky (1995).

(5)   a.   red pen – pen that writes in red or that is red on the surface
      b.   red apple – apple whose skin is red
      c.   quick meal – meal that is quick to eat or to prepare

To account for these interpretations Pustejovsky argues that the lexical entry for content words (including nouns) should include what he called a Qualia Structure with four features, each corresponding to a *quale*: FORMAL, CONSTITUTIVE, AGENTIVE, and TELIC. The FORMAL quale

characterizes the general ontological properties of an object; the CONSTITUTIVE, its parts; the AGENTIVE, how it comes into existence; and the TELIC, its function. Pustejovsky proposes that adjectives can restrict the denotation of a noun by placing conditions on the values of the different qualia in the noun's semantic representation.

The logical representations in (6) illustrate how this approach can be used in the sorts of modification that interest us here. In (6-a), the primitive AGENT specifies the relation between Canada and the visit; in (6-b), the primitive CONSTITUTIVE acts as an operator on the representation of *apple* to retrieve indirectly a part of the apple to which the property denoted by *red* can be ascribed.

(6)   a.   *Canadian visit*: $\lambda e[\textbf{visit}(e) \wedge \text{AGENT}(e, \textbf{Canada})]$
       b.   *red apple*: $\lambda x \exists y[\textbf{apple}(x) \wedge \text{CONSTITUTIVE}(\textbf{apple}) = \text{PART-OF}(y, x) \wedge \textbf{red}(y)]$

The use of semantic primitives to capture modification relations has two main advantages. First, it speaks to the very strong intuitions that the literature reports about default or productive interpretation processes (see e.g. Levi, 1978, 84-86). Second, similar defaults are observed cross-linguistically – for example, Pustejovsky's theory has been applied to various languages, and Levi observes that she found evidence for a similar set of primitives in a study of Modern Hebrew (Levi, 1978, 86). Clearly, there is something to be captured in these data.

However, the use of primitives of any sort, at least as the only compositional strategy, has also long been argued to be problematic. On the one hand, it is clearly too strong insofar as no set of necessary and sufficient primitives can be provided to account for all cases.[2] Levi herself observes (p. 84; see also p. 238ff.) that the goal of her study is to account for patterns of modification that are *productive*, as opposed to *possible*: In other words, her aim was a theory of why, even if we *can* interpret, for example, a phrase such as *Korean passengers* as 'passengers on Korean Airlines',[3] our first inclination is arguably not to do so but rather to interpret it as 'passengers from Korea'. On the other hand, the use of primitives is too weak. As e.g. Clark (1992) and Murphy (2002) observe, even when such primitives might apply, they are insufficiently granular: There are cases in which they provide too little information about the exact nature of the relation instantiated by any given primitive. This is already apparent in the analysis of *red apple* in (6). The CONSTITUTIVE quale introduces a part of the apple, but it does not specify which part, and so the inference that it is the skin of the apple (or, more generally, its surface) is not directly accounted for. A representative example involving a relational adjective is *an electrical fire*, which could be paraphrased as 'a fire caused by electricity': This case is even more problematic than the *apple* example insofar as the paraphrase does not capture the fact that the term is used to refer to fires caused by malfunctions in electrical systems and not, for example, by lightning.

Given these problems, a second line of approach to modification has involved sacrificing the coarse generalizations embodied in primitives in favor of broader empirical coverage. On one version of the approach, modification is mediated by a maximally underspecified relation whose value, much like that of a pronoun, is resolved in context (examples include McNally and Boleda, 2004, Kennedy and McNally, 2010). On another (see e.g. Bosch, 1983, Rothschild and Segal, 2009) adjectives have as their lexical content functions from contexts to contents, that is, Kaplanian

---

[2]Observations to this effect with respect to productive compounding, which shares important properties with modification of the sort discussed by Levi, can be found as far back as Jespersen (1942). See Gagné et al. (this volume), Westerlund and Pylkkänen (this volume), and references cited in these works for further discussion of the complexities involved in modification.

[3]Example taken from http://www.koreatimes.co.kr/www/news/biz/2014/09/373_135962.html.

characters (Kaplan, 1989). These analyses are respectively illustrated in (7).

(7)   a.   *Canadian visit*: $\lambda e[\mathbf{visit}(e) \wedge R_i(e, \mathbf{Canada})]$
      b.   *red apple*: $\lambda x[(\mathbf{red}_i(\mathbf{apple}))(x)]$

Again, this approach has both advantages and disadvantages. On the positive side, it is appropriately flexible: There is no interpretation that cannot be accommodated under such an analysis. However, its flexibility is arguably also a disadvantage: It has nothing to say about the strength of default interpretations or the fact that we tend to generalize them to new examples (a point made, as noted above, by Levi). Moreover, these analyses have provided no substantive theory of how context intervenes to yield the interpretations that arise.

We know of only one explicit proposal that contemplates the possibility of combining these two general approaches to resolving modification, namely that in Asher (2011). Asher combines a classical, model-theoretically interpreted intensional logic with a separate, proof-theoretic logic of types that is intended to mirror language users' systems of concepts. The latter is used to compute and resolve the basic relations between predicates in composition – for example, it will allow us to determine that, in principle, it must be possible to infer that *red* picks out a type that, when combined with the type picked out by *pen*, yields a type that corresponds to a pen that writes in red.[4] Though he does not develop the possibility in detail, he (p. 226) suggests that Pustejovsky's qualia could be introduced into his system as type-shifting operators that mediate in this process: For example, WRITE could be the output of a general type coercion operator TELIC applied to the type *pen*, and this information could then be exploited in the semantic composition process. In addition, alongside the possibility of such operators, Asher's system contemplates the possibility of contextually-valued type coercion operators for cases where the discourse structure makes it clear that default value operators such as TELIC would not apply.

The proposal we develop in the rest of this paper shares with Asher's the intuition that there are (at least) two distinct sorts of composition processes involved in computing the interpretation of a sentence. Our contributions will consist in laying out the proposal in more explicit terms, providing new empirical support for this dual system, the use of distributional semantics an alternative to Asher's logic of types, and a specific proposal for formalizing the distinction using Discourse Representation Theory (Kamp, 1981).

3   A dual system for semantic composition

3.1   Conceptually vs. referentially afforded composition

We begin with a very programmatic proposal concerning two ways in which the construction of meaning can be mediated. Our proposal is based on the following assumption.

Assumption: The construction of meaning draws on connections we make between linguistic expressions and our conceptual structure, on the one hand, and the world, on the other.

---

[4]Asher's logic of types builds on an approach to type theory, now sometimes referred to as "modern" type theory, developed within intuitionistic logic by Martin-Löf (1984). Due to space limitations, we refer the reader to Asher's book for details concerning the logic of types.

This assumption is of course familiar from traditional semiotic models and also resonates with the "dual content" model recently proposed in Del Pinal (2015), which provides (p. 44ff.) a useful overview of the different ways in which language, conceptual structure, and the world have been related to each other both in the philosophy of language and cognitive psychology literature. The assumption also underlies the classic Fregean model that distinguishes *sense* (*Sinn*), which Frege suggests forms part of the 'common treasure of thoughts that [humanity] transmits from one generation to another', and *reference* (*Bedeutung*) (Frege (1892), p. 29[5]). However, in modern formal semantics in the Montagovian tradition, despite its Fregean roots, conceptual structure has largely been set aside. In this latter tradition, Fregean sense has largely been substituted for the notion of *intension*, modeled non-psychologistically as, for example, a function from possible worlds to truth values.

We recover the classically Fregean notion of sense as including conceptual information and propose that both conceptual and referential aspects of meaning play a role in composition. Specifically, we can think of them as affording concept combination in different ways. Our use of the term *affordance* is based on Chemero's (2003) development of the notion, originally due to Gibson (1979); it is also inspired in Rietveld's (2008) extension of the notion to higher cognition. Chemero defines affordance as a relation between features of situations and abilities of organisms, and argues that to perceive an affordance is to recognize that the feature in question facilitates an action by the organism. The classic example is a mug with a handle: If a person who has never seen a mug gets to interact with it, it is very likely that she will grab it by the handle. The mug, by its shape, affords the grabbing-by-the-handle action on the part of the person.

Our extension of this idea to the case of language is very simple. We take the connection to concepts, on the one hand, and to the world, on the other, to be distinct features of language, each of which facilitates – that is, affords – a distinct composition process. If we posit that language users have access to both of these features and the corresponding processes that they facilitate, the tension we observed between default and highly context dependent interpretations in Section 2 disappears.

The default interpretations, we argue, can be understood as the result of conceptually-afforded concept composition. These are the interpretations that are immediately available in the absence of discourse context, and they are productive, suggesting that they build on regularities in our lexical knowledge – that is, the connections between words and concepts. For instance, the fact that physical objects typically have colored surfaces will afford the interpretation of a color term modifying a noun denoting a physical object as describing surface color, as in *red box* (see example (3-a)) and *red apple* (5-b). The fact that a writing instrument produces text or images with a particular color, and that this color may vary from one writing instrument to the next – part of our concept of what a writing instrument is like – affords the interpretation of *red pen* as a pen that writes in red (example (5-a)). Note that the use of a color term with *pen* is easily extended to other writing instruments with the same general properties, such as *pencil*, *crayon*, *marker*. Similarly, the fact that analyses are carried out using instruments, and that microscopes are instruments, affords the interpretation of *microscopic analysis* provided in (4-a) above. Different species of animals tend to require different climates, so again, the interpretation of (4-b), with *tropical* describing a climate, is in our view conceptually afforded.

---

[5]We cite the translation by Max Kölbel, published in Byrne and Kölbel (2009) which includes references to the original pagination by Frege.

Note that these interpretations arise from very detailed conceptual knowledge, presumably accessible from the words involved. The primitive-based analyses discussed in the previous section are too coarse-grained to allow for these interactions; the lack of conceptual information in typical formal semantic approaches doesn't allow for them either. Thus we need a richer and more nuanced lexical representation; in Section 4 we show how distributional semantics can serve this purpose.

The notion of conceptual affordance also allows us to make predictions about combinations of modifiers and nouns that will be infelicitous out of the blue. Interestingly, Vecchi et al. (2011) developed a computational model (using, it is worth noting, distributional semantics) that was able to partially distinguish between (out of the blue) deviant versus possible adjective-noun phrases. Vecchi, et al. randomly selected a set of phrases that were unattested in a very large corpus and tested whether their model would group them in ways that correlated with whether or not the phrase was acceptable to human judges. Examples of unattested but semantically acceptable phrases included *vulnerable gunman*, *huge joystick*, and *blind cook*; deviant phrases included, for instance, *blind pronunciation*, *parliamentary potato*, and *sharp glue*. The acceptable phrases are similar to those we have hypothesized above to involve conceptually afforded composition. For instance, joysticks are physical objects and have a size dimension that can be modified by *huge*. In contrast, it is not obvious in the absence of a specific context along which conceptual dimension a pronunciation could be blind, what kind of relation might exist between potatoes and parliaments, or what it would mean for glue, which is not rigid, to be sharp.

Now, it is possible to find a semantic interpretation for the allegedly deviant phrases. For instance, imagine that potatoes were thrown at parliamentary members in a protest concerning the recent economic crisis in Spain, and that one of the potatoes knocked out the president of the parliament and was retrieved and put on display. This object could well be dubbed *the parliamentary potato*. We submit that such interpretations are the result of referentially-afforded concept composition: they are retrievable only once we have a specific candidate (or small set of candidates) for who or what is being referred to with the phrase, along with a salient set of candidate properties that could be described by the modifier.

To further illustrate referentially-afforded composition, let us return to the use of "red box" in (3-b) and the "Canadian visit" example in (2). In the first case, the situation presents the hearer with two brown cardboard boxes. The speaker can assume that the hearer knows that the boxes each have a context-specific property of being destined to hold objects of a specific color. The use of *red* to modify the *box* color in this case is incongruent with what we can arguably consider the basic concept associated with box – the concept cannot afford any meaningful interpretation of the modifier – but the box referents and their context-specific properties can. In the case of (2), recall that the problem is that ethnic adjectives tend to express the agent when combined with eventive nouns. Thus, by default we expect *Canadian visit* to describe a visit made by Canadians. However, in (2) the interpretation on which Canada is the location visited is afforded by specific information about individuals in the context, namely, that Prince Edward and his wife are royalty members, that *Canada* denotes a place, and that Prince Edward and his wife are the agents of the action of beginning a (Canadian) visit.

In order for the distinction between conceptually and referentially afforded concept composition to have bite, we should have independent criteria for identifying the components of the specific concepts being combined. For now, we limit ourselves to the claim that once such criteria are established, it should be possible to predict when a combination of modifier and noun is easily interpreted in the absence of a specific discourse context.

3.2 Empirical evidence supporting the distinction

Despite the caveat made in the preceding paragraph, we consider it promising that the distinction between conceptually and referentially afforded concept composition accounts for puzzling data that we gathered in previous work and for which we had no explanation at the time. We now summarize these data and explain how our proposal makes sense of it. Note that this section does not present new empirical results but rather a novel, unified interpretation of previously obtained results.

The modification data reviewed so far point to the fact that modifier-noun combinations can have very plastic interpretations. We posit that a large part of this plasticity corresponds to referentially afforded composition. We use relational adjectives, introduced in Section 2, to illustrate.

As noted above, relational adjectives are typically denominal and, crucially, the adjective-forming morphology has been claimed to be essentially transparent (e.g. Spencer, 1999). The only contribution of the adjectival morphology, then, would be to make explicit that there is some relation between the referent of the noun from which the adjective is derived and the referent of the noun that the relational adjective modifies.[6]

By hypothesis, relational adjectives provide a way to pick up on, in a maximally condensed fashion, the myriad possible relations between the referent of the modified noun and the referent of the adjective's nominal stem (e.g. 'Canada', in the case of *Canadian*). These relations can be identified thanks to general knowledge (e.g. *national anthem*) or be inferred from the meaning of the modified noun (particularly when the noun is deverbal, as in *chemical reaction*); however, we posit that in many cases the relations are in fact afforded by specific information we have about the referents in question in the discourse (e.g. *Canadian visit* in example (2) above).

Boleda (2007) reported that, in Catalan, relational adjectives appear much more often in definite noun phrases than do other types of adjectives: Specifically, in an analysis of a 16.5 million-word, balanced Catalan corpus, relational adjectives appeared almost 60% of the time in definite noun phrases (59%, with a standard deviation of 15%), while other types of adjectives did so a little over 30% of the time. Definite noun phrases are used to refer to individuals that are familiar either from the context or from prior discourse, and referentially afforded concept composition is only possible when the referent is known. So, the high proportion of uses of relational adjectives inside definite noun phrases is suggestive of a tendency towards referential affordance in the composition of relational adjectives and nouns.[7] Without a distinction between the two kinds of composition, it is hard to account for the data in Boleda (2007).

Boleda and colleagues provided more data in the same direction in a statistical study of the British National Corpus (Boleda et al., 2012). The study compared nominal modification with a subclass of relational adjectives, namely ethnic adjectives (*Canadian*), to prepositional phrase

---

[6]This claim does not exclude the possibility that, over time, a relational adjective might come to be associated with a specific property or properties, such as the systematic use of ethnic adjectives to pick out characteristic properties of the individuals of a particular origin, as in (i).

(i)   Park's parents immigrated to the United States in the 1950s....Park says, "My parents thought the best way to help us succeed was to become very American...".
        (W. D. Thomas, *Korean Americans*, Benchmark Books, 2009, p. 59)

[7]Of course, that doesn't imply that other types of adjectives can't enter into referentially afforded composition (cf. the *red box* example in (3-b)), but relational adjectives, because of their properties, do so more.

modification (*from Canada*, *to Canada*, etc.). The results showed that ethnic adjectives are used especially when the discourse makes the semantic relationship between the head noun and the adjective explicit, that is, in contexts where previous information about the referent is available. Factors correlated with the use of these adjectives in the corpus, as opposed to their prepositional phrase counterparts, included, again, the definiteness of the DP containing the ethnic adjective, but also others like the occurrence of *visited Canada* before *Canadian visit* in a given discourse. To summarize, both studies constitute evidence for referentially afforded concept composition and show some effects that the use of this composition strategy by language users has on their linguistic output.

A second piece of evidence concerns one prediction of our proposal, namely, that the more context dependent (or referentially afforded) concept composition is, the more difficult it can be expected to be to reconstruct out of context. The results of a study involving computational modeling of adjectival modification, reported in Boleda et al. (2013), are in line with this prediction.

Boleda et al. (2013) used distributional semantic methods to produce meaning representations for adjective-noun phrases. They built representations for phrases like *former commentator* in two ways. On the one hand, they constructed a representation of the phrase directly from the corpus, that is, recording the contexts in which *former commentator* is used and building a representation from those contexts. We will call this representation the *observed* representation. On the other, they combined the distributional representations for the words using a computational algorithm. For example, this algorithm took the representation for *former* and that for *commentator* to build a semantic representation for *former commentator*. Boleda et al. then compared this *predicted* representation with the observed one, to see how accurate the prediction was.

Their results showed that the more typical the property denoted by the adjective is for the entity described by the noun, the easier it is to model the semantics of the phrase, that is, to approximate the observed representation for the phrase from the representation of its component words. For instance, *former* can be said to be a typical attribute for role-denoting nouns such as *commentator*, *father-in-law*, *teacher*, or *president*, insofar as the concepts associated with these nouns arguably include the specification that the role has a potentially limited duration. Information about duration is supplied by adjectives like *former*, *current*, or *future*. And indeed, the predicted representation for *former commentator* obtained by combining the representations for its two component words was very similar to the observed one. In contrast, the representation for phrases like *former colour* were more difficult to predict from the representations of the component words alone: *Colour* does not denote a concept with a clear temporal specification, and the relationship between *former* and *colour* will depend on the object whose colour is being referred to.[8]

We conclude that modification of nouns by adjectives describing typical attributes corresponds to conceptually afforded composition, and at least some uses of atypical modifiers correspond to referentially afforded composition. In the latter cases, without additional evidence from the specific discourse context (note that distributional semantic representations are aggregates over all uses of a word in a corpus and do not reflect discourse structure), it is hard to make sense of the semantic relationship between the adjective and the noun. Again, the distinction between conceptual and referential affordance in modification helps explain the results in Boleda et al. (2013).

Classical formal analyses of semantic composition involving adjectives are not well-equipped

---

[8]Typicality correlates positively with frequency, but the contrasting effects we mention are attested with phrases of similar frequency, including the ones used in the text. Thus, the results of this study cannot be explained simply as a byproduct of the frequency of the phrases in question.

to take into account the degree of fit or typicality relation between the property denoted by the adjective and general features of the concept associated with the noun. Theories like Pustejovsky's Generative Lexicon were designed to do this to some extent; however, as noted in Section 2, such theories cannot help with highly context-dependent meaning relations. Thus, the challenge is to find a way to incorporate the distinction between conceptually and referentially afforded concept composition into semantic theory, so as to broaden the theory's empirical coverage. As a first step in addressing this challenge, we turn to compositional distributional semantics.

4 Conceptually afforded composition with distributional semantics

We propose distributional semantics as a framework to account for conceptually afforded composition because we do not consider other approaches (e.g. standard formal semantics or primitive-based approaches such as the Generative Lexicon) to offer a rich enough representation of a word's meaning to account for the range of effects discussed. However, as it is beyond the scope of this paper to exhaustively compare these different approaches, we limit ourselves here to simply providing enough background on distributional semantics for the reader to be able to follow the formalization presented in the next section, leaving more thorough discussion for future work.

Distributional semantic analyses (Landauer and Dumais, 1997, Turney and Pantel, 2010) represent the semantics of a word as a function of the contexts it occurs in. Context can be defined in various ways, but the most typical approach is to define context as the words surrounding the target word in a corpus. A distributional representation for a word will then be a list of context counts, aggregated over the whole corpus and suitably transformed, that is, it is a *vector*. Figures 1 and 2 offer a toy example. Figure 1 is intended to illustrate how even a small context window reveals repeated examples of co-occurrences between a target word (here, *moon*) and other words that are suggestive of our knowledge about the target. Figure 2 exemplifies a partial vectorial representation for the words *moon*, *sun*, and *dog*.[9] The vectors show how the distributional representation mirrors some semantic similarities and differences between these words: All three can appear with *shadow*, but, while *moon* and *sun* appear with words such as *planet* or *shine*, *dog* does not. *Moon* and *sun* are similar in representation, but not identical: for example, *full* and *crescent* occur primarily with *moon*, *shine* is a more typical context for *sun* than for *moon*.

curtains open and the	moon	shining in on the barely
rough the night with the	moon	shining so brightly, it
surely under a crescent	moon	, thrilled by ice-white
is dazzling snow, the	moon	has risen full and cold
and the temple of the	moon	, driving out of the hug
in the dark and now the	moon	rises, full and amber
bird on the shape of the	moon	over the trees in front

Figure 1: The basic data for distributional semantic representations: contexts.

---

[9]Figures 1-4 are adapted from materials by Marco Baroni.

	planet	night	full	shadow	shine	crescent
moon	10	22	43	16	29	12
sun	14	10	4	15	45	0
dog	0	4	2	10	0	0

Figure 2: Semantic representation: vectors of context counts.

A vector for a word as used in distributional semantics ranges from a few hundred to a few thousand dimensions (contexts or transformations thereof), thus providing a very rich, flexible representation for word meaning. However, this makes it difficult to inspect it manually.[10] The power of distributional semantics lies in its use of well defined linear algebra techniques to manipulate these vectors, yielding useful information about the semantics of the involved words. We visualize one kind of technique used in Figure 3, where simple, two-dimensional vectors for the words *moon*, *sun*, and *dog* are visually represented. The two dimensions depicted in the graph (corresponding to word contexts) are *shadow* and *shine*, with the values shown in the left part of Figure 3. The geometric distance (e.g., the Euclidean distance; see discontinuous lines) between the vectors for *moon* and *sun* is smaller than the distance between the vectors for *moon* and *dog*. Crucially, the algebraic techniques that we can visualize with two dimensions generalize to any number of dimensions. Thus, in distributional semantics, geometric distance corresponds to semantic distance.

Distributional semantic methods are highly successful at modeling word meaning because they are based on linguistic data naturally produced by humans, as manifest in large text corpora drawn from the internet and other sources. The representations are rich, with hundreds or thousands of dimensions providing different bits of contextual information. Also, distributional representations are naturally graded; for instance, the notion of semantic distance is a continuum, with words being more or less distant. This makes them useful for semantic phenomena such as the typicality effect observed in the previous section.

Recently, researchers have begun to explore *compositional* distributional semantics, giving a distributional representation not only to words but also to phrases and even sentences (Mitchell and Lapata, 2010, Coecke et al., 2011, Socher et al., 2012, Baroni et al., 2014a, among many others); the previous work we presented at the end of Section 3.2 falls into this line of research. Here, the challenge is typically framed as capturing how composition changes the values of the vectors. For instance, *blood* is not a relevant context for *moon*, but when *red* modifies *moon* it does become relevant (see Figure 4). This kind of effect is achieved by applying composition operations to build the meaning representation of the phrase from the representations of its constituents. A very simple but stubbornly effective method is to simply add up the word vectors, as in Figure 4 (Mitchell and Lapata, 2010), but more sophisticated methods have been designed that sometimes yield better results (Baroni and Zamparelli, 2010). Nothing we say in this paper depends on the chosen method for composition, hence we will simply use **comp**($\overrightarrow{red}$, $\overrightarrow{moon}$) for the distributional representation of the phrase *red moon* obtained by applying a composition function to its constituent word vectors, $\overrightarrow{red}$ and $\overrightarrow{moon}$ (we represent word vectors with an overhead arrow, e.g. $\overrightarrow{moon}$).

---

[10]The context counts in a real distributional model are also typically further operated upon to remove noise in the data and make them more compact; see the references in this section for more information. Also, recently, neural network or deep learning models have been shown to outperform traditional count-based methods such as the ones explained in this section (Baroni et al., 2014b). Since nothing we say in this paper hinges on the particular type of model chosen, for clarity we present count-based models only.
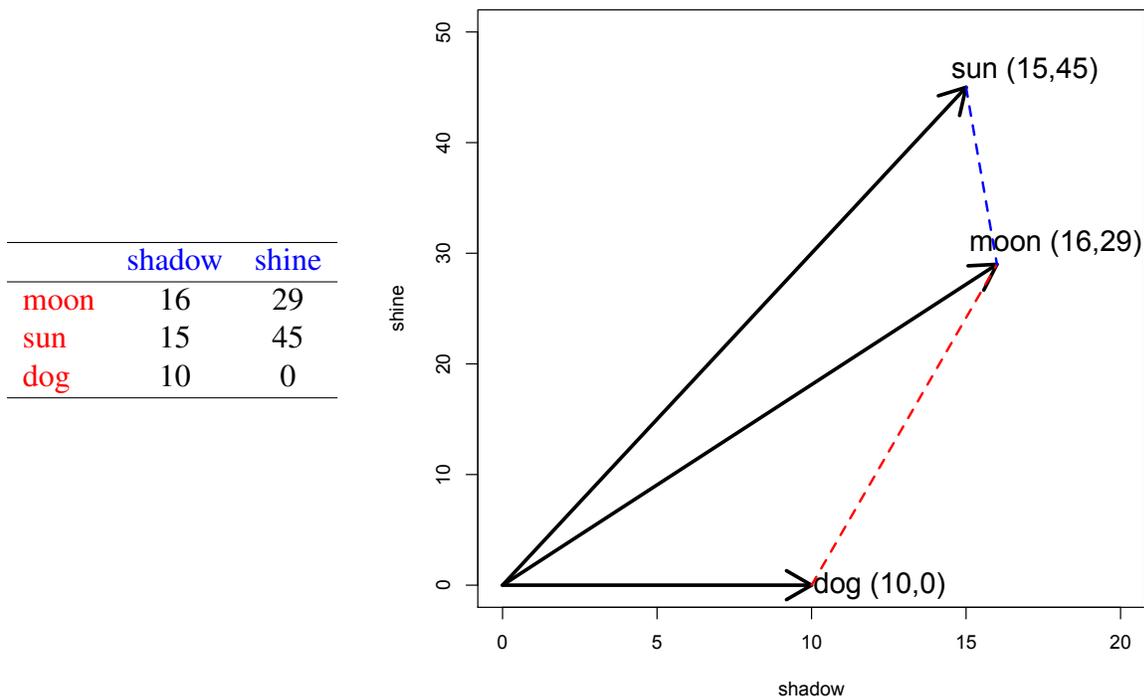
	shadow	shine
moon	16	29
sun	15	45
dog	10	0

Figure 3: Semantic distance as geometric distance.

	planet	night	blood
red	15	3	21
moon	24	15	1
red + moon	39	18	22

Figure 4: Vector composition by addition.

Note, finally, that there is an alternative method for obtaining a distributional representation for a phrase, namely, to directly extract it from the corpus, just as representations for words are generated (Figure 5). Because it is based on counts for actual occurrences of phrases in corpora, this representation should be a faithful rendering of the meaning of the phrase, and this is why we used it as a benchmark in the research in Boleda et al. (2013), to compare the result of compositionally obtained (predicted) vectors. However, this technique can only be used for sufficiently frequent phrases. Since many possible phrases will not occur frequently or even at all, composition is still needed to build a representation for many phrases.

Because of their data-driven nature and their rich representation of meaning, compositional distributional representations for phrases are able to account for subtle nuances of meaning arising from the combination of modifiers and nouns. For instance, Baroni and Zamparelli (2010) report that the most similar element in a large semantic space to the phrase *historical introduction* is *historical background*; to *small drop*, *droplet*; to *common understanding*, *common vision*. Though crude and incomplete as an approximation of what concepts are (as the discussion in Barsalou, this

a large	red moon	, Campana			planet	night	blood
a blood	red moon	hung over		red moon	34	20	31
glorious	red moon	turning
The round	red moon	, she 's
a blood	red moon	emerged
rains ,	red moon	blows,
monstrous	red moon	had climb
A very	red moon	rising is
under the	red moon	a vampire

Figure 5: Corpus-extracted distributional representation for phrase *red moon*.

volume, will make apparent), these representations have the advantage of being easy to construct and incorporable into a testable interpretive model. We therefore adopt them as the technique for modeling conceptually afforded composition in the formalization we offer in the next section.

5 A mixed model for two types of semantic composition

We next sketch how the mechanisms of conceptual and referential affordance can both be incorporated into a single, mixed interpretive model.[11]

We will use Discourse Representation Theory (DRT) as the scaffolding for our semantics. We use DRT because 1) the notion of discourse referent is crucial for implementing referentially afforded composition, and 2) the most recent research on compositional distributional semantics has not yet been able to show how such models can provide effective analyses of referential grounding or discourse dynamics (Bernardi et al., 2015, Sadrzadeh and Purver, 2015). This latter state of affairs leads us to tentatively hypothesize that compositional distributional semantics is best used to model only those parts of semantic composition that are, in our terms, conceptually afforded.

For reasons of space, we must assume basic familiarity with DRT; the reader is referred to e.g. Kamp (1981) or Kamp and Reyle (1993) for background. Our implementation of DRT will be entirely standard, with just three exceptions. First, we need a means of connecting distributional semantic representations to Discourse Representation Structures (DRSs). Second, as a result of doing this we will introduce minor modifications in our treatment of nominal and adjectival predication with respect to what is more generally assumed. Finally, we will need a way to distinguish conceptually afforded from referentially afforded composition.

We incorporate distributional semantics by building on the idea in Zamparelli (1995) that nouns (and not just certain kinds of generic noun phrases) denote Carlsonian kinds (Carlson, 1977).[12] The crucial step is to use distributional semantic representations rather than atomic abstract entities as models for kinds. However, as with the classic treatment of kinds as abstract entities, these distribu-

---

[11]Some of the basic discussion in this section concerning the integration of distributional semantics and formal semantics is drawn from McNally (2015); however, the application to DRT is new here, as is the idea of referentially-afforded composition. See Garrette et al. (2011) for a different approach to combining distributional semantics and Discourse Representation Theory, and see Pelletier, this volume, for extensive discussion of "two-tiered" semantic theories, of which he takes this proposal to be an example.

[12]Zamparelli posited that nouns come to denote sets of entities only in the semantic composition process, and used type-shifting operations licensed by functional morphosyntax to do this. See below.

13

tional representations will be coded in the DRSs as constants. Since distributional representations are mathematically-speaking vectors, the constants we use for them will be indicated with an overhead arrow (e.g. $\overrightarrow{\text{box}}$), as noted in the previous section.

We further extend Zamparelli's idea to adjectives, also interpreting them as vectors (e.g. $\overrightarrow{\text{red}}$). Since adjectives are not assumed to denote natural kinds but rather to pick out properties, this proposal can be seen as generalizing Zamparelli's, substituting concepts for both kinds and properties along the way. Distributional vectors will thus serve as very crude representations for concepts.[13] The crucial step will be to allow the distributional representations for nouns and adjectives to combine with each other to yield new representations of the same type, whose role in the DRT part of the semantics is exactly analogous to the role of the representations for unmodified nouns.

In the previous section we briefly sketched how the composition of two vectors works. We assume that the grammar of a language indicates when semantic composition for certain phrases involves the composition of vectors, as opposed to other sorts of semantic operations. The composition of vectors happens outside of the DRT model, but as the result is also a vector, it can, like the component vectors, be associated with a constant in a DRS, which we will represent as e.g. $\mathbf{comp}(\overrightarrow{\text{red}}, \overrightarrow{\text{box}})$. In other words, constants of the form $\overrightarrow{\text{red}}$, $\overrightarrow{\text{box}}$, and $\mathbf{comp}(\overrightarrow{\text{red}}, \overrightarrow{\text{box}})$ are all of the same type. Thus, distributional semantics will give us a relatively concrete algebraic model for simple and complex concepts on which both sorts of concepts are of fundamentally the same nature, much in the way lattice-theoretic structures serve as models for treating atomic entities and pluralities as fundamentally similar types of objects (Link, 1983).

The next piece we need is a way to exploit nouns and adjectives with such interpretations within DRT, so that referents can be associated with the concepts that nouns and adjectives pick out. Zamparelli used Carlson's (1977) realization relation, which we represent here as **Realize**, to do this: This relation holds between an object and a kind just in case the object constitutes an instance of the kind.[14] Again following Zamparelli, we assume that the **Realize** relation is introduced by (possibly abstract) functional morphosyntax that turns a noun into an expression that denotes a set of entities. As a first approximation, then, we can represent a referential expression such as *a box* as in (8), where u is the discourse referent introduced by the phrase, which must satisfy the condition that it is a realization of the concept $\overrightarrow{\text{box}}$.

(8)

u
$\mathbf{Realize}(u, \overrightarrow{\text{box}})$

Now consider modification. Prior to the point in the syntax at which the **Realize** relation is introduced, the composition operations at work will combine vector-denoting expressions; this corresponds to concept composition. We model conceptually-afforded composition as the result of

---

[13]We should insist that we are not making anything like the claim that concepts, whatever they are, consist only of distributional information, even if such information may play a role in concept formation. Rather, we are using distributional representations to model concepts primarily because they have certain properties that we hypothesize concepts to share and because they have various attractive methodological features, such as that of allowing us to make testable predictions of various kinds.

Note also that, to our knowledge, the model-theoretic semantics literature has largely avoided the deeper question of what kinds are and how they relate to concepts. Our reinterpretation of Zamparelli should not be viewed as reflecting any particular position on how kinds have been understood in this literature. See McNally (2015) for further discussion.

[14]Carlson's ontology also included stages (spatiotemporal slices) of individuals, which could also instantiate kinds, but these will not play a role in our discussion.

composing adjective and noun vectors directly into a new vector, corresponding to a complex concept, which can then stand in the **Realize** relation to a discourse referent, as in (9).

(9)
$$\boxed{\begin{array}{c} u \\ \hline \mathbf{Realize}(\text{u}, \mathbf{comp}(\overrightarrow{\text{red}}, \overrightarrow{\text{box}})) \end{array}}$$

The syntactic rules of the language will have to make it clear when this sort of composition can be appealed to and when not, but studies of the syntax of modification clearly indicate that syntax could, indeed, encode this kind of information. The fact that relational adjectives in various languages are known to have to occur closer to the noun they modify than other sorts of adjectives is just one sort of indicator that could be appealed to.

Now let us consider referentially-afforded concept composition. As mentioned in Section 3, this is attested only when the referent of the nominal is already familiar in the discourse. This referent therefore plays a role in the interpretation of the combination of the modifier and noun. We see two ways in which this could be implemented. One would be to take the referent to modulate the composition operation that combines the adjective and noun vectors. This could be represented as in (10), where the subscript u indicates modulation by referent u.

(10)
$$\boxed{\begin{array}{c} u \\ \hline \mathbf{Realize}(\text{u}, \mathbf{comp}_{\text{u}}(\overrightarrow{\text{red}}, \overrightarrow{\text{box}})) \end{array}}$$

On this view, the concept associated with *red* in the context would be exactly the same across all contexts, but its interaction with the concept contributed by the noun would vary from one context to the next, for example by the use of varying weights on the sums or products of the vectors.

Alternatively, the vector corresponding to the adjective could be modified as a function of the referent, i.e. reinterpreted as an ad hoc, referent-mediated property, as could be represented in (11).

(11)
$$\boxed{\begin{array}{c} u \\ \hline \mathbf{Realize}(\text{u}, \mathbf{comp}(f(\text{u}, \overrightarrow{\text{red}}), \overrightarrow{\text{box}})) \end{array}}$$

On this view, the composition operation as such is not altered in any way; rather the input to that operation is. In other words, *red* in this example would simply be associated with a different concept in the context in question. Further research will be needed to determine which of these options constitutes a better analysis of the facts, or, indeed, if they are empirically distinguishable. However, it is worth noting that this latter approach closely resembles the indexical interpretations of adjectives proposed by Bosch (1983) and Rothschild and Segal (2009), briefly introduced in Section 2.

These analyses do not offer an account of *how* context intervenes to determine the referentially afforded interpretation; in this, unfortunately, we are in good company, as no theory we know of offers such an account, and the area is one in which much more research is needed.

We close this section with some very brief, speculative comments on how the proposed analysis relates to classical analyses of adjective modification of nouns within formal semantics. Such modification has been analyzed in two ways: Either by treating the adjective as a second-order property that takes the noun as its argument, or by treating it as a first-order property that is combined via conjunction or set intersection with the (first-order) property denoted by the noun (see e.g. Kamp,

1975, Siegel, 1976, Larson, 1998, among many others for proposals and discussion). The latter analysis is appropriate specifically for cases of so-called intersective modification, when the adjectival and nominal properties are each entailed to hold of the individual being described. The former is more general and can be used not only for intersective modification but also for non-intersective modification, namely, subsective modification, where the adjectival property does not obviously directly hold of the referent but the noun property does (cf. *molecular biologist*); and intensional modification, where the nominal property is not entailed to hold (or is entailed not to hold) at the time or world of ascription (*former mayor* or *alleged thief*).

All of the above-sketched implementations of concept composition are counterparts of non-intersective modification. In no case is the concept contributed by the adjective directly related to the referent. Moreover, as we have set things up, our analyses of concept composition directly captures the intuition developed in Landman (2001) and Partee (2010) that all adjective-noun combinations, even intensional modification, are, in some sense, subsective, that is, the nominal description is always somehow used to identify the referent, insofar it contributes positively to the eventual complex description that the referent is related to via the **Realize** relation. Of course, it remains to explore how to reproduce the entailment effects of the world and temporal parameters that have played a role in traditional analyses of intensional adjectives, but we note that one surprising result of the study in Boleda et al. (2013) was that intensional adjectives turned out to be no more difficult to model in distributional semantics than other kinds of adjectives, insofar as, all other things being equal, compositional distributional semantic techniques could predict the semantic representation for phrases containing intensional adjectives from the representations of the component words just as well as they could for phrases containing nonintensional adjectives (see Section 4, above).

6   Conclusions

Semantic composition is a dynamic process that cannot be understood without simultaneously considering what we are referring to and the concepts associated with the words we are using. Concepts, and thus the words associated with them, encode significant regularities. At the same time, they are plastic, insofar as we must use a finite vocabulary to describe a potentially infinite variety of situations and generalizations in the world. However, once a word is applied to a referent, that word is grounded in a very specific manner, and the referent can influence the way we understand the word and its associated concept(s) in the context of use. This interplay between our conceptual structure and the world is what motivated the first contribution of this paper, namely to propose that modification works in two ways: It can be conceptually afforded, when the modifier and the head introduce concepts that fit to form a complex concept, and the speaker and the hearer use this fit in their interpretations; or referentially afforded, when the result of combining the modifier with the noun depends on specific properties of the referent. This proposal has an antecedent in Asher (2011), but we have made it more explicit and have proposed a specific analysis combining distributional semantics and DRT.

Along the way we hope to have made a case for further exploring distributional representations within semantic theory. They are automatically induced (and thus easy to construct and empirically well-founded), have some psychological plausibility (Landauer and Dumais, 1997), and offer a wealth of empirical data. Distributional semantic representations also avoid some of the weaknesses of semantic primitives: Since they generally encode a relatively large number of features

16

with continuous values,[15] they can express many more nuances of meaning than a small set of discrete features, while at the same time accounting for default interpretations. The key is to recognize their limitations. In this respect, we consider promising the division of labor between distributional semantics and a referential semantic framework like DRT.

References

Asher, N. (2011). *Lexical Meaning in Context*. Cambridge University Press, Cambridge.

Bally, C. (1944). *Linguistique générale et linguistique française*. A. Francke, Berne.

Baroni, M., Bernardi, R., and Zamparelli, R. (2014a). Frege in space. *Linguistic Issues in Language Technology*, 9(6):5–110.

Baroni, M., Dinu, G., and Kruszewski, G. (2014b). Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 238–247, Baltimore, Maryland. Association for Computational Linguistics.

Baroni, M. and Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of EMNLP*, pages 1183–1193, Boston, MA.

Barsalou, L. W. (2016). Cognitively plausible theories of concept composition. In Winter, Y. and Hampton, J., editors, *Compositionality and concepts in linguistics and philosophy*. Springer, Berlin. This volume.

Bernardi, R., Boleda, G., Fernandez, R., and Paperno, D. (2015). Distributional semantics in use. In *Proceedings of EMNLP 2015 Workshop LSDSem 2015: Linking Models of Lexical, Sentential and Discourse-level Semantics*. To appear.

Boleda, G. (2007). *Automatic acquisition of semantic classes for adjectives*. PhD thesis, Universitat Pompeu Fabra.

Boleda, G., Baroni, M., Pham, N. T., and McNally, L. (2013). Intensionality was only alleged: On adjective-noun composition in distributional semantics. In *Proceedings of IWCS 2013*, Potsdam.

Boleda, G., Evert, S., Gehrke, B., and McNally, L. (2012). Adjectives as saturators vs. modifiers: Statistical evidence. In Aloni, M., Kimmelman, V., Roelofsen, F., Sassoon, G. W., Schulz, K., and Westera, M., editors, *Logic, Language and Meaning - 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, Lecture Notes in Computer Science 7218, pages 112–121. Springer.

---

[15]In the examples we gave in Section 4, values are discrete; however, in actual distributional semantic models values are typically continuous as a result of the operations made on the original counts. See the references in that section for details.

Bosch, P. (1983). "Vagueness" is context-dependence. A solution to the Sorites Paradox. In Ballmer, T. and Pinkal, M., editors, *Approaching Vagueness*, pages 189–210. North Holland, Amsterdam.

Byrne, D. and Kölbel, M., editors (2009). *Arguing about Language*. Routledge, London.

Carlson, G. N. (1977). *Reference to Kinds in English*. PhD thesis, University of Massachusetts at Amherst.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, 15:181–195.

Clark, H. (1992). *Arenas of Language Use*. University of Chicago Press, Chicago.

Coecke, B., Sadrzadeh, M., and Clark, S. (2011). Mathematical foundations for a compositional distributed model of meaning. *Linguistic Analysis*, 36:345—384.

Del Pinal, G. (2015). Dual Content Semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics*, 8(7):1–53.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50.

Gagné, C. L., Spalding, T. L., and Kostelecky, M. (2016). Conceptual combination, property inclusion, and the Aristotelian-Thomistic view of concepts. In Winter, Y. and Hampton, J., editors, *Compositionality and concepts in linguistics and philosophy*. Springer, Berlin. This volume.

Garrette, D., Erk, K., and Mooney, R. (2011). Integrating logical representations with probabilistic information using Markov logic. In *Proceedings of IWCS 2011*.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin, Boston, MA.

Jespersen, O. (1942). *A Modern English Grammar on Historical Principles*. Einar Munksgaard, Copenhagen.

Kamp, H. (1981). A theory of truth and semantic representation. In Groenendijk, J., Janssen, T., and Stokhof, M., editors, *Formal Methods in the Study of Language*, volume 1, pages 277–322. Mathematisch Centrum, Amsterdam.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic*. Kluwer, Dordrecht.

Kamp, J. (1975). Two theories about adjectives. In Keenan, E., editor, *Formal semantics of natural language*, pages 123–155. Cambridge University Press, Cambridge.

Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes from Kaplan*, pages 481–614. Oxford University Press, New York.

Kayne, R. (1984). *Connectedness and Binary Branching*. Foris, Dordrecht.

Kennedy, C. and McNally, L. (2010). Color, context, and compositionality. *Synthese*, 174:79–98.

Landauer, T. and Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2):211–240.

Landman, M. (2001). Adjectival modification restricted. Ms, University of Massachusetts Amherst.

Larson, R. (1998). Events and modification in nominals. In Strolovitch, D. and Lawson, A., editors, *Proceedings from SALT VIII*, pages 145–168, Ithaca, NY. CLC Publication.

Lenci, A. (2008). Distributional approaches in linguistic and cognitive research. *Italian Journal of Linguistics*, 20:1–31.

Levi, J. (1978). *The Syntax and Semantics of Complex Nominals*. Academic Press, New York.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretic approach. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, Use, and Interpretation*, pages 302–323. de Gruyter, Berlin.

Martin-Löf, P. (1984). *Intuitionistic type theory: Notes by Giovanni Sambin of a series of lectures given in Padua in June 1980*. Studies in proof theory. Bibliopolis, Napoli.

McNally, L. (2015). Kinds, descriptions of kinds, concepts, and distributions. Ms., Universitat Pompeu Fabra.

McNally, L. and Boleda, G. (2004). Relational adjectives as properties of kinds. In Bonami, O. and Cabredo Hofherr, P., editors, *Empirical Issues in Syntax and Semantics*, volume 5, pages 179–196. http://www.cssp.cnrs.fr/eiss5.

Mitchell, J. and Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, 34:1388–1429.

Murphy, G. L. (2002). *The Big Book of Concepts*. MIT Press, Cambridge (etc.).

Ogden, C. K. and Richards, I. A. (1923). *The Meaning of Meaning*. Kegan Paul, London.

Partee, B. H. (2010). Privative adjectives: Subsective plus coercion. In Bäuele, R., Reyle, U., and Zimmermann, T. E., editors, *Presuppositions and Discourse: Essays offered to Hans Kamp*, pages 273–285. Emerald Group Publishing, Bingley, UK.

Pelletier, F. (2016). Compositionality and concepts A perspective from formal semantics and philosophy of language. In Winter, Y. and Hampton, J., editors, *Compositionality and concepts in linguistics and philosophy*. Springer, Berlin. This volume.

Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press, Cambridge, MA.

Rietveld, E. (2008). *Unreflective Action. A Philosophical Contribution to Integrative Neuroscience.* PhD thesis, University of Amsterdam.

Rothschild, D. and Segal, G. (2009). Indexical predicates. *Mind and Language*, 24(467–493).

Sadrzadeh, M. and Purver, M. (2015). Geometry of meaning from words to dialogue acts. In *Proceedings of the IWCS 2015 Workshop on Advances in Distributional Semantics*, London, UK.

Siegel, M. E. (1976). *Capturing the Adjective*. PhD thesis, University of Massachusetts at Amherst.

Socher, R., Huval, B., Manning, C. D., and Ng, A. Y. (2012). Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the joint meeting of the Conference on Empirical Methods in Natural Language Processing and the Conference on Computational Natural Language Learning (EMNLP-CoNLL)*, Jeju Island, Korea. http://aclweb.org/anthology//D/D12/D12-1110.pdf.

Spencer, A. (1999). Transpositions and argument structure. In Booij, G. and van Marle, J., editors, *Yearbook of Morphology 1998*, pages 73–101. Kluwer, Dordrecht.

Turney, P. and Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37:141–188.

Vecchi, E. M., Baroni, M., and Zamparelli, R. (2011). (linear) maps of the impossible: Capturing semantic anomalies in distributional space. In *Proceedings of the Workshop on Distributional Semantics and Compositionality*, pages 1–9, Portland, Oregon, USA. Association for Computational Linguistics.

Westerlund, M. and Pylkkänen, L. (2016). How does the left anterior temporal lobe contribute to conceptual combination? Interdisciplinary perspectives. In Winter, Y. and Hampton, J., editors, *Compositionality and concepts in linguistics and philosophy*. Springer, Berlin. This volume.

Zamparelli, R. (1995). *Layers in the Determiner Phrase*. PhD thesis, U. Rochester.