Scales in natural language

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

July 1, 2013

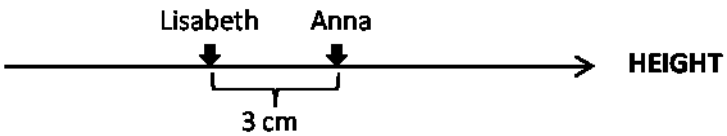
Abstract

The meaning of a wide variety of natural language expressions can be stated with respect to degrees on scales. This paper provides an overview of the contribution of scales to linguistic meaning: What are scales? Where do they come from? What is their role in language? Some initial steps are taken towards a general model of scale structure. The relation between scales of the semantic variety and those underlying pragmatic inferences – 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 Lisabeth, a situation depicted visually in (1b).

(1) a. Anna is 3 cm taller than Lisabeth.
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 & Vennemann (1973); Cresswell (1977); von Stechow (1984); Cruse (1986); Bierwisch (1989); Klein (1991), and this area (now typically called 'degree semantics') has seen an upsurge of interest in recent years. This work has produced a large number of important insights; but the current landscape is characterized largely by discrete topics pursued independently from one another. The goal of this paper is first of all to consolidate these various divergent threads of research into a unified 'state-of-our-knowledge' picture of

the nature and role of the scales that underlie meaning in language. As a part of this, I will outline the first steps towards a general model of scale structure, based on a comprehensive set of parameters of structural variation. In doing so, I will highlight some lesser understood aspects of scales and their structures, where further research has the potential to clarify as yet unexplained linguistic phenomena.

The organization of the paper is the following: Section 2 addresses the fundamental questions of what scales are and where they come from. Section 3 surveys the linguistic domains where scalarity has been found to play a role, including possible cross-linguistic variation in this. Section 4 returns to the question of scale structure, investigating several parameters on which scales vary in their structure, and the linguistic effects this gives rise to. Some preliminary observations towards a unified model of scale structure are offered. Finally, Section 5 briefly considers the relationship between the scales of the sort exemplified in (1) and the so-called Horn scales involved in the calculation of pragmatic inferences.

2 The ontology of scales

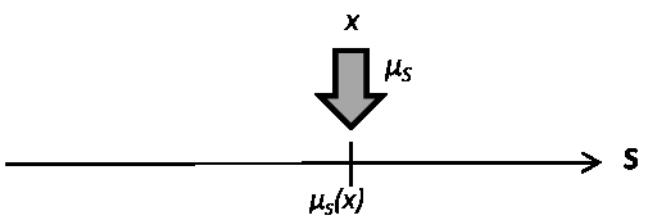
As a starting point, let us take the fairly common view (e.g. Kennedy 2007) that a scale S is a triple of the following form:

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

- D is a set of degrees,
- $\succ$ is an ordering relation on that set, 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 saturated by degree modifiers such as *too*, *very*, *that*, etc., or by comparative or superlative morphology. In the case of the unmodified 'positive' form of the adjective (as in *Anna is tall*), it is typical to take this role to be played by a null degree morpheme *pos*, which introduces a standard of comparison

(Cresswell, 1977; Kennedy, 2007; von Stechow, 2006; Heim, 2006). With the appropriate entry for the comparative morpheme *-er* (cf. Beck, 2011), (1) receives the interpretation in (5):

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

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

But the schema in (2) places only loose restrictions on what scales might look like, and leaves open a wide space 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 sorts 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 origin. 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, 2011; Lassiter, 2011). Specifically, we begin 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 between degrees is then derived from the relation between individuals as follows: for equivalence classes (i.e. degrees) $\overline{a}$ and $\overline{b}$ containing individuals a and b, respectively, $\overline{a} \succ \overline{b}$ iff $R(a,b)$.

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 (see Kranz et al. 1971; Roberts 1979 for basic introductions to measurement theory). The benefit of this is that it makes clear predictions about possible scale structures. If scales and degrees are taken to constitute a separate and purely abstract domain, then no *a priori* commitments are made as to the structure of that domain. But

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

on the measurement theoretic view, matters are more constrained. The scale derived via the process 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. But some linguistic phenomena require more than an ordinal scale. While degrees-as-equivalence-classes can be assigned numerical values in an order-preserving way, this cannot be done in such a way to make statements of the following form meaningful:

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

These require a stronger scale structure than ordinal. In measurement theoretic terms, (6a) requires an **interval** scale (featuring a fixed unit of measure), while (6b) requires a **ratio** scale (fixed unit of measure plus non-arbitrary zero point).

The derivation procedure outlined above can be augmented to produce a stronger (ratio level) scale, by introducing a concatenation operation on the domain of individuals and a corresponding mathematical operation on degrees. To continue with height as an example, 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 measure phrases (e.g. an object of exactly 1 foot in length).

The question is whether all of the scales relevant to natural language can be given a formal basis in this manner. The procedure described above extends readily to other physical dimensions such as width, depth, weight, etc., and perhaps to some abstract ones as well (e.g. cost). Furthermore, the mathematical operation corresponding to concatenation need not be addition. Temperature, for example, might represent an example of what Lassiter (2011) calls an "intermediate" dimension: the temperature of two bowls of soup poured together is intermediate between the temperatures of the two individual bowls. Dimensions such as beauty, intelligence and importance are however more problematic. Some authors assume their scales are only ordinal, and indeed it is difficult to imagine what the relevant concatenation operations might be. But examples of the following form are quite acceptable (and readily found in corpora):

(7) Anna is twice as beautiful/smart/important as Lisabeth.

This cannot simply be a matter of *twice as* having a loose interpretation on which it can combine with <u>any</u> gradable adjective. It is infelicitous with negative dimensional adjectives (??*twice as short/narrow/light*), which Sassoon (2010) attributes to these lexicalizing an interval-level scale, lacking a zero point. This suggests that the scales corresponding to the adjectives in (7) are ratio level, but it is unclear how such a scale structure is to be derived.

Note also the following contrast:

(8) 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 Lisabeth.

If the scale of height is ratio level, we are led to conclude that those underlying *beautiful* etc. have a weaker structure, apparently something between ordinal and ratio.

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 at least some 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 the question of scale structure in Section 4, where I take up the issue of variations in scale structure and their linguistic consequences, and take some initial steps towards a revised and more linguistically oriented typology of scale structures. 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) have underlying scalar semantics (in the sense of introducing a degree argument). Bierwisch (1989) assumes degree-based entries only for dimensional adjectives such as *tall*, while taking evaluative adjectives such as *smart* or *beautiful* at the lexical level to denote simple one-place predicates. Furthermore, while the meaning of an example such as (1a) must clearly be stated in terms of a scale of heights, the same is not necessarily the case for corresponding examples involving the bare positive form, or even the unmodified comparative (*Anna is taller than Lisabeth*). An alternate theoretical framework, the so-called delineation approach associated in particular with Klein (1980, 1982), takes gradable adjectives to denote context-dependent, partial one-place predicates, 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 (1a), 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, 2011a; 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 & Vennemann (1973) and Cresswell (1977):

(9) a. Arabella is more beautiful than Tom.
b. Chicago has more residents than Austin.
c. We bought more milk than wine.

If (9a) expresses the comparison of two degrees on the scale of beauty, it seems that (9b,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 scales. In examples like (9b), the scale is the number line, that is, the integers, or perhaps the rational or real numbers (see below), while in (9c), it is one corresponding to a mass dimension, here presumably 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)).

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

Here $|x|$ is a set cardinality operator, i.e. a counting measure function, which maps a plurality to a number that corresponds to its cardinality. This semantics allows a standard analysis of adjectival comparatives such as (1a) and (9a) to be extended directly to cases of quantificational *more* such as (9b), which is decomposed into *many* + *-er* and (9c), which can be analyzed as *much* + *-er* (cf. Bresnan 1973).

Number words (e.g. *three*) and modified numerals (*more than three*, *at least four*) are also readily analyzed in a scalar semantics. Compositionally this can be handled in a variety of ways: some authors (Partee, 1987; Krifka, 1999) posit a counting measure function as a part of the lexical meaning of the numeral itself, 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 extended to measure phrases corresponding to mass dimensions, such as *two liters (of wine)* and *five ounces (of gold)*. Here it is most common (e.g. Krifka, 1989) to take the measure noun to encode a measure function. *Ounce*, for example, 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 include a degree argument (Cresswell, 1977), with measure phrases themselves analyzed as simply denoting degrees on the relevant scale.

Other degree-based work on quantity and amount includes Schwarzschild (2006); Rett (2008); Hackl (2009); Solt (2009).

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

(11) a. Joe read a lot of books.
b. Anne went to the movies a lot 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 what is still the most comprehensive discussion 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 (11b), and the duration reading in cases like (11c). Conversely, individual level verbs 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 (11d). As evidence for the relevance of the g- vs. q-position distinction, some degree expressions are restricted to one or the other type of scale. For example, *a lot* in (11b,c) (as well as (11a)) would be translated by Dutch *veel* 'much', while in (11d) the translation would be *erg* 'badly', which also occurs with gradable adjectives.

The preceding discussion already shows that verbal scalarity is a multifaceted phenomenon. But examples of the sort in (11b-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:

(12) 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), this group includes degree achievement verbs such as *widen*, *cool* and *lengthen* (13a), directed motion verbs such as *ascend* and *descend* (13b), and incremental theme verbs such as *eat* (13c):

(13) 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 below). As one example, Kennedy & 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 (14), 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 (again, note the parallel to the adjectival example (4)).

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

Kennedy & Levin (2008) note that this is readily extended to directed motion verbs as in (13b). Incremental theme verbs as in (13c), however, present more of a challenge,

particularly in capturing the semantic contribution of the object NP. There is a long tradition of analyses of such VPs 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 & Nicolas, 2005; Piñon, 2008) taking it to be part of the content of the verb itself, while others (e.g. Rappaport Hovath, 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 & 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've 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:

(15) huge idiot; big smoker; utter disaster; real hero

(16) 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 a huge 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 (16).

While nominal gradability is undoubtedly real, we again face the issue of how it should be analyzed. One possibility is that certain nouns such as *idiot* and *disaster* are like gradable adjectives in themselves having a degree argument (per Morzycki 2009); but this raises the question of why they do not combine with the same degree modifiers found with adjectives (**very/so/too idiot*), and how then the noun versus adjective distinction should be characterized at all. Furthermore, examples of the sort in (15) and (16) plausibly involve different sorts of gradability (see Sassoon 2011; Morzycki 2012 for discussion). *Huge idiot/more of an idiot* seem to involve a scale similar to that underlying the adjective *idiotic*; *real hero* seems by contrast to involve a scale of prototypicality; and the comparative in (16b) could be metalinguistic. Thus a unified analysis of nominal gradability looks unlikely.

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 seen in the following examples, which show that they can form comparatives and compose with degree modifiers and measure phrases:

(17) 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), following 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, specifically a finitely additive probability space. 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 referenced above were drawn from English. But comparable examples could be found in any number of other languages. This raises the question of whether <u>all</u> 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:

(18) **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 e.g. in *Mary is taller than 1.70 m* or *Mary is 2 cm taller than Frank*).

Motu, a language of Papua New Guinea that forms comparatives via the so-called conjunctive strategy (19), elicits a negative answer to all of these questions, and is therefore concluded to have a negative setting for the DSP.

(19) Mary na lata, to Frank na kwadoḡi.
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 argue that its semantics is essentially context-dependent, such that (19) is true if there is a context in which Mary counts as tall while Frank counts as short.

Another language argued not to have predicates based on degree arguments is Fijian (Pearson, 2010). Furthermore, it is well known that there are languages with no number words beyond 'one', 'two', 'three/a few' and 'many' (Dixon, 1980), suggesting that they lack even the scale of number.

If there are are languages that express gradability completely without degrees, it is interesting to ask whether there might be a trace 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

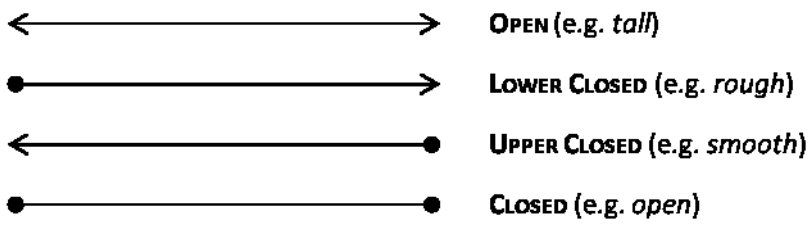
Having concluded that the meanings of certain items of natural language must be stated in terms of scales, it does not necessarily follow that scales themselves are a legitimate object of study in formal semantics. However, work in recent years has increasingly shown that the nature of scales does matter in 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, whose importance was brought to light especially by Rotstein & Winter (2004); Kennedy & McNally (2005); Kennedy (2007), relates to the presence or absence of scalar **endpoints**. The motivating data come from the distribution of degree modifiers, which exhibit seemingly ideosyncratic patterns of acceptability with different sorts of gradable adjectives:

(20) a. completely open/smooth/*rough/*tall
b. slightly open/*smooth/rough/*tall

The authors show that 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).

(21)



The lexical semantics of *completely* requires the adjective have a scale with a maximum (to be *completely 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 (21) 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 *open*, *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, while that for *rough* is defined in terms of the minimum point on the roughness scale. Relative adjectives such as *tall*, whose scales are totally open, have instead contextually determined standards, which are typically dependent on a comparison class of some sort.

Differences of this sort in scale structure impact sentence processing (Frazier et al., 2008), and even children are sensitive to them (Syrett et al., 2010). And the open/closed distinction is relevant beyond the domain of gradable adjectives. In particular, this parameter has been productively applied to account for facts relating to verbal telicity. For example, (22a) favors a telic reading, while (22b) has only an atelic reading, which can be explained in that the process of emptying has an endpoint (complete emptiness), while the process of widening does not (Kennedy & Levin, 2008).

(22) a. The sink emptied.
b. The gap between the boats widened.

The nature and role of scalar endpoints is an ongoing topic of research, and some important questions remain open, among others regarding the nature of the diagnostics employed for the presence of scalar maxima/minima, and the extent to which this aspect of scale structure determines standard of comparison (for discussion, see Kagan & Alexeyenko, 2010; Lassiter, 2011; Klecha, 2012). Some authors (notably McNally, 2011) have also developed alternative, non-scale-structure based accounts of the relative vs. absolute distinction. It is fair to say that the final words have not been written on this topic. But the more general conclusion remains, namely that through investigating the structure of scales, and how these vary, we open up a productive path towards explanation of linguistic facts.

Perhaps surprisingly given the amount of ongoing research in this area, there is as of yet no general model of the structure of scales, nor, as noted by Kennedy & McNally (2005), is there a clear understanding of the full set of linguistically relevant parameters of structural variation. As discussed in Section 2, the typology of levels of measurement (ordinal/interval/ratio) does not seem to adequately characterize the variation in natural language scales. What is called for is a linguistically oriented model. While I cannot present a fully developed model of this sort here, in the next section I offer some preliminary remarks on what it will need to encompass.

4.2 Towards a general model

In Section 2 it was proposed that a scale can be conceptualized as a triple consisting of a set of degrees D, an ordering relation on that set $\succ$, and a dimension *DIM*. The discussion

below is organized around variation in each of these three components, though they cannot always be treated as independent of one another.

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, and whether there is variation in this. Degrees are most commonly assumed to be points on a scale, but various authors over the years have argued they should instead be conceptualized as intervals or extents (e.g. Seuren, 1978; Bierwisch, 1989). As a recent 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), which he proposes allows 'cross-polar anomalies' of the form in (23) to be ruled out, in that the two types of degrees cannot be compared.

(23) *Anna is taller than Lisabeth is short.

Schwarzschild & Wilkinson (2002) propose a somewhat more radical interval-based semantic theory to account for examples involving quantifiers in comparatives, where they argue point-based theories do not derive correct truth conditions. Other approaches have treated degrees as vectors (Zwarts & Winter, 2000; Winter, 2001), or as some other sort of complex entity (Grosu & Landman, 1998).

It is not clear however that these more complex treatments of degrees are actually necessary. In subsequent work, the phenomena used to justify interval-based semantics have been addressed within a more traditional point-based view, generally by treating intervals as sets of degrees-as-points (see especially Büring, 2007; Beck, 2011). I thus provisionally assume that degrees can be treated as points.

A further aspect of the structure of D that has been 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, duration and such, but Fox & 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 & Hackl argue that the assumption of universal scale density allows for a unified account of a variety of phenomena involving maximalization operations. One example involves scalar implicatures. Unmodified numerals on some accounts give rise to upper-bounding implicatures; for example, (24a) implicates that John does not have 4 children (i.e. that 3 is the maximum number n such that he has n children). But (24b) lacks the corresponding implicature that he does not have more than 4 children. A similar absence of implicature is observed in (25), but here there is an intuitively clear reason: since the scale of weight is dense, there <u>is</u> no maximum degree d such that John weighs more than d. Crucially, this

explanation can only be extended to (24b) if we assume that the scale of cardinalities is likewise dense rather than discrete.

(24) 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

(25) John weighs more than 120 pounds

Similar reasoning can be extended to asymmetries with overt *only* and certain constraints on *wh*-movement (see also Nouwen 2008 on implicatures with negated numerical comparatives). Other authors, however, have argued that these phenomena can receive an explanation that do not rely on assumptions about density (Mayr, to appear). As such, it is not yet conclusively established whether density is a universal property of measurement scales, or whether dense vs. discrete is a parameter on which scales vary.

Another meaningful point of variation is whether or not there exists a standard unit of measure, 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 in (8) suggests that this factor also impacts what sort of proportional comparisons can be expressed with an adjective: those with corresponding numerical measures (e.g. *tall*) allow precise ratio comparisons, while those without (e.g. *beauty*) do not. 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 names or 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 aspect of D that is related to the previous two is **granularity**, a notion introduced by Krifka (2007) to account for the approximate interpretation of round numbers. For example, (26a) allows or even favors an approximate reading ('about 100'), while (26b) is necessarily interpreted precisely, a pattern Krifka attributes to a difference in the granularity of the underlying measurement scale: (26b) is interpreted relative to the unit scale (the integers), while (26a) on its approximate interpretation involves a coarser-grained scale, e.g. one based on units of 10.

(26) 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, years and so forth.

The mechanism of granularity has also been applied to other phenomena, such as the semantics of approximators like *exactly* and *approximately* (Sauerland & Stateva, 2007).

To date, granularity has not been incorporated into a more general model of scale structure. There is first of all no consensus as to whether it is a semantic phenomenon (per

Sauerland & Stateva, 2007) or a pragmatic one (Lasersohn, 1999; Lauer, 2012). That is, does (26a) 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: via entirely distinct scales, per Krifka's account (an analysis which would contradict universal scale density); via granularity functions that map scale points to intervals of varying width around them (Sauerland & Stateva, 2007); or perhaps some other way. Also to investigate is what further phenomena can be explained in terms of granularity. For example, if I utter it *The theater is empty tonight* to describe a situation where there are unexpectedly few people present, is this an instance of the scalar endpoint being interpreted at a very coarse-grained level (Kennedy, 2007) or of 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 & McNally (2005), the scales lexicalized by antonym pairs such as *tall/short* and *flat/bumpy* 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 (27): the degrees corresponding to Anna's and Elizabeth's 'tallness' are the same as those corresponding to their 'shortness', but the ordering between them is reversed.

(27) Anna is taller than Lisabeth $\Leftrightarrow$ Lisabeth is shorter than Anna

Another potential area of parametric 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 (i.e. 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 $a > b + \epsilon$. The result is a so-called semi-order (Luce, 1956), an ordering where $\succ$ is transitive but indifference is not, such 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 real-life comparison actually works, 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), and 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 similar sort of tolerance, an example being implicit comparisons such as *Anna is tall compared to Lisabeth*, which is infelicitous if the difference between the two heights is small. Fults (2011) analyzes such examples via a semi-ordered scale based on psychological models of humans' approximate numerical capabilities, and Solt (2011b) proposes something similar underlies the semantics of the quantifier *most* (for a non-scalar account based on tolerance, see also Cobreros et al. 2012). Another phenomenon that might lend itself to explanation in terms of partially ordered scales involves multidimensional adjectives

such as *clever*, which seem not to impose a total ordering on a domain. It must be acknowledged, however, that allowing structures of this sort would represent a significant departure from the more usual notion of a scale.

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. We have already seen some possibilities.

As one example, it was shown above that dimensions with a corresponding **numerical measurement system** behave differently than those without one. As another, dimensions could potentially be classified according to how measurement interacts with **concatenation**: as noted in Section 2, some dimensions (e.g. height) are additive w.r.t. concatenation, while others (e.g. temperature) are intermediate, and there may be other possibilities as well.

Beyond this, in Section 3 we saw that the distinction between degrees of **quantity and grade** has linguistic relevance, determining for example the distribution of Dutch degree modifiers *erg* and *veel*. Schwarzschild (2006) introduces a related distinction between dimensions such as weight that are **monotonic** on the part/whole relation between entities and those like purity that are not; here monotonicity means that the measure of any proper subpart of an individual is less than that of the individual as a whole, a property that intuitively corresponds to dimensions of quantity. Only measure phrases corresponding to monotonic dimensions occur in partitives (e.g. *20 grams of gold* vs. **20 karats of gold*). But the distinction is a slippery one: 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*). Thus the nature of this potential parameter, and its role, remains to be better understood.

Finally, there is another intuitive but difficult-to-define distinction between **evaluative and non-evaluative** dimensions. Again, adjectives corresponding to the two classes pattern differently. For example, in contrast to the equivalence illustrated in (27) for *tall/short*, that in (28) for *smart/dumb* does not seem to hold, in that the second sentence is infelicitous if both Anna and Lisabeth are judged smart.

(28) Anna is smarter than Lisabeth $\not\Leftrightarrow$ Lisabeth is dumber than Anna

There have been various attempts to capture the inherent 'norm-relatedness' of comparatives like *dumber* in (28) (e.g. Bierwisch, 1989; Rett, 2008). One scale-based approach (based on Cruse, 1986) 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 scales tracking evaluative vs. non-evaluative dimensions, corresponding more directly to the difference between objective measurement and subjective, speaker-dependent judgment.

5 Scales in pragmatics

The scales we have been considering to this point might be called semantic scales. A different sort of scale plays a role in pragmatic inferencing. By way of illustration, (29a) 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 pragmatic 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 (29b). From the speaker's choice to use the weak 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).

(29) a. Anna read some of the books on the reading list.
b. $\langle all, some \rangle$

A wide variety of lexical items form Horn scales, including quantifiers (as above), logical connectives (e.g. $\langle and, or \rangle$), gradable adjectives ($\langle beautiful, pretty \rangle$, $\langle hot, warm \rangle$), and verbs ($\langle know, believe \rangle$). 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 (e.g. in some contexts *I typed the letter* implicates I didn't yet mail it). But not all ordered sets give rise to implicatures (for example, *if* doesn't 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 & Levinson, 1981; Matsumoto, 1995).

At first glance, Horn scales are very different sorts of entities from the scales that have been the focus of this paper. Most importantly, Horn scales are composed of linguistic expressions, while semantic scales are part of the domain that provide such expressions with their content. But the two are not unconnected. In the case of adjectival Horn scales, the associated semantic scale provides for their ordering (e.g. *hot* corresponds to a range higher on the heat scale than *warm*). The connection is even closer in the case of number words, where it is generally assumed that the scale that provides them with their semantic content is the same as that against which implcatures are calculated (cf. the above discussion of Fox & Hackl 2006). But there is evidence this is not the case: Cummins et al. 2012 demonstrate that modified numerals such as *more than 100* give rise to implicatures relative to a coarser level of granularity than that underlying their semantic meaning.

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

References

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

Bale, Alan Clinton. 2008. A universal scale of comparison. *Linguistics and Philosophy* 31(1). 1–55. doi:10.1007/s10988-008-9028-z.

Bale, Alan Clinton. 2011. Scales and comparison classes. *Natural Language Semantics* 19(2). 169–190. doi:10.1007/s11050-010-9068-0.

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

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

Beck, Sigrid, Svetlana Krasikova, Daniel Fleischer, Remus Gergel, Christiane Savelsberg nd aJohn Vanderelst & Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9. 1–66.

Bierwisch, Manfred. 1989. The semantics of gradation. In Manfred Bierwisch & Ewald Lang (eds.), *Dimensional adjectives*, 71–261. Berlin: Springer Verlag.

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

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

Büring, Daniel. 2007. Cross-polar nomalies. In *Semantics and Linguistic Theory (SALT) XVII*, CLC Publications.

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

Caudal, Patrick & David Nicolas. 2005. Types of degrees and types of event structures. In C. Maienborn & A. Wöllstein (eds.), *Event arguments: Foundations and application*, 277–299. Tübingen: Max Niemeyer Verlag.

Cobreros, Pablo, Paul Egré, David Ripley & Robert van Rooij. 2012. Tolerant, classical, strict. *Journal of Philosophical Logic* 41(2). 347–385.

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

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

Cummins, Chris, Uli Sauerland & Stephanie Solt. 2012. Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy* 35(2). 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. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English*: Leiden University Ph.D. dissertation.

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

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

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

Filip, Hana & Susan Rothstein. 2006. Telicity as a semantic parameter. In J. Lavine, S. Franks, Tasseva-Kurktchieva & H. Filip (eds.), *Formal approaches to Slavic linguistics 14 (the Princeton meeting).*, 139–156. Ann Arbor: Michigan Slavic Publications.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. doi:10.1007/s10988-006-9004-4.

Frazier, Lyn, Jr. Charles Clifton & Britta Stolterfoht. 2008. Scale structure: processing minimum standard and maximum standard scalar adjectives. *Cognition* 106(1). 299–324.

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

Gawron, Jean Mark. 2005. Generalized paths. In E. Georgala & J. Howell (eds.), *Proccedings of semantics and linguistic theory xv*, CLC Publications.

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

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

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

Hackl, Martin. 2000. *Comparative quantifiers*: Massachusetts Institute of Technology Ph.D. dissertation.

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 & Beth Levin. 1999. Scalar structure underlies telicity in "degree achievements". In T. Mathews & D. Strolovitch (eds.), *Proceedings of semantics and linguistic theory (SALT) 9*, Ithaca, NY: CLC Publications.

Heim, Irene. 2000. Degree operators and scope. In Brendan Jackson & Tanya Matthews (eds.), *Proceedings of the 9th semantics and linguistic theory conference*, CLC Publications.

Heim, Irene. 2006. *Little*. In Masayuki Gibson & Jonathan Howell (eds.), *Proceedings of the 16th semantics and linguistic theory conference.*, CLC Publications.

Hirschberg, Julia. 1985. *A theory of scalar implicature*: University of Pennsylvania dissertation.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in English*: University of California, Los Angeles dissertation.

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

Kagan, Olga & Sascha Alexeyenko. 2010. Degree modification in Russian morphology: The case of the suffix *-ovat*. In Yehuda N. Falk (ed.), *Proceedings of israel association for theoretical linguistics 26*, Bar Ilan University.

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 predicates. *Linguistics and Philosophy* 30. 1–45. doi:10.1007/s10988-006-9008-0.

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

Kennedy, Christopher & Beth Levin. 2008. Measure of change: The adjectival core of verbs of variable telicity. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*, 156–182. Oxford: Oxford University Press.

Kennedy, Christopher & Louise McNally. 2005. Scale structure and the semantic typology of gradable predicates. *Language* 81. 1–37. doi:10.1353/lan.2005.0071.

Klecha, Peter. 2012. Positive and conditional semantics for gradable modals. In Ana Aguilar Guevara, Anna Chernilovskaya & Rick Nouwen (eds.), *Proceedings of sinn und bedeutung 16*, vol. 2, 363–376.

Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4(1-45). doi:10.1007/BF00351812.

Klein, Ewan. 1982. The interpretation of adjectival comparatives. *Journal of Linguistics* 18. 113–136. doi:10.1017/S0022226700007271.

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

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

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

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

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

Krifka, Manfred. 2007. Approximate interpretations of number words: A case for strategic communication. In *Cognitive foundations of interpretation*, chap. Approximate interpretation of number words: a case for strategic communication, 111–126. Koninklijke Nederlandse Akademie van Wetenschapen.

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

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75(3). 522–551.

Lassiter, Dan. 2011. *Measurement and modality: the scalar basis of modal semantics*: New York University Ph.D. dissertation.

Lauer, Sven. 2012. On the pragmatics of pragmatic slack. In Ana Aguilar Guevara, Anna Chernilovskaya & Rick Nouwen (eds.), *Proceedings of sinn und bedeutung 16*, vol. 2, .

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. to appear. Implicatures of modified numerals. In Ivano Caponigro & Carlo Cecchetto (eds.), *From grammar to meaning: The spontaneous logicality of language*, 139–171. Cambridge University Press.

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

Morzycki, Marcin. 2009. Degree modification of gradable nouns: size adjectives and adnominal degree morphemes. *Natural Language Semantics* 17(2). 175–203.

Morzycki, Marcin. 2012. Adjectival extremeness: Degree modification and contextually restricted scales. *Natural Language and Linguistic Theory* 30(2). 567–609.

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

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

Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Groenendijk et al. (ed.), *Studies in discourse representation theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris.

Pearson, Hazel. 2010. How to do comparison in a language without degrees: a semantics for the comparative in Fijian. In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Proceedings of sinn und bedeutung 14*, 356–372.

Piñon, Christopher. 2008. Aspectual composition with degrees. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics and discourse*, 183–219. Oxford: Oxford University Press.

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

Rett, Jessica. 2008. *Degree modification in natural language*: Rutgers University Ph.D dissertation.

Roberts, Fred S. 1979. *Measurement theory with applications to decisionmaking, utility, and the social sciences*. Cambridge: Cambridge University Press.

van Rooij, Robert. 2011. Measurement, and interadjective comparisons. *Journal of Semantics* 28(3). 335–358. doi:10.1093/jos/ffq018.

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

Rotstein, Carmen & 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(1). 151–180. doi: 10.1007/s11229-009-9687-5.

Sassoon, Galit. 2011. Adjectival vs. nominal categorization processes. *Belgium Journal of Linguistics* 25. 104–147.

Sauerland, Uli & Penka Stateva. 2007. Scalar vs. epistemic vagueness. In Masayuki Gibson & Tova Friedman (eds.), *Proceedings of semantics and linguistic theory (salt) 17*, Ithaca, NY: CLC Publications.

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

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

Seuren, Pieter A.M. 1978. The structure and selection of positive and negative gradable adjectives. In *14th regional meeting of the chicago linguistic society*, .

Solt, Stephanie. 2009. *The semantics of adjectives of quantity*: City University of New York Ph.D. dissertation.

Solt, Stephanie. 2011a. Notes on the comparison class. In Rick Nouwen, Robert van Rooij, Uli Sauerland & Hans-Christian Schmitz (eds.), *Vagueness in Communication (ViC2009), Revised Selected Papers (LNAI 6517)*, 189–206. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-18446-$8_1$1.

Solt, Stephanie. 2011b. Vagueness in quantity: Two case studies from a linguistic perspective. In Petra Cintula, Chris Fermüller, Lluis Godo & Petr Hajek (eds.), *Understanding vagueness. logical, philosophical and linguistic perspectives*, 157–174. London: College Publications.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3. 1–77. doi:10.1093/jos/3.1-2.1.

von Stechow, Arnim. 2006. Times as degrees: *früh(er)* 'early(er)', *spät(er)* 'late(r)', and phrase adverbs. Unpublished manuscript.

Stevens, Stanley Smith. 1946. On the theory of scales of measurement. *Science* 103. 677–680. doi:10.1126/science.103.2684.677.

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

Syrett, Kristen, Christopher Kennedy & 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.

Winter, Yoad. 2001. Measure phrase modification in a vector space semantics. In K. Megerdoomina & L. A. Bar-el (eds.), *Proceedings of wccfl 20*, 607–620. Somerville, MA: Cascadilla.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916–937.

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