C. Kennedy Northwestern University
Ling 471 Fall 2003

Vector Space Semantics

1 Measurement across grammatical categories

So far, we have focused on the use and interpretation of measurement expressions in adjectives, nouns and verbs:

(1) *Adjectives*
  a. Kim is very tall.
  b. Kim is much taller than Lee.
  c. Kim is 2 meters tall.
  d. Kim is 50 centimeters taller than Lee.

(2) *Nouns*
  a. Kim bought a lot of fish.
  b. Kim bought 3 kilos of fish.
  c. Kim bought a 3 kilo fish.

(3) *Verbs*
  a. Kim worked a lot.
  b. Kim worries a lot.
  c. Kim raised the blind 3 centimeters.

As we observed at the beginning of the quarter, though, we also see expressions of measurement with prepositions that express spatial or temporal relations:

(4) *Prepositions*
  a. The bird is way above/behind/outside the house.
  b. Kim left long before/after Lee.
  c. The bird is 10 meters above/behind/outside the house.
  d. Kim left 10 minutes before/after Lee.

For the first three uses of measurement expressions, we assumed that the sentences crucially involve some sort of measure function, which maps objects to an appropriate scale. Can we do the same thing for measurement of directions/times in PP, and if so, what does this imply about preposition meanings?

Alternatively, might we take prepositions and the spatial (or temporal) relations they express as somehow basic, and try to analyze everything else in terms of them?

2 Varieties of spatial meaning

Perhaps one reason we might think of spatial meaning as somehow 'basic' is that all grammatical categories have terms whose meanings are spatial. As Zwarts (2000) points out, prepositions 'specialize' in this sort of meaning, but we find the spatial meanings *place* and *change of place* encoded in adjectives, nouns, verbs (and adverbs) as well:

(5) *Place terms*

a. The milk is in the refridgerator.
b. We met behind the stature.
c. He lives in the vicinity of the library.
d. He lives far from here.

(6) *Change of place terms*

a. They moved into the apartment.
b. I led him across the yard.
c. They entered the bar.
d. The airplane went down/descended.

Zwarts also suggests that we can think of the following categories as basically spatial:

(7) *Size*

a. a thin hair
b. the height of the building
c. to shorten a rope
d. to grow

(8) *Orientation*

a. an oblique line
b. to stand upright
c. He kicked it over.
d. two rotations/to rotate

(9) *Shape*

a. a straight stick
b. a circle
c. to straighten/bend a stick

(10) *Spatial parts*
top, backside, corner, center, bump, notch

Furthermore, as I mentioned last week, Jackendoff (1996) argues that the semantic similarity that unites the three classes of verbs in (11) is (possibly metaphoric) movement along a path.

(11) *Creation/consumption verbs*

a. Kim ate <u>rice</u> for an hour. ATELIC
b. Kim ate <u>a bowl of rice</u> in an hour. TELIC

(12) *Directed motion verbs*

a. <u>The balloon</u> ascended for an hour. ATELIC
b. <u>The submarine</u> ascended in an hour. TELIC

(13) *Degree achievements*

a. The dripping water lengthened <u>the icicle</u> for an hour. ATELIC
b. The tailor lengthened <u>my pants</u> in an hour. TELIC

Let's explore the possibility that we can build a unified semantics for expressions of measurement on top of a basic semantics for space.

3 Vector space semantics

3.1 The basics

The basic idea: terms from the domains of place, size, orientation and spatial parts are all interpreted with respect to the same spatial structure, which has the form of a three-dimensional *vector space*, where vectors can be thought of most generally as directed line segments between points in space.

The initial motivation for using vectors comes from the interpretation of measurement terms with PPs. An initially plausible analysis of the meaning of a PP like *behind the tractor* is that it denotes a general region of space that stands in a particular relation to the tractor.

However, according to Zwarts, if we don't assign any more structure than this to the meaning it would be difficult to handle modified PPs like the following:

(14) a. 3 meters behind the tractor
b. very far behind the tractor
c. just behind the tractor
d. straight behind the tractor

Intuitively, what the different modifiers in (14) are doing is measuring or further specifying the properties of the 'bare' spatial relation expressed by *behind the tractor*. But if this relation is one that just gives us a general region of space, what exactly would these phrases be doing?

If we think of the PP as giving us instead a set of vectors — those that extend from the tractor in directions that lead away from it relative to some reference position in front — then we can come up with a very clear analysis of the modifiers: they measure vectors.

(15) a. $[\![\text{behind the tractor}]\!] = \{v \mid v$ *originates at the tractor and extends beyond it*$\}$
b. $[\![\text{2 meters behind the tractor}]\!] = \{v \mid v$ *originates at the tractor and extends beyond it and v measures 2 meters*$\}$
c. $[\![\text{straight behind the tractor}]\!] = \{v \mid v$ *originates at the tractor and extends beyond it and v is perpindicular to the facing plane of the tractor*$\}$

The formal components of VSS are as follows (see Zwarts 1995; Zwarts and Winter 1997; Winter 2001):

1. A vector space V over the real numbers $\mathbf{R}$.
2. An addition operation $+$ on elements of V.
3. A zero element 0 that satisifes $v + 0 = v$ for al $v \in V$.
4. An opposite element -v for each $v \in V$, such that $v + (-v) = 0$.
5. A scalar multiplication operator $\cdot$ between real numbers in $\mathbf{R}$ and elements in V such that for all $s \in \mathbf{R}$ and $v \in V$, $s \cdot v \in V$.
6. A norm function $|\ |$ from V to non-negative elements of $\mathbf{R}$.

Zwarts 2000 makes a distinction between two types of vectors: *place vectors*, which represent the relative position of an object wrt a point of reference; and *axis vectors*, which represent the axial orientation of an object.

Some basic notation:

(16) a. $place(x, v) = x$ is placed at vector v (v is directed towards x)
b. $place(v, y) =$ vector v is placed at y (y is the origin of v)
c. $place(x, v, y) = v$ points from y to x
d. $axis(x, v) = x$ has an axis v

3.2 The semantics of locative PPs

As we saw informally above, PPs in this framework are analyzed as sets of located vectors ('vector descriptions'). Below are some more precise definitions of a few different kinds of PPs.[1]

(17) a. $[\![\text{inside NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |v| \prec 0\}$
b. $[\![\text{outside NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |v| \succ 0\}$
c. $[\![\text{on NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |v| = 0\}$
d. $[\![\text{near NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |v| \prec r\}$, where r is a contextually specified 'small' amount.

To handle directional prepositions, we add the *axis functions* VERT, FRONT, LAT (others?), their inverses (-VERT, etc.), and their orthogonal complements ($\perp$VERT, etc.):

(18) a. $[\![\text{above/over NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |\textsc{vert}(v)| \succ |\perp \textsc{vert}(v)|\}$
b. $[\![\text{below/under NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |\text{-}\textsc{vert}(v)| \succ |\perp \text{-}\textsc{vert}(v)|\}$
c. $[\![\text{in front of NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |\textsc{front}(v)| \succ |\perp \textsc{front}(v)|\}$
d. $[\![\text{behind NP}]\!] = \{v \mid place(v, [\![\text{NP}]\!]) \wedge |\text{-}\textsc{front}(v)| \succ |\perp \text{-}\textsc{front}(v)|\}$

Measure phrases and PP-degree modifiers can also be analyzed as sets of (located) vectors:

(19) In the context $[_{\text{PP}}$ __ $[_{\text{P}'}$ ...]]:
a. $[\![\text{3 meters}]\!] = \{v \mid |v| = 3m\}$
b. $[\![\text{(very) far}]\!] = \{v \mid |v| > r\}$, where r is a contextually specified 'large' amount
c. $[\![\text{just}]\!] = \{v \mid |v| \approx 0\}$
d. $[\![\text{straight}]\!] = \{v \mid \textsc{vert}(v)\}$

This lets us deal with modified PPs in terms of regular old restrictive modification (set intersection):

(20) $[\![[_{\text{PP}} \text{ XP } [_{\text{P}'} \text{ ... }]]]\!] = [\![\text{XP}]\!] \cap [\![[_{\text{P}'} ...]]\!]$

(21) $[\![[_{\text{PP}} \text{ 3 meters [behind the tractor]}]]\!] = [\![\text{3 meters}]\!] \cap [\![\text{[behind the tractor]}]\!]$
$= \{v \mid |v| = 3m\} \cap \{v \mid place(v, [\![\text{the tractor}]\!]) \wedge |\text{-}\textsc{front}(v)| \succ |\perp \text{-}\textsc{front}(v)|\}$
$= \{v \mid place(v, [\![\text{the tractor}]\!]) \wedge |\text{-}\textsc{front}(v)| \succ |\perp \text{-}\textsc{front}(v)| \wedge |v| = 3m\}$

[1] An important part of the system that I'm not going to represent here is that only vectors that start from 'edges' of objects are considered. This basically ensures that in something like *10 meters behind the house*, we start measuring from the edge of the house, not from somewhere in the middle (see Zwarts 1995; Zwarts and Winter 1997; Winter 2001 for discussion).

It's clear that the framework gives us what looks to be exactly what we want (though we need to say a bit more about how we turn sets of vectors into properties of individuals, or alternatively, how we compose an individual and a set of vectors to get a proposition), but if we're interested in asking whether VSS provides a good framework for representing measurement, isn't just saying $= 2m$ kind of cheating?

I'll come back to this later, but first let's see how the basic sytem would extend to the analysis of other categories.

3.3 A VSS for gradable adjectives

3.3.1 First try

Zwarts (2000) observes the following semantic similarities (I'll focus mainly on the case of gradable adjectives, though Zwarts talks about how the system can extend to other categories as well):

(22) a. x is far from y = the distance between x and y is long
b. x is close to y = the distance between x and y is short

(23) a. x is long = one end of x is far from the other end
b. x is short = one end of x is close to the other end

Zwarts tries to capture this similarity by analyzing the prepositions in terms of place vectors and the adjectives in terms of axis vectors; in all of these examples r is some contextually defined standard:

(24) a. x is far from $y = \exists v[place(x, v, y) \wedge |v| > r]$
b. x is long $= \exists v[axis(v, x) \wedge |v| > r]$

(25) a. x is close to $y = \exists v[place(x, v, y) \wedge |v| < r]$
b. x is short $= \exists v[axis(v, x) \wedge |v| < r]$

It should be fairly clear that this is a little too simplistic for the adjectives. First, what we really want to know is what the meaning of the *adjective* is, not what we get at the end of the day. In particular, we want to know about markedness (e.g., the difference in 'bias' between *how far/long* vs. *how close/short*).

The logical solution is to say that adjective meanings are just sets of vectors, but then we have to move away from talking in terms of axis vectors, I think.

Second, we need more content to the meanings if we want to distinguish $long_{space}$ from $long_{time}$, *long* from *wide* from *deep*. The former type of distinction is actually something that VSS may provide a nice handle on, since here we can take advantage of the various 'perspective functions':

(26) a. x is long $= \exists v[axis(v, x) \wedge \text{FRONT}(v) \wedge |v| > r]$
b. x is wide $= \exists v[axis(v, x) \wedge \text{LAT}(v) \wedge |v| > r]$
c. x is deep $= \exists v[axis(v, x) \wedge \text{VERT}(v) \wedge |v| > r]$

Maybe temporal adjectives can be handled in terms of vector paths? What about adjectives that involve non-spatial gradable properties like *expensive, intelligent, happy*, etc? We're going to need more than just perspective functions for these guys: we're going to need dimensions.

3.3.2 Second try

Winter (2001) provides an extension of VSS for gradable adjectives that has the features we want (which is itself a variant of the extension developed in Faller 2000). The crucial definitions are in (27).

(27) Let V be a vector space over $\mathbf{R}$ with norm function $|\ |$. Let $U \subset V$ be a (finite) set of unit vectors called *scale units*. A pair $S = \langle u_S, X_S \rangle$ where $u_S \in U$ and $X_S \subseteq \mathbf{R}$ is called a *scale* over V. The set of *values* of a scale S is the set of vectors $\{s \cdot u_S \mid s \in X_S\}$.

The crucial parts of this are the scale unit component and the definition of scalar values.

Scale units in effect define a bunch of different vector spaces — as many vector spaces as we can come up with scale dimensions. As Winter says: "...the *height* scale for the adjectives *tall* and *short* is defined using a unit vector u_h.... A tempora scale for the adjectives *early* and *late* is similarly defined using a unit vector u_t (*different than* u_h)...." (my emphasis).

So, the restriction that the 'values' of a scale come from the set $\{s \cdot u_S \mid s \in X_S\}$ is basically a way of sorting them, i.e., of turning them into what we have been calling 'degrees'.

Furthermore, the scalar multiplication operation gives us the sort of linear structure we want, since this has the effect of 'extending' a basic vector (see Zwarts 1995).

Winter can now give definitions for adjective meanings along the following lines:

(28) For the length scale $L = \langle u_L, (0, \infty) \rangle$ and contextual standard r:

a. $[\![\text{long}]\!] = \{s \cdot u_L \mid s > 0 \wedge s > r\}$

b. $[\![\text{short}]\!] = \{s \cdot u_L \mid s > 0 \wedge s < r\}$

In other words, an AP headed by *long* denotes the set of length-unit vectors whose size is greater than the standard, and *short* denotes the set of length-unit vectors whose size is less than the standard.

Aside: I actually think this is the wrong analysis of polar opposition, and moreover don't think that the standard of comparison should be part of *adjective* meaning (as opposed to the AP meaning), but that's sort of beside the point. I suspect that my preferred way of thinking about things could be incorporated into this system one way or the other.

This sort of system can also handle the difference between e.g., *long, wide* and *deep*, and presumably would take advantage of some of the same semantic primitives used by spatial prepositions. Certainly if the same underlying representational system is being used in the adjectival domain and the prepositional one, we expect such overlap.

However, the actual lexical semantics of gradable adjectives and prepositions is not quite the same: preopositions don't have a 'scale unit' component, and this component crucially allows for the 'dimensional variablity' we want with adjectives.

3.4 Measurement in VSS

One potential worry about this system might be that the scale unit component generates an expectation that we should have measurement units defined for all scales — e.g., *10 hedons of happiness, 98 Einsteins of intelligence*, and so forth.

Well, maybe we do...? However, this prediction is not really there, because scale units aren't exactly the same thing as measure units.

Winter (2001): "Measure units such as *meter* or *year* are assumed to specify constant real numbers, defined relative to the norm of the vector space."

In other words, measure units are arbitrarily chosen values, and (vector) measurement can be formalized in terms of the norming operation:

(29) $[\![\text{two meters}]\!] = \{v \in V \mid |v| = 2m\}$, where $2m$ is a value in $\mathbf{R}$.

We have to say a bit more, though, to capture unit/scale correlations like those in (30), and to explain the (general) absence of MPs like *98 Einsteins.*

(30)
a. 2 meters/??years tall
b. 2 ??meters/years old
c. 2 meters/??years in front of the barn
d. 2 ??meters/years after the fall of Rome

If unit terms just mapped to real numbers, then it's unclear what rules out the bad examples in (30). Here is Winter's semantics for *two meters tall*:

(31)
$$[\![\text{two meters tall}]\!] = \{v \in V \mid |v| = 2m\} \cap \{t \cdot u_H \mid t \succ 0 \wedge t \succ r_H\}$$
$$= \{t \cdot u_H \mid t = 2m \wedge t \succ r_H\}$$

It's not clear to me what rules out *2 years tall*: at the end of the day, we're going to get (32), which should be perfectly interpretable if $2y$ is just a number.

(32)
$$[\![\text{two years tall}]\!] = \{v \in V \mid |v| = 2y\} \cap \{t \cdot u_H \mid t \succ 0 \wedge t \succ r_H\}$$
$$= \{t \cdot u_H \mid t = 2y \wedge t \succ r_H\}$$

So we need to allow measure units to be defined only for specific types of scale units. Something like the following?

(33)
a. $[\![\text{meter}]\!] = \{t \cdot u_L \mid t = m\}$
b. $[\![\text{year}]\!] = \{t \cdot u_T \mid t = y\}$

This ought to feed into the same sort of compositonal semantics for MPs that we were considering earlier in the quarter.

4 Concluding thoughts

Quantities vs. measures

References

Faller, Martina. 2000. Dimensional adjectives and measure phrases in vector space semantics. In *Formalizing the dynamics of information*, ed. Martina Faller, Stefan Kaufmann, and Marc Pauly, pp–pp. Stanford, CA: CSLI Publications.

Jackendoff, Ray. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory* 14:305–354.

Winter, Yoad. 2001. Measure phrase modification in a vector space semantics. In *Proceedings of WCCFL 20*, ed. K. Megerdoomina and L. A. Bar-el, 607–620. Somerville, MA: Cascadilla Press.

Zwarts, Joost. 1995. The semantics of relative position. In *Semantics and Linguistic Theory 5*, 405–422.

Zwarts, Joost. 2000. Vectors across spatial domains: From place to size, orientation, shape and parts. To appear in Emile van der Zee and John Slack, eds., *Axes and Vectors in Language and Space*. Oxford: Oxford University Press.

Zwarts, Joost, and Yoad Winter. 1997. A semantic characterization of locative PPs. In *Semantics and Linguistic Theory 7*, ed. Aaron Lawson, 294–311. Ithaca, NY: CLC Publications.