Measurement scales in natural language

Stephanie Solt
Centre for General Linguistics (ZAS), Berlin
solt@zas.gwz-berlin.de

August 1, 2014

Measurement scales in natural language

August 1, 2014

Abstract

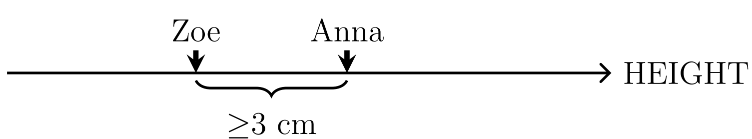
The meaning of a wide variety of natural language expressions can be stated in terms of degrees on scales. This paper provides an overview of the contribution of scales to linguistic meaning. I discuss the ontology of scales, survey the role of scalarity within and across languages, and summarize recent research into the linguistically relevant features of scales, organizing these findings into a preliminary framework for a comprehensive model of scale structure. The relation between the scales underlying the interpretation of gradable expressions and those involved in the calculation of scalar implicatures – so-called Horn scales – is also briefly discussed.

1 Introduction

The interpretation of (1a) involves a scale. It might be paraphrased as saying that the point on the scale of height that corresponds to Anna is (at least) 3 cm above or beyond the point corresponding to Zoe, a situation depicted visually in (1b).

(1) a. Anna is 3 cm taller than Zoe.

b.



This paper is about the contribution of scales to natural language meaning. Semantic approaches that make reference to scales have a long tradition, going back to foundational works including Bartsch and Vennemann (1973); Seuren (1973); Cresswell (1977); von Stechow (1984); Cruse (1986); Bierwisch (1989); and Klein (1991), and this area (now typically called 'degree semantics') has seen an upsurge of interest in recent years. It is now well established that scalarity of the sort in (1) is ubiquitous in language, and that a wide variety of linguistic phenomena can be explained in terms of properties of scales and variation in their structures. The goal of this review is to consolidate the quite divergent threads of research on this topic into a unified 'state-of-our-knowledge' picture of the nature and role of the scales that underlie meaning in language. In doing so, I will highlight some lesser understood aspects of scales and their structures, where further research has the potential to answer still open questions in semantics and pragmatics.

The organization of the paper is as follows: Section 2 addresses some fundamental ontological questions relating to the nature and origin of scales as they are invoked in semantic analysis. Section 3 surveys the linguistic domains where scalarity is present, and considers possible cross-linguistic variation in the role of degrees and scales. Section 4 summarizes recent research into the linguistically relevant properties of scales, organizing these findings into a preliminary framework for a general model of scale structure. Finally, Section 5 briefly discusses the relationship between measurement scales of the sort in (1) and the so-called Horn scales involved in the calculation of pragmatic inferences.

2 The ontology of scales

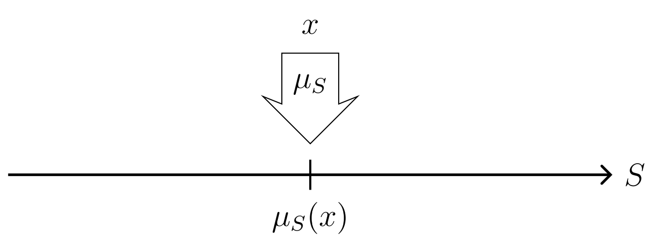
As a starting point, let us take the fairly common view (e.g. Kennedy 2007) that a scale S is a triple of the following form:

(2) $S = \langle D, \succ, DIM \rangle$, where

- D is a set of degrees,
- $\succ$ is an ordering relation on D, and
- DIM is a dimension of measurement.

Degrees are linked to individuals via measure functions: μ_S is the function that maps an individual x to the degree on the scale S that represents x's measure with respect to the dimension *DIM*:

(3)



To continue with our previous example, *tall* can be taken to have a lexical entry based on a HEIGHT measure function, i.e. a function that maps individuals to their heights, as in (4) (from Heim, 2000). The degree argument in (4) is supplied by some form of degree morpheme; this might be a degree modifier (e.g. *too*, *very*), a measure phrase (e.g. *1.7 meters*), or simply comparative or superlative morphology. In the case of the unmodified 'positive' form of the adjective (as in *Anna is tall*), a common assumption is that this role is played by a phonologically null degree morpheme *pos*, which introduces a context-dependent threshold or standard of comparison (Cresswell, 1977; Heim, 2006; Kennedy, 2007; von Stechow, 2009).

With the appropriate entry for the comparative morpheme *-er* (cf. Beck, 2011), (1) thus receives the interpretation in (5):

(4) $[\![tall]\!] = \lambda d \lambda x . \mu_{HEIGHT}(x) \succeq d$

(5) $\mu_{HEIGHT}(Anna) \succeq \mu_{HEIGHT}(Zoe) + 3\ cm$

But the schema in (2) places only loose restrictions on what scales might look like, and leaves open a wide range of possibilities for how they could vary in their structures. For example, nothing in (2) requires that a scale have a corresponding unit of measurement, though some scales (e.g. that associated with height) clearly do. Nor is it specified what sort of things degrees are, or what properties the relation $\succ$ must, or may, possess.

The question of the structure of scales is closely related to the question of their ontological status. Here, a range of perspectives can be distinguished.[1] Some authors take degrees and scales to be some sort of abstraction, which embody our ability to judge magnitudes and make comparisons. For von Stechow, 'whatever they are, they are highly abstract objects' (von Stechow, 1984, p. 47). Somewhat similarly, Kennedy and others describe degrees as 'abstract representations of measurement' (Kennedy, 2007, p. 3). A related view takes degrees to be numbers (Krifka, 1989). Although this point is not always made explicit, underlying this view is the assumption that degrees are a primitive component of the ontology, which have an existence independent of the entities whose measurements they encode. This is often reflected by assigning degrees their own semantic type, type *d*.

Other authors propose that degrees and scales are derived in some way from elements already assumed as part of the ontology. Bierwisch (1989) takes a step in this direction when he proposes that degrees, while being 'mental entities', are produced via the operation of comparing individuals: "there is no degree without comparison and no comparison without degree" (p. 112). A yet more concrete and still very influential view is due to Cresswell (1977), who notes that degrees can be construed as equivalence classes of individuals (for other work in this tradition, see Klein, 1991; Bale, 2008, 2011; Sassoon, 2010; van Rooij, 2011a; Lassiter, 2011). The underlying intuition is that Anna's 'degree' of height, for instance, can be understood as the set of individuals of the same height as Anna. Formally this is implemented by starting with a comparison relation R between individuals, such as a relation of 'taller than' or 'more beautiful than'. The equivalence classes under this relation become the degrees of the scale (where a and b are equivalent under R iff for all c, *R(a,c)* iff *R(b,c)*, and *R(c,a)* iff *R(c,b)*). A relation $\succ$ between degrees is then defined on the basis of the original relation between individuals. In this way, a scale is derived without the need to posit a pre-existing abstract notion of degree.

Authors who have adopted Cresswell's degree-as-equivalence-class view have typically sought to ground degree-based semantic frameworks in measurement theory, the branch of applied mathematics concerned with the numerical representation of properties of and relationships between entities (for basic introductions to measurement theory, see Kranz et al. 1971; Roberts 1985). The benefit of such an approach is that it makes clear predictions

[1] The discussion in this section owes a debt to Klein (1991); Sassoon (2010); van Rooij (2011a); Lassiter (2011).

about possible scale structures, and as such has the potential to serve as the starting point for a linguistically relevant typology of scales, something that does not emerge from an approach that takes scales and degrees to constitute a separate and purely abstract domain.

The scale derived via the equivalence class procedure described above is a simple linearly ordered set of points. In the typology of measurement levels introduced by Stevens (1946), it is an **ordinal** scale. This is sufficient for the interpretation of a simple comparative such as (6a). But while degrees on an ordinal scale can be assigned numerical values in an order preserving way, this does not suffice to make examples such as (6b,c) meaningful:

(6) a. The Empire State Building is taller than the Washington Monument.
 b. The Empire State Building is 695 feet taller than the Washington Monument.
 c. The Empire State Building is twice as tall as the Washington Monument.

These require a stronger or more informative scale structure: (6b) requires an **interval** scale, on which distances between scale points are meaningful; (6c) requires a **ratio** scale, featuring in addition a non-arbitrary zero point.

To derive a stronger (ratio level) scale, the equivalence class procedure must be augmented by introducing a concatenation operation on the domain of individuals and a corresponding mathematical operation on degrees. To continue with height as an example, the concatenation of two individuals can be conceptualized as stacking them one on top of the other, with the corresponding mathematical operation being addition: the height of two individuals stacked together is the sum of their individual heights. A standard element can then be selected to form the basis for a unit of measurement (e.g. an object exactly 1 foot tall). By this means, the structure needed for the interpretation of examples such as (6b,c) can be obtained.

There is evidence that natural language is sensitive to something like this hierarchy of measurement levels. For example, Sassoon (2010) shows that contrasts such as those in (7) can be attributed to distinctions of this sort: positive adjectives such as *tall* are based on ratio scales, while their negative antonyms (e.g. *short*) are based on interval scales, lacking a fixed zero point.

(7) a. twice as tall / 5 feet tall
 b. *twice as short / *5 feet short

There are other data, however, that are more problematic for a classification grounded in measurement theory. For example, the felicity of (8) would suggest that the scales tracking evaluative dimensions such as beauty, intelligence and importance are ratio level, the same as the scale of height.

(8) Anna is twice as beautiful/smart/important as Zoe.

The first puzzle is what sort of concatenation operation could be invoked in these cases to derive a scale of this structure. Beyond this, if it is concluded from (8) that both *tall* and *beautiful* etc. invoke ratio scales, the following contrast is quite unexpected:

(9) a. The Empire State Building is 2.14 times as tall as the Washington Monument.
b. ??Anna is 2.14 times as beautiful/smart/important as Zoe.

Apparently there is in fact some difference in structure between the scales lexicalized by the two sorts of adjectives; but this does not correspond to any distinction standardly made in measurement theory.

As this discussion shows, it has not yet been established that the scales underlying natural language meaning can uniformly be given a measurement-theoretic basis, at least in the form described. Beyond this, there is evidence that the typology of scale structures assumed in measurement theory does not adequately characterize the range of variation in natural language scales.

I return to this topic in Section 4, where I take up the issue of variations in scale structure and their linguistic consequences, and outline a preliminary framework for a more linguistically oriented typology of scales. Before this, we will take a detour to investigate the linguistic domains where scalarity plays a role.

3 Scales across language(s)

3.1 Which categories are scalar?

I began this paper with an adjectival example because gradable adjectives such as *tall* are typically taken to be the paradigm case of scalarity.

Here it should be pointed out that it is not universally accepted that all adjectives that are gradable (i.e., that form comparatives and combine with degree expressions) necessarily have underlying scalar semantics (in the sense of introducing a degree argument). An alternate theoretical framework, the so-called delineation approach associated in particular with Klein (1980, 1982), takes gradable adjectives to denote context-dependent, partial functions from individuals to truth values, which induce a partition of a comparison class into a positive extension (e.g. the tall things), a negative extension (the not-tall things) and an extension gap (the things in between). A notion of scale can be layered on top of this basic approach to account for examples such as (1), but plays no role in the underlying lexical semantics of the adjective.

Kennedy (2007) provides arguments in favor of the degree approach over the delineation one, but the debate remains an ongoing one (see e.g. McNally, 2011; Solt, 2011; Burnett, 2012); I refer the reader to the cited works for further discussion.

Leaving this issue aside, it has been recognized since Sapir (1944) and Bolinger (1972) that scalarity also plays an important role outside of the adjectival domain. The following briefly surveys some findings in these areas.

Quantity and amount. Parallels of the following sort were already discussed by Bartsch and Vennemann (1973) and Cresswell (1977):

(10) a. Arabella is more beautiful than Tom.

b. Chicago has more residents than Austin.

c. We bought more milk than wine.

If (10a) expresses the comparison of two degrees on the scale of beauty, it seems that (10b,c) must likewise be analyzed as involving comparison of degrees on some appropriate scales. In fact, most current work on the semantics of quantity expressions assumes either explicitly or implicitly that their meaning must be stated with reference to degrees. In examples like (10b), the scale is the number line, that is, the integers, or possibly the rational or real numbers (see Section 4.2.1); in (10c), it is one corresponding to a mass dimension, here perhaps liquid volume.

An influential work in this vein is Hackl (2000), who proposes that *many* be analyzed as a parameterized quantifier whose first argument is a degree argument (note the parallel to the adjectival entry in (4)).

(11) $[\![many]\!] = \lambda d \lambda P \lambda Q . \exists x \, [P(x) \wedge Q(x) \wedge |x| \succeq d]$

Here $|x|$ is a cardinality operator, i.e. a counting measure function, which maps a plurality to a number that corresponds to its cardinality. This semantics allows the standard analysis of adjectival comparatives such as (1a) and (10a) to be extended directly to quantificational *more*, which can be decomposed as *many* + *-er* in (10b) and as *much* + *-er* in (10c) (cf. Bresnan 1973). Hackl (2009) applies a similar approach to account for facts relating to superlative *(the) most*, and Rett (2008) and Solt (2014b) develop proposals according to which *many* and similar words are not quantifiers at all, but rather have purely degree-based semantics.

Number words (e.g. *three*) and modified numerals (*more than three*, *at least four*) are also readily analyzed in a scalar semantics. Compositionally these can be handled in a variety of ways: some authors posit a counting measure function as a part of the lexical meaning of the numeral itself (Partee, 1987; Krifka, 1999), while others take numerals at the most basic level to denote degrees, with the measure function encoded elsewhere, perhaps in a phonologically null syntactic element (Landman, 2004; Nouwen, 2010), perhaps in the plural noun itself (Krifka, 1989). This approach can also be applied to measure phrases corresponding to mass dimensions, such as *two liters (of wine)* and *five ounces (of gold)*. Here it is most common to take the measure noun to encode a measure function (e.g. Krifka, 1989). *Ounce* on this view is based on a function that maps a portion of matter to a number that reflects its weight in ounces; *five ounces of gold* thus comes to denote a predicate true of portions of gold weighing 5 ounces. Alternately, mass nouns might also be analyzed as including a degree argument (Cresswell, 1977), with measure phrases themselves simply denoting degrees on the relevant scale.

Verbal semantics. A further extension of the degree-based approach is suggested by examples such as the following, based on Doetjes (1997):

(12) a. Joe read <u>a lot</u> of books.

b. Anne went to the movies <u>a lot</u> last year.

c. I slept a lot.

d. Joe appreciates Lisa a lot.

Cross-linguistically, it is common for quantity expressions to also occur as intensifiers in the verbal domain. Doetjes (1997) – who offers one of the most comprehensive discussions of verbal gradability of this sort – accounts for this pattern by proposing that verb phrases can contain two types of scalar argument positions, which may be saturated by a degree expression such as *a lot*. Stage level VPs – like nouns – contain a *q*-position (for 'quantity'), which is associated with the event argument, and is responsible for the frequency reading in examples like (12b), and the duration reading in cases like (12c). Conversely, individual level VPs as well as some stage level ones – like gradable adjectives – contain a *g*-position (for 'grade'), which is responsible for the intensity reading in examples like (12d). As evidence for the relevance of the *g*- versus *q*-position distinction, some degree expressions are restricted to one or the other type of scale. For example, *a lot* in (12a-c) would be translated by Dutch *veel* 'much', while in (12d) the translation would be *erg* 'badly', which also occurs with gradable adjectives.

The close relationship between quantity in the nominal domain and verbal gradability of the sort in (12b,c) is explored further by Nakanishi (2004a,b, 2007); Wellwood et al. (2012); Wellwood (2012), who propose analyses based on the idea that measure functions can apply both to individuals and to events, in the latter case measuring either the event directly or an entity related to the event (cf. Krifka 1989). The parallel between adjectival gradability and verbal examples of the sort in (12d) has however received less attention from this perspective.

The preceding discussion already shows that verbal gradability is a multifaceted phenomenon. But examples such as (12b-d) do not exhaust the complexity observed in the verbal domain. Another case involves measure verbs such as *weigh*, whose meaning presumably must be stated in terms of a measure function:

(13) The suitcase weighs 18 pounds / costs $60 / measures 26 inches in length.

Added to this are a wide and varied class of verbs that can be understood as expressing gradual change along some scalar dimension; per Hay et al. (1999), these include degree achievement verbs such as *widen*, *cool* and *lengthen* (14a), directed motion verbs such as *ascend* and *descend* (14b), and incremental theme verbs such as *eat* (14c):

(14) a. The lake cooled 4 degrees.

b. The plane descended 1000 feet.

c. John ate the apple/apples.

There is now a sizable body of work that extends degree-based semantic frameworks to these phenomena as well; this has proven fruitful in particular in developing a unified account of telicity effects found with all of these classes (see Section 4.1). As one example, Kennedy and Levin (2008) propose that degree achievement verbs lexicalize a special sort of derived measure function, namely a measure of change function. *Cool*, for example, might be given the entry in (15), where $cool_\Delta$ is a function that takes an object x and an event e and returns a degree representing the amount to which x changes in coolness as a result of participating in e (observe again the parallel to the adjectival entry (4)).

(15) $[\![cool]\!] = \lambda d \lambda x \lambda e.cool_{\Delta}(x)(e) \succeq d$

Kennedy and Levin (2008) note that this is readily extended to directed motion verbs as in (14b). Incremental theme verbs as in (14c) present more of a challenge, particularly in capturing the semantic contribution of the object NP. There is a long tradition of analyses of such verb phrases that invoke some sort of notion of scalarity (Krifka, 1989; Tenny, 1994; Filip, 1999), and in recent work they too have been analyzed explicitly in terms of measure functions from events and/or individuals to degrees. Here a topic of debate is where this measure function is lexicalized, with some authors (e.g. Caudal and Nicolas, 2005; Piñón, 2008) taking it to be part of the content of the verb itself, but others (e.g. Rappaport Hovav, 2008; Kennedy, 2012) arguing that it originates elsewhere. For other recent scale-based work on verbal semantics, see also Filip (2005, 2008); Gawron (2005); Filip and Rothstein (2006); Rothstein (2008).

As this discussion shows, scalarity in the verbal domain is an active topic of research in semantics. What is yet to emerge is an account that unifies all of the various verbal phenomena surveyed here.

Nouns. Nouns, too, can seemingly be gradable. We have already seen one sort of nominal scalarity, in the form of a quantity argument that is part of the semantics of the noun itself or some projection of it. Another sort is exemplified below:

(16) enormous idiot; utter disaster; real hero

(17) a. Clyde is more of an idiot than Floyd.
b. Clyde is more phonologist than phonetician.

As discussed by Morzycki (2009, 2012), to call someone an enormous idiot is not to say something about his physical size, but rather his degree of idiocy. Similarly, *utter disaster* conveys the degree to which something qualifies as a disaster. Nouns also allow comparison, as in (17).

While nominal gradability is undoubtedly real, we again face the question of how it should be analyzed. One possibility is that nouns such as *idiot* and *disaster* are like gradable adjectives in having a degree argument (per Morzycki 2009). But as discussed by Sassoon (2007, 2011); Constantinescu (2011); Morzycki (2012), gradability in the nominal domain patterns quite differently from the adjectival variety. In particular, in contrast to the well-delineated class of adjectival degree morphemes (*too/so/very*/etc.), there does not seem to be a coherent class of adnominal degree modifiers. Furthermore, examples of the sort in (16) and (17) plausibly represent distinct varieties of gradability: *enormous idiot/more of an idiot* appear to invoke a scale similar to that underlying the adjective *idiotic*; *real hero* seems instead to convey prototypicality or epistemic certainty; and the comparative in (17b) could be metalinguistic. Thus a unified analysis of nominal gradability looks unlikely, and better empirical coverage may be achieved by an approach that is only partially degree-based

(Morzycki, 2012) or that analyzes nouns as lacking degree structure entirely (Constantinescu, 2011).

Modal expressions. An area where degree-based frameworks have been applied more recently is in the analysis of modality. Modal expressions of various sorts are gradable, as evidenced by the fact that they form comparatives and compose with degree modifiers and measure phrases:

(18) a. I need to go on vacation more than I need to finish this work.
b. It is very likely that we've missed our train.
c. We are completely unable to fulfill your request.
d. It is 95% certain that our team will win.

Examples such as these are challenging for standard analyses of modality (e.g. Kratzer, 1981). To account for them, Lassiter (2011), building on Yalcin (2010), develops a comprehensive theory in which modal expressions of all sorts – even those that are not overtly gradable – are analyzed as functions that map propositions to points on a scale and compare them to a threshold value. In the case of epistemic modality, the relevant scale is a probability scale; for deontic and bouletic modality, it is a scale derived from a preference relation on propositions.

This discussion has not exhausted all sorts of natural language scalarity (we have not for example considered prepositions, e.g. *10 meters above*, and a discussion of Horn scales is postponed to Section 5), but this should be sufficient to illustrate that scales play a pervasive role throughout language.

3.2 Do all languages make use of scales?

For purposes of exposition, the linguistic data cited above were drawn from English. But comparable examples could be found in any number of other languages. This raises the question of whether all natural languages contain expressions that are scalar in the sense assumed here, i.e. whose meanings must be stated in terms of degrees on a scale. The answer seems to be no. While all languages apparently can express gradable concepts, not all do so via explicit reference to degrees.

In a questionnaire-based study of comparison constructions in 14 languages, Beck et al. (2009) argue for the following as a parameter of semantic variation between languages:

(19) **Degree Semantics Parameter (DSP):**
A language {does/does not} have gradable predicates (type $\langle d, \langle e, t\rangle\rangle$ and related), i.e. lexical items that introduce degree arguments.

They propose a set of tests to determine whether a language has the positive setting for this parameter, including the existence of expressions that manipulate degree arguments (e.g.

comparative/superlative morphology; words like *too* and *enough*), or that refer to degrees (as in *taller than 1.70 m* or *2 cm taller*).

Motu, a language of Papua New Guinea that forms comparatives via the so-called conjunctive strategy (20), elicits a negative answer to all of these questions, and is therefore concluded to have a negative setting for the DSP.

(20) Mary na lata, to Frank na kwadogi.
Mary TOP tall, but Frank TOP short
'Mary is taller than Frank.'

Thus in contrast to the lexical entry for English *tall* in (4), that for Motu *tall* does not introduce a degree argument. Rather, the authors propose that its semantics is essentially context dependent, such that (20) is true if there is a context in which Mary counts as tall while Frank counts as short.

Other languages argued to have a negative setting for this parameter are Fijian (Pearson, 2010) and Washo (Bochnak, 2103). Beyond this, it is well known that there are languages with no number words beyond 'one', 'two', 'three/a few' and 'many' (Dixon, 1980), suggesting that these lack even the scale of number.

If there are languages that express gradability completely without degrees, it is interesting to ask whether there might be traces of something similar even in languages like English that do have degrees as part of the ontology (cf. the above discussion of the delineation approach). This area is ripe for further research.

4 The structure of scales

4.1 The linguistic relevance of scale structure

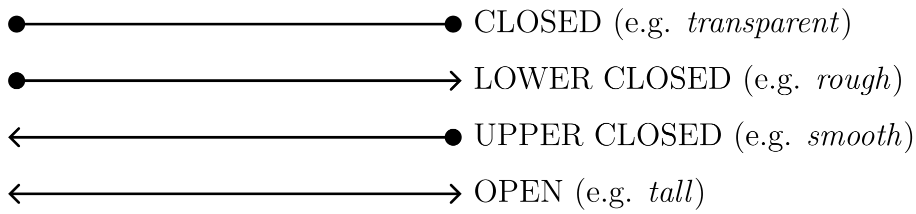
Having concluded that the meanings of certain expressions of natural language must be stated in terms of scales, it does not necessarily follow that scales themselves are a legitimate object of study in formal semantics. However, work in recent years has increasingly shown that the nature of scales does matter to the study of meaning, in that a range of linguistic phenomena lend themselves to explanation in terms of properties of scales and variation in their structures.

The best-known illustration of the linguistic relevance of scale structure relates to the presence or absence of scalar endpoints (see especially Rotstein and Winter, 2004; Kennedy and McNally, 2005; Kennedy, 2007). The motivating data come from the distribution of degree modifiers, which exhibit seemingly idiosyncratic patterns of acceptability with different sorts of gradable adjectives:

(21) a. perfectly transparent/smooth/*rough/*tall
b. slightly transparent/*smooth/rough/*tall

As the above authors show, these facts can be accounted for in terms of the structure of the scale lexicalized by the adjective, specifically whether it has minimum and/or maximum points. On this basis, a typology of four scale types can be derived: totally closed (both maximum and minimum), lower closed (minimum only), upper closed (maximum only) and open (neither).

(22)



The lexical semantics of *perfectly* requires the adjective have a scale with a maximum (to be *perfectly smooth* is to have the maximum value on the scale of smoothness). Conversely, *slightly* requires a scale with a minimum (to be *slightly rough* is to have a degree of roughness just beyond the minimum value).

Importantly, the explanatory potential of the typology in (22) is not restricted to facts relating to degree modification. Kennedy (2007) in particular argues that the standard of comparison for gradable adjectives in their positive form is determined by the structure of the underlying scale. So-called absolute adjectives like *transparent*, *smooth* and *rough* – which lexicalize partially or totally closed scales – have endpoint-oriented standards: that for *smooth*, for example, is defined in terms of the maximum point on the scale of smoothness. Relative adjectives such as *tall*, whose scales are totally open, have instead contextually determined standards.

Differences of this sort in scale structure impact online sentence processing (Frazier et al., 2008), and sensitivity to them begins in childhood (Syrett et al., 2010). The open/closed distinction has also been productively applied to account for phenomena outside the adjectival domain, a case in point being verbal telicity. To take an example involving degree achievement verbs, (23a) favors a telic reading, as evidenced by the felicity of an *in*-adverbial; (23b) by contrast has only an atelic reading. This can be explained with reference to the corresponding scales: the process of emptying has an endpoint (complete emptiness), while the process of widening does not (Kennedy and Levin 2008; see also Kearns 2007 for an alternate scale-based account).

(23) a. The sink emptied (in an hour).
b. The gap between the boats widened (*in an hour).

The nature and role of scalar endpoints is an ongoing topic of research, and some important questions remain open. Recent work has shed new light on the status of modifiers like *slightly* as diagnostics for minima/maxima (Bogal-Allbritten, 2012; Sassoon, 2012; Solt, 2012), and on the relation of this aspect of scale structure to the standard invoked by the positive form (Kagan and Alexeyenko, 2011; Lassiter, 2011; Klecha, 2012; Bylinina, 2014). Alternative, non-scalar accounts of the relative versus absolute distinction have also been developed (McNally, 2011; Lassiter and Goodman, 2013). It is fair to say that the final words on this topic are yet to be written. But the more general conclusion remains, namely that through investigating the properties of scales and how these vary, we open up a productive path towards explanation of linguistic facts.

Perhaps surprisingly given the amount of work in this area, there is at present no fully general model of the structure of the scalar domain, nor is there a clear understanding of the complete range of scalar features that natural language is sensitive to. We saw in Section 2 that the typology of levels of measurement (ordinal/interval/ratio) does not seem to fully characterize the variation in the scales that underlie meaning in language. What is called for is a linguistically oriented perspective. In the next section I review recent research on the linguistically relevant properties of scales, with the goal of outlining a framework for a comprehensive model of scale structure.

4.2 Towards a general model

In Section 2, a scale S was defined as a triple consisting of a set of degrees D, an ordering relation $\succ$ on that set, and a dimension *DIM*. A model of scale structure might be organized around variation in each of these three components.

4.2.1 Variation in D

We have already seen that the presence or absence of scalar **endpoints** – that is, of minimum and/or maximum elements of D – is a meaningful parameter of variation in scale structure.

An even more basic aspect of D that must be incorporated into a general model involves the **nature of degrees** themselves. Degrees are most commonly assumed to be points on a scale, but various authors have argued they should instead be conceptualized as intervals or extents (e.g. Seuren, 1978; Bierwisch, 1989). As an example, Kennedy (2001) develops an interval-based system that distinguishes between positive degrees (intervals extending from the scalar origin to some midpoint) and negative degrees (intervals extending from a midpoint to infinity). This distinction allows 'cross-polar anomalies' such as (24) to be ruled out as sortal mismatches, in that the two kinds of degrees cannot be compared.

(24) *Anna is taller than Zoe is short.

Schwarzschild and Wilkinson (2002) propose a more radical interval-based theory to account for examples involving quantifiers in comparatives, where they argue that point-based theories yield incorrect truth conditions. Other approaches have treated degrees as vectors (Zwarts and Winter, 2000; Winter, 2001), directed scalar segments (Schwarzschild, 2012), or other sorts of complex entities (Grosu and Landman, 1998), or have replaced degrees with tropes (Moltmann, 2009).

It is not fully clear that such more complex treatments of degrees are necessary. In particular, the phenomena used to justify interval-based semantics have in subsequent work been addressed within a more traditional point-based framework, generally by treating intervals as sets of degrees-as-points (see especially Büring, 2007; Beck, 2011). In what follows I thus assume the point-based view.

A further aspect of the structure of D that is argued to have linguistic consequences is whether it is **dense or discrete**, where density means that for any two degrees a, b such that $a \succ b$, there is a third degree c such that $a \succ c \succ b$. Density is intuitively plausible for dimensions such as height, weight, temperature and duration, but Fox and Hackl (2006) make the surprising claim that this property holds even for the dimension of cardinality, such that the relevant scale is not the natural numbers but rather something like the rational or real numbers.

Fox and Hackl argue that assuming 'universal density of measurement' allows for a unified account of a variety of phenomena involving maximalization operations. One example has to do with scalar implicatures. Unmodified numerals on some accounts give rise to upper-bounding implicatures; for example, (25a) implicates that John does not have 4 children (i.e. that 3 is the maximum number of children he has). But (25b) lacks the corresponding implicature that he does not have more than 4 children. A similar absence of implicature is observed in (26), but here there is an intuitively clear reason: since the scale of weight is dense, there is no maximum degree d such that John weighs more than d. Crucially, this explanation can only be extended to (25b) if we assume that the scale of cardinalities is likewise dense rather than discrete.

(25) a. John has 3 children. $\rightarrow$ John does not have 4 children.

b. John has more than 3 children. $\nrightarrow$ John does not have more than 4 children.

(26) John weighs more than 120 pounds.

Similar reasoning can be applied to asymmetries with overt *only* and with *wh*-movement, and to aspects of the interpretation of negated numerical comparatives (for the latter, see Nouwen 2008). Other authors, however, have argued that these phenomena can be explained without relying on assumptions about density (e.g. Mayr, 2013). As such, it is not yet conclusively established whether density is a universal property of measurement scales, or whether dense versus discrete is a parameter on which scales vary.

Another meaningful point of variation is whether or not there exists a standard unit of measurement, such that the members of D can be associated with **numerical values**; degrees of height, for example, can, while degrees of beauty cannot. Bale (2008) observes that the existence of a system of numerical measures is crucial to the availability of direct interadjective comparisons such as *the table is longer than it is wide*. The previously-noted contrast between *2.14 times as tall* and **2.14 times as beautiful* suggests that this factor also impacts the expression of proportional comparison: only adjectives with corresponding numerical measures (e.g. *tall*) allow precise ratio comparisons. Surprisingly, there seems to be no established view as to the nature and source of numerical degrees (see Klein 1991; Bale 2008; Sassoon 2011 for discussion): are numerical values simply labels given to degrees (which have independent existence), or should degrees be equated with numbers? Clarifying this could shed light on how and why this aspect of scale structure impacts meaning.

A final important property of D is **granularity**, a notion introduced by Krifka (2007) to account for the approximate interpretation of round numbers. For example, (27a) allows or even favors an approximate reading ('about 100'), while (27b) is necessarily interpreted precisely, a pattern Krifka attributes to a difference in the granularity of the underlying measurement scales: (27b) is interpreted relative to the unit scale (the integers), while (27a) on its approximate interpretation involves a coarser-grained scale, e.g. one based on units of 10.

(27) a. There were 100 people at the meeting. APPROXIMATE
b. There were 99 people at the meeting. PRECISE

In the case of cardinality, coarser-grained scales are typically based on powers of ten. For other dimensions, available granularity levels are shaped by the corresponding measurement system; for example, distinct granularity levels for time measurement are based on units of minutes, hours, days and so forth.

The mechanism of granularity has also been applied to the semantics of approximators like *exactly* and *roughly* (Sauerland and Stateva, 2007), and of degree modifiers such as *slightly* and *completely* (Sassoon and Zevakhina, 2012).

To date, granularity has not been incorporated into a more general formal model of scales. There is first of all no consensus as to whether it is a semantic phenomenon (per Sauerland and Stateva, 2007) or a pragmatic one (Lasersohn, 1999; Lauer, 2012). That is, does (27a) have an interpretation on which it is true in the situation where 99 people attended, or is it literally false – though perhaps 'close enough' – in this case? If semantic, there are various possibilities for how different granularity levels might be modeled, e.g. via entirely distinct scales (Krifka, 2007), or via granularity functions that map scale points to intervals around them of varying widths (Sauerland and Stateva, 2007). Also to investigate is what further phenomena can be explained in terms of granularity. As an example (discussed by Kennedy and McNally, 2005, among others), if I utter *The theater is empty tonight* to describe a situation where there are unexpectedly few people present, does this involve a very coarse-grained interpretation of the scalar endpoint, or a non-endpoint-oriented standard?

4.2.2 Variation in $\succ$

The most obvious parameter of variation in the ordering relation $\succ$ is **direction**. Per Kennedy and McNally (2005) and others, the scales lexicalized by antonym pairs such as *tall/short* and *smooth/rough* can be analyzed as sharing a dimension *DIM* and a set of degrees D, but differing in the direction of the relation $\succ$ on that set. This allows an account of equivalences such as (28): the degrees corresponding to Anna's and Zoe's 'tallness' are the same as those corresponding to their 'shortness', but the ordering between them is reversed.

(28) Anna is taller than Zoe. $\Leftrightarrow$ Zoe is shorter than Anna.

Another potential point of variation is in the properties of $\succ$ itself, involving what might be termed **ordering strength**. Most authors assume at least that the degrees on a scale are totally ordered, such that for any distinct degrees a and b, either $a \succ b$ or $b \succ a$. But we might imagine that this requirement could be relaxed, a possibility envisioned already by Cresswell (1977). Consider for example a scale where $\succ$ does not correspond to the mathematical operation $>$, but rather $a \succ b$ iff $a > b + \epsilon$ for some fixed value ϵ. The result is a so-called semiorder (Luce, 1956; van Rooij, 2011b), an ordering where $\succ$ is transitive but indiscriminability is not, meaning that there could be three degrees a, b and c where $\neg(a \succ b)$ and $\neg(b \succ c)$ but $a \succ c$.

A scale with this sort of structure plausibly captures an important aspect of how speakers actually make comparisons, namely tolerance. Across a wide range of perceptual properties, our ability to differentiate two stimuli – for example, to determine which of two weights is heavier – is subject to ratio-dependent thresholds of discriminability (Stevens, 1975). Something similar is observed in perception of approximate number (Dehaene, 1997), and in preference between options (Luce, 1956). Furthermore, certain linguistic expressions of comparison exhibit a parallel sort of tolerance, an example being implicit comparatives such as *Anna is tall compared to Zoe*, which is infelicitous if the difference between the two heights is small (Kennedy, 2007). Fults (2011) analyzes such examples via a semi-ordered scale based on psychological models of humans' approximate numerical abilities, and Solt (2014a) argues that something similar underlies aspects of the interpretation of the quantifier *most* (for a non-scalar account based on tolerance, see also Cobreros et al. 2012). Whether such scale structures might play a broader role requires further investigation.

4.2.3 Variation in *DIM*

At the most obvious level, the dimension *DIM* is what separates the scale of, say, height from those of length, weight or beauty. A more profound question is whether there are meaningful subclasses of dimensions to which the grammar is sensitive.

Krifka (1989) calls attention to one important such subclass, those dimensions whose associated measure functions are **extensive** or additive with respect to concatenation (cf. the discussion in Section 2). Schwarzschild (2006) introduces a related but slightly weaker notion of **monotonicity**, where a function is monotonic on the part/whole relation between entities if the measure of any proper subpart of an individual is less than that of the individual as a whole. Volume, for example is extensive and thus also monotonic (the volume of two cups of water put together is the sum of their individual volumes); temperature, by contrast, is neither extensive nor even monotonic (the temperature of a part of a cup of water is not necessarily lower than that of the entire cup). This classification has grammatical consequences: only measure phrases corresponding to extensive/monotonic dimensions occur in partitives, while only those that are non-monotonic may be used attributively:

(29) a. 3 gallons/*20 degrees of water
b. *3 gallon/20 degree water

The distinction here is also intuitively related to that between Doetjes' dimensions of **quantity and grade**, which likewise has linguistic relevance, determining for example the distribution of the Dutch degree modifiers *erg* and *veel* (cf. Section 3.1).

It seems that we have here the basis for a meaningful classification of dimensions; however, assigning a given dimension to one or the other class is not completely straightforward. Schwarzschild notes that monotonicity depends on what the context determines to be the relevant part/whole relationship. In talking about falling snow, for example, the salient notion of parthood is based on layers and sublayers; depth is monotonic on this, allowing examples such as *2 feet of snow*. In the context of snow removal, on the other hand, we might be more concerned with three-dimensional volumes; relative to this way of partitioning, volume is monotonic, allowing *2 million cubic feet of snow*. Furthermore, Doetjes observes that what is conceptually the same dimension can be construed as either a quantity (e.g. *much luck*) or a grade (*very lucky*). We might also ask if the crucial property is additivity or monotonicity, and if there might also be other types of linguistically relevant relationships between measurement and parthood and/or concatenation. As one possibility, Lassiter (2011) argues for the existence of dimensions that are 'intermediate' with respect to concatenation. Temperature seems to be an example: the temperature of two bowls of soup poured together is intermediate between the temperatures of the individual bowls. Which dimensions share this property, and whether they form a coherent class from a linguistic perspective, are open questions.

A second intuitive but difficult-to-define distinction is that between **evaluative and non-evaluative** dimensions. Adjectives corresponding to the two classes again pattern differently. For example, in contrast to the equivalence illustrated in (28) for the non-evaluative pair *tall/short*, that in (30) for the evaluative *smart/dumb* does not seem to hold, in that the second sentence is infelicitous if both Anna and Zoe are judged smart.

(30) Anna is smarter than Zoe. $\not\Leftrightarrow$ Zoe is dumber than Anna.

There have been various attempts to capture the distinction between evaluative and non-evaluative adjectives, and in particular the inherent 'norm-relatedness' of evaluative comparatives like *dumber* in (30) (e.g. Bierwisch, 1989; Rett, 2008). One scale-based approach (inspired by Cruse, 1986; Bierwisch, 1989) would be to say that the scales underlying evaluative adjective pairs differ not only in the ordering relation $\succ$ but also in that their sets of degrees D are not fully coextensive. There may also be more fundamental differences between the two sorts of scales, corresponding more directly to the difference between objective measurement and subjective, speaker-dependent judgment (for relevant discussion see Kennedy 2013).

4.3 Summary

The central conclusion from the preceding discussion is that natural language is sensitive to a wide range of scalar properties. The investigation of scale structure has thus proven to be a fruitful direction in research in semantics. The current state of our knowledge is perhaps not yet such that the various results in this area can be brought together into a fully general model of the scalar domain, but the above overview gives an indication of the components and parameters of variation that such a model must include. I hope this brief overview has shown that there is benefit to be had in pursuing this line of research further.

5 Scales in pragmatics

The scales we have been considering up to this point could be called measurement scales. A different sort of scale plays a role in pragmatic inferencing. By way of illustration, (31a) entails that Anna read at least some of the relevant books, but carries the implication that she did not read all of them. A leading account of such examples treats them as scalar implicatures, calculated with reference to a scale of alternatives ordered by strength – a Horn scale (named after Horn, 1972) – as depicted in (31b). From the speaker's choice to use the weaker scalar term *some*, it can be inferred that he/she is not in a position to use the stronger alternative *all* (Grice, 1975; Horn, 1989; Levinson, 2000).

(31) a. Anna read some of the books on the reading list.

b. $\langle \textit{some}, \textit{all} \rangle$

It is beyond the scope of this paper to discuss theories of scalar implicature in general; see Sauerland (2012) for a recent overview of this topic. Here I will confine myself to some brief remarks on the nature of the scales that they are based on, and their relationship to the scales that have been the focus of this paper

The first observation to be made is that Horn scales are different sorts of entities from measurement scales: the former are composed of linguistic expressions (Levinson, 2000), while the latter are part of the domain that provides such expressions with their content. A wide variety of lexical items form Horn scales, including quantifiers (as above), logical connectives (e.g. $\langle \textit{or}, \textit{and} \rangle$), gradable adjectives ($\langle pretty, beautiful \rangle$, $\langle warm, hot \rangle$), modal expressions ($\langle might, must \rangle$), and verbs ($\langle try, succeed \rangle$). Classic examples of Horn scales are based on entailment (the stronger term unidirectionally entails the weaker one), but Hirschberg (1985) points out that other sorts of ordered pairs/sets also do so, including rank orders, spatial orderings and process stages. For example, in some contexts *I typed the letter* would implicate that I didn't yet mail it. But not all ordered sets give rise to implicatures (e.g. *if* does not implicate 'not if and only if'), and there have been a number of attempts to formulate conditions on well-formed Horn scales, e.g. the terms must be equally lexicalized, from the same semantic field, or salient in the discourse (for discussion, see the above-cited works as well as Gazdar, 1979; Atlas and Levinson, 1981; Matsumoto, 1995).

Yet despite the fundamental difference between these two varieties of scalarity, there are nonetheless connections between them. In the case of adjectival Horn scales, the associated measurement scale provides for their ordering (e.g. *hot* corresponds to a higher range on the heat scale than *warm*). The connection is even closer in the case of number words, where it is typically assumed that the scale that provides them with their semantic content is the same as that which serves as the basis for implicature calculation (cf. the above discussion of Fox and Hackl 2006). Recent experimental work has highlighted further interactions between the scales underlying gradability and the availability of implicatures. Doran et al. (2009) demonstrate that different types of Horn scales differ in the frequency with which implicatures arise: higher levels are observed for quantificational scales and cardinal numerals, and lower levels for rank orderings and adjectival scales. Within the adjectival group, the structure of the underlying measurement scale plays a role. van Tiel et al. (2013) show that implicatures are more likely to be calculated relative to scalar endpoints (e.g. from *cheap* to *not free*) than non-endpoints (e.g. *big* to *not enormous*); scalar distance between Horn scale members also has an effect (see also Beltrama and Xiang, 2012). From a different perspective, Cummins et al. (2012) demonstrate the existence of implicatures from modified numerals such as *more than 100* that are dependent on scale granularity.

These observations suggest that a full account of the scales involved in natural language meaning must also address how they contribute to pragmatic inferencing.

Acknowledgements

I would like to thank Carla Umbach and Hazel Pearson for thoughtful feedback on earlier versions of this review. Thanks also to two anonymous Language and Linguistics Compass reviewers and to editor Maribel Romero, whose helpful comments have made for a much improved paper. This work was supported by the Deutsche Forschungsgemeinschaft under grant 1157/1-1.

References

Atlas, Jay David, and Stephen C. Levinson. 1981. *It*-clefts, informativeness, and logical form: radical pragmatics (revised standard version). In *Radical pragmatics*, ed. Peter Cole, 1–61. New York: Academic Press.

Bale, Alan Clinton. 2008. A universal scale of comparison. *Linguistics and Philosophy* 31:1–55.

Bale, Alan Clinton. 2011. Scales and comparison classes. *Natural Language Semantics* 19:169–190.

Bartsch, Renate, and Theo Vennemann. 1973. *Semantic structures: A study in the relation between syntax and semantics*. Frankfurt: Athenäum Verlag.

Beck, Sigrid. 2011. Comparison constructions. In *Semantics: An international handbook of natural language meaning, Vol. 2*, ed. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 1341–1389. Berlin: De Gruyter Mouton.

Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst, and Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9:1–66.

Beltrama, Andrea, and Ming Xiang. 2012. Is *good* better than *excellent*? an experimental investigation on scalar implicatures and gradable adjectives. In *Proceedings of Sinn und Bedeutung 17*, ed. Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, 81–98.

Bierwisch, Manfred. 1989. The semantics of gradation. In *Dimensional adjectives*, ed. Manfred Bierwisch and Ewald Lang, 71–261. Berlin: Springer Verlag.

Bochnak, M. Ryan. 2103. The non-universal status of degrees: Evidence from Washo. In *NELS 42: Proceedings of the 42nd Annual Meeting of the North East Linguistic Society*, ed. Stefan Keine and Shayne Sloggett, 79–92. Amherst, MA: GLSA.

Bogal-Allbritten, Elizabeth. 2012. *Slightly* coerced: Processing evidence for adjectival coersion by minimizers. In *Proceedings of the 48th Meeting of the Chicago Linguistic Society*.

Bolinger, Dwight. 1972. *Degree words*. The Hague: Mouton.

Bresnan, Joan W. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4:275–343.

Büring, Daniel. 2007. Cross-polar nomalies. In *Proceedings of the 17th Semantics and Linguistic Theory Conference*, 37–52. Ithaca, NY: CLC Publications.

Burnett, Heather Susan. 2012. The grammar of tolerance: On vagueness, context-sensitivity and the origin of scale structure. Doctoral Dissertation, University of California Los Angeles.

Bylinina, Lisa. 2014. The grammar of standards: Judge-dependence, purpose-relativity, and comparison classes in degree constructions. Doctoral Dissertation, Utrecht University.

Caudal, Patrick, and David Nicolas. 2005. Types of degrees and types of event structures. In *Event arguments: Foundations and application*, ed. Claudia Maienborn and Angelika Wöllstein, 277–300. Tübingen: Max Niemeyer Verlag.

Cobreros, Pablo, Paul Égré, David Ripley, and Robert van Rooij. 2012. Tolerant, classical, strict. *Journal of Philosophical Logic* 41:347–385.

Constantinescu, Camelia. 2011. Gradability in the nominal domain. Doctoral Dissertation, Leiden University.

Cresswell, Max J. 1977. The semantics of degree. In *Montague grammar*, ed. Barbara H. Partee, 261–292. New York: Academic Press.

Cruse, D.A. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Cummins, Chris, Uli Sauerland, and Stephanie Solt. 2012. Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy* 35:135–169.

Dehaene, Stanislas. 1997. *The number sense: How the mind creates mathematics*. Oxford: Oxford University Press.

Dixon, Robert M. W. 1980. *The languages of Australia*. Cambridge University Press.

Doetjes, Jenny Sandra. 1997. Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English. Doctoral Dissertation, Leiden University.

Doran, Ryan, Rachel E. Baker, Yaron McNabb, Meredith Larson, and Gregory Ward. 2009. On the non-unified nature of scalar implicature: an empirical investigation. *International Review of Pragmatics* 1:211–248.

Filip, Hana. 1999. *Aspect, eventuality types and nominal reference*. New York and London: Garland.

Filip, Hana. 2005. Measures and indefinites. In *Reference and quantification: the Partee effect*, ed. Gregory N. Carlson and Francis Jeffrey Pelletier, 229–288. Stanford: CSLI Publications.

Filip, Hana. 2008. Events and maximalization: The case of telicity and perfectivity. In *Theoretical and crosslinguistic approaches to the semantics of aspect*, ed. Susan Rothstein, 217–256. Amsterdam: John Benjamins.

Filip, Hana, and Susan Rothstein. 2006. Telicity as a semantic parameter. In *Formal Approaches to Slavic Linguistics 14 (The Princeton Meeting)*, ed. James Lavine, Steven L. Franks, Mila Tasseva-Kurktchieva, and Hana Filip, 139–156. Ann Arbor: Michigan Slavic Publications.

Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.

Frazier, Lyn, Charles Clifton Jr., and Britta Stolterfoht. 2008. Scale structure: processing minimum standard and maximum standard scalar adjectives. *Cognition* 106:299–324.

Fults, Scott. 2011. Vagueness and scales. In *Vagueness and language use*, ed. Paul Égré and Nathan Klinedinst, 25–50. New York: Palgrave Macmillan.

Gawron, Jean Mark. 2005. Generalized paths. In *15th Semantics and Linguistic Theory Conference*. University of California, Los Angeles.

Gazdar, Gerald. 1979. *Pragmatics: implicature, presupposition, and logical form*. New York: Academic Press.

Grice, H. Paul. 1975. Logic and conversation. In *Syntax and semantics 3: Speech acts*, ed. Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.

Grosu, Alexander, and Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6:125–170.

Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, Massachusetts Institute of Technology.

Hackl, Martin. 2009. On the grammar and processing of proportional quantifiers: *Most* versus *more than half*. *Natural Language Semantics* 17:63–98.

Hay, Jennifer, Christopher Kennedy, and Beth Levin. 1999. Scalar structure underlies telicity in "degree achievements". In *Proceedings of the 9th Semantics and Linguistic Theory Conference*, ed. Tanya Matthews and Devon Strolovitch, 127–144. Ithaca, NY: CLC Publications.

Heim, Irene. 2000. Degree operators and scope. In *Proceedings of the 10th Semantics and Linguistic Theory Conference*, ed. Brendan Jackson and Tanya Matthews, 40–64. Ithaca, NY: CLC Publications.

Heim, Irene. 2006. *Little*. In *Proceedings of the 16th Semantics and Linguistic Theory Conference*, ed. Masayuki Gibson and Jonathan Howell, 35–58. Ithaca, NY: CLC Publications.

Hirschberg, Julia. 1985. A theory of scalar implicature. Doctoral Dissertation, University of Pennsylvania.

Horn, Laurence R. 1972. On the semantic properties of logical operators in English. Doctoral Dissertation, University of California, Los Angeles.

Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Kagan, Olga, and Sascha Alexeyenko. 2011. Degree modification in Russian morphology: the case of the suffix *ovat*. In *Proceedings of Sinn and Bedeutung 15*, ed. Ingo Reich, Eva Horch, and Dennis Pauly, 321–335. Saarbrücken: Saarland University Press.

Kearns, Kate. 2007. Telic senses of deadjectival verbs. *Lingua* 117:26–66.

Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24:33–70.

Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30:1–45.

Kennedy, Christopher. 2012. The composition of incremental change. In *Telicity, change, and state: a cross-categorical view of event structure*, ed. Violeta Demonte and Louise McNally, 103–121. Oxford: Oxford University Press.

Kennedy, Christopher. 2013. Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry: An Interdisciplinary Journal of Philosophy* 56:258–277.

Kennedy, Christopher, and Beth Levin. 2008. Measure of change: The adjectival core of degree achievements. In *Adjectives and adverbs: Syntax, semantics, and discourse*, ed. Louise McNally and Christopher Kennedy, 156–182. Oxford: Oxford University Press.

Kennedy, Christopher, and Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language* 81:345–381.

Klecha, Peter. 2012. Positive and conditional semantics for gradable modals. In *Proceedings of Sinn und Bedeutung 16. MIT Working Papers in Linguistics*, ed. Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, volume 2, 363–376.

Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4:1–45.

Klein, Ewan. 1982. The interpretation of adjectival comparatives. *Journal of Linguistics* 18:113–136.

Klein, Ewan. 1991. Comparatives. In *Semantics: An international handbook of contemporary research*, ed. Arnim von Stechow and Dieter Wunderlich, 673–691. Berlin: de Gruyter.

Kranz, David H, R. Duncan Luce, Patrick Suppes, and Amos Tversky. 1971. *Additive and polynomial representations*, volume I of *Foundations of Measurement*. New York: Academic Press.

Kratzer, Angelika. 1981. The notional category of modality. In *Words, worlds, and contexts: new approaches in word semantics*, ed. Hans-Jürgen Eikmeyer and Hannes Rieser, 38–74. Berlin: de Gruyter.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and contextual expressions*, ed. Renate Bartsch, Johan van Benthem, and Peter von Emde Boas, 75–115. Dordrecht: Foris.

Krifka, Manfred. 1999. At least some determiners aren't determiners. In *The semantics/pragmatics interface from different points of view*, ed. Ken Turner, 257–291. Oxford: Elsevier.

Krifka, Manfred. 2007. Approximate interpretations of number words: A case for strategic communication. In *Cognitive foundations of interpretation*, ed. Gerlof Bouma, Irena Krämer, and Joost Zwarts, 111–126. Koninklijke Nederlandse Akademie van Wetenschappen.

Landman, Fred. 2004. *Indefinites and the type of sets*. London: Blackwell.

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75:522–551.

Lassiter, Dan. 2011. Measurement and modality: the scalar basis of modal semantics. Doctoral Dissertation, New York University.

Lassiter, Dan, and Noah D. Goodman. 2013. Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Proceedings of the 23rd Semantics and Linguistic Theory Conference*.

Lauer, Sven. 2012. On the pragmatics of pragmatic slack. In *Proceedings of Sinn und Bedeutung 16. MIT working papers in linguistics*, ed. Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, volume 2, 389–402.

Levinson, Stephen C. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Luce, R. Duncan. 1956. Semiorders and a theory of utility discrimination. *Econometrica* 24:178–191.

Matsumoto, Yo. 1995. The conversational condition on Horn scales. *Linguistics and Philosophy* 18:21–60.

Mayr, Clemens. 2013. Implicatures of modified numerals. In *From grammar to meaning: The spontaneous logicality of language*, ed. Ivano Caponigro and Carlo Cecchetto, 139–171. Cambridge: Cambridge University Press.

McNally, Louise. 2011. The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In *Vagueness in Communication (ViC2009), Revised Selected Papers (LNAI 6517)*, ed. Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, 151–168. Berlin, Heidelberg: Springer.

Moltmann, Friederike. 2009. Degree structure as trope structure: a trope-based analysis of positive and comparative adjectives. *Linguistics and Philosophy* 32:51–94.

Morzycki, Marcin. 2009. Degree modification of gradable nouns: size adjectives and adnominal degree morphemes. *Natural Language Semantics* 17:175–203.

Morzycki, Marcin. 2012. The several faces of adnominal degree modification. In *Proceedings of the 29th West Coast Conference on Formal Linguistics*, ed. Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, and Alex Trueman, 187–195. Somerville, MA: Cascadilla.

Nakanishi, Kimiko. 2004a. Domains of measurement: Formal properties of non-split/split quantifier constructions. Doctoral Dissertation, University of Pennsylvania.

Nakanishi, Kimiko. 2004b. On comparative quantification in the verbal domain. In *Proceedings of the 14th Semantics and Linguistics Theory Conference*, ed. Robert B. Young, 179–196. Ithaca, NY: CLC Publications.

Nakanishi, Kimiko. 2007. Measurement in the nominal and verbal domains. *Linguistics and Philosophy* 30:235–276.

Nouwen, Rick. 2008. Upper-bounded *no more*: the exhaustive interpretationof non-strict comparison. *Natural Language Semantics* 16:271–295.

Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3:1–41.

Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof, 115–143. Dordrecht: Foris.

Pearson, Hazel. 2010. How to do comparison in a language without degrees: a semantics for the comparative in Fijian. In *Proceedings of Sinn und Bedeutung 14*, ed. Martin Prinzhorn, Viola Schmitt, and Sarah Zobel, 356–372.

Piñón, Christopher. 2008. Aspectual composition with degrees. In *Adjectives and adverbs: Syntax, semantics, and discourse*, ed. Louise McNally and Christopher Kennedy, 183–219. Oxford: Oxford University Press.

Rappaport Hovav, Malka. 2008. Lexicalized meaning and the internal temporal structure of events. In *Theoretical and crosslinguistic approaches to the semantics of aspect*, ed. Susan D. Rothstein, 13–42. Amsterdam/Philadelphia: John Benjamins.

Rett, Jessica. 2008. Degree modification in natural language. Doctoral Dissertation, Rutgers University.

Roberts, Fred S. 1985. *Measurement theory with applications to decisionmaking, utility, and the social sciences*. Cambridge: Cambridge University Press.

van Rooij, Robert. 2011a. Measurement, and interadjective comparisons. *Journal of Semantics* 28:335–358.

van Rooij, Robert. 2011b. Vagueness and linguistics. In *Vagueness: a guide*, ed. Giuseppina Ronzitti, 123–170. Dordrecht, Heidelberg, London, New York: Springer.

Rothstein, Susan D. 2008. Telicity, atomicity and the Vendler classification of verbs. In *Theoretical and crosslinguistic approaches to the semantics of aspect*, ed. Susan D. Rothstein, 43–78. Amsterdam: John Benjamins.

Rotstein, Carmen, and Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12:259–288.

Sapir, Edward. 1944. Grading: a study in semantics. *Philosophy of Science* 11:93–116.

Sassoon, Galit. 2010. Measurement theory in linguistics. *Synthese* 174:151–180.

Sassoon, Galit. 2011. Adjectival vs. nominal categorization processes: The rule vs. similarity hypothesis. *Belgium Journal of Linguistics* 25:104–147.

Sassoon, Galit. 2012. A *slightly* modified economy principle: stable properties have non-stable standards. In *Proceedings of the Israel Association of Theoretical Linguistics (IATL) 27. MIT Working Papers in Linguistics*.

Sassoon, Galit, and Natalia Zevakhina. 2012. Granularity shifting: Experimental evidence from degree modifiers. In *Proceedings of the 22nd Semantics and Linguistic Theory Conference*, ed. Anca Chereches, 226–246. Ithaca, NY. CLC Publications.

Sassoon, Galit Weidmann. 2007. Vagueness, gradability and typicality, a comprehensive semantic analysis. Doctoral Dissertation, Tel Aviv University.

Sauerland, Uli. 2012. The computation of scalar implicatures: pragmatic, lexical or grammatical. *Language and Linguistics Compass* 6:36–49.

Sauerland, Uli, and Penka Stateva. 2007. Scalar vs. epistemic vagueness: evidence from approximators. In *Proceedings of the 17th Semantics and Linguistic Theory Conference*, ed. Masayuki Gibson and Tova Friedman, 228–245. Ithaca, NY: CLC Publications.

Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9:67–110.

Schwarzschild, Roger. 2012. Directed scale segments. In *Proceedings of the 22nd Semantics and Linguistic Theory Conference*, ed. Anca Chereches, 65–82. Ithaca, NY: CLC Publications.

Schwarzschild, Roger, and Karina Wilkinson. 2002. Quantifiers in comparatives: A semantics of degree based on intervals. *Natural Language Semantics* 10:1–41.

Seuren, Pieter A.M. 1973. The comparative. In *Generative grammar in Europe*, ed. Ferenc Kiefer and Nicholas Ruwet, 528–564. Dordrecht: Riedel.

Seuren, Pieter A.M. 1978. The structure and selection of positive and negative gradable adjectives. In *14th Regional Meeting of the Chicago Linguistic Society*, 336–346.

Solt, Stephanie. 2011. Notes on the comparison class. In *Vagueness in Communication (ViC2009), Revised Selected Papers (LNAI 6517)*, ed. Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, 189–206. Berlin, Heidelberg: Springer.

Solt, Stephanie. 2012. Comparison to arbitrary standards. In *Proceedings of Sinn und Bedeutung 16. MIT Working Papers in Linguistics*, ed. Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, volume 2, 557–570.

Solt, Stephanie. 2014a. On measurement and quantification: the case of *most* and *more than half*. Accepted for publication in *Language*.

Solt, Stephanie. 2014b. Q-adjectives and the semantics of quantity. *Journal of Semantics* Published online February 3, 2014.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3:1–77.

von Stechow, Arnim. 2009. The temporal degree adjectives *früh(er)/spät(er)* 'early(er)'/'late(r)' and the semantics of the positive. In *Quantification, definiteness, and nominalization.*, ed. Anastasia Giannakidou and Monika Rathert, 214–233. Oxford: Oxford University Press.

Stevens, Stanley Smith. 1946. On the theory of scales of measurement. *Science* 103:677–680.

Stevens, Stanley Smith. 1975. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.

Syrett, Kristen, Christopher Kennedy, and Jeffrey Lidz. 2010. Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* 27:1–35.

Tenny, Carol L. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.

van Tiel, Bob, Emiel van Miltenburg, Natalia Zevakhina, and Bart Geurts. 2013. Scalar diversity. Submitted for publication.

Wellwood, Alexis. 2012. Back to basics: *more* is always *much-er*. In *Proceedings of Sinn und Bedeutung 17*, ed. Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, 599–618.

Wellwood, Alexis, Valentine Hacquard, and Roumyana Pancheva. 2012. Measuring and comparing individuals and events. *Journal of Semantics* 29:207–228.

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.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5:916–937.

Zwarts, Joost, and Yoad Winter. 2000. Vector space semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information* 9:169–211.