

Gradability, Vagueness, and Scale Structure: From the Armchair to the Lab



Elena Castroviejo, Louise McNally and Galit W. Sassoon

Abstract In this chapter we present an overview of three main issues that have surrounded the study of gradable properties—vagueness, measurement, and dimensionality—and how they have been pursued from the perspectives of philosophy, linguistics, and psychology. We then provide a brief summary of each chapter in the volume, together with a guide to how the chapters relate to each other thematically.

Keywords Semantics · Properties · Adjectives · Vagueness · Measurement Dimensionality

1 Introduction

Many properties that we use to describe individuals or categories of individuals—dimension, texture, emotions, worth, are just a few examples—manifest themselves to a greater or lesser degree, and it is often relevant to group, order, or compare individuals according to the degrees they possess of the properties in question. This

E. Castroviejo (✉)
Department of Linguistics and Basque Studies,
University of the Basque Country (UPV/EHU),
Vitoria-Gasteiz, Spain
e-mail: elena.castroviejo@ehu.eus

E. Castroviejo
Ikerbasque, Basque Foundation for Science, Bilbao, Spain

L. McNally
Department of Translation and Language Sciences, Universitat Pompeu Fabra,
Barcelona, Spain
e-mail: louise.mcnally@upf.edu

G. W. Sassoon
Department of English Literature and Linguistics, Bar Ilan University,
Ramat Gan, Israel
e-mail: galit.weidman-sassoon@biu.ac.il

fact has raised important questions in the fields of philosophy, linguistics, and the study of cognition.

First, there is the vagueness question. If a property (take size as an example) can be held to different degrees, on what basis do we partition a set of individuals based on their size—how do we draw the line between the large and the small? What is different about vague predicates like *long* versus those that are arguably not vague (or needn't be vague), such as *spotted*, for which drawing the line is much easier?

Second, there is the measurement question. By what means do we order or compare properties, particularly when these are not reducible to a single measurable dimension, as is the case with e.g. beauty? How can we best model the semantics of degree, measure, and comparison constructions in language?

Third, logically prior to the measurement question is the dimension selection question. Many properties can be compared on more than one dimension. For example, we can compare the intensity of a color, or its extension on a surface. This and other sorts of variability have led to the study of what has come to be known as *scale structure*, which can be understood in the broadest sense as the study of how lexical semantics—including not only the multidimensionality of properties, but also other factors as well—is connected to more formal properties of gradability, such as whether a property can be held to a minimal or maximal degree (i.e., is associated with a *bounded* or *unbounded* scale), and what the nature is of the *standard* or threshold for truthful ascription of the property.

These problems have been amply explored over the years from a highly theoretical/conceptual perspective (“the armchair”). However, with the increasing presence of experimental methods both in philosophy and linguistics, theoretical analyses are now being tested experimentally. The chapters in this volume, contributed by philosophers, linguists and psychologists, all reflect this move from the armchair to the lab. In the remainder of this introduction, we first present a very brief review of some of the most important perspectives on the semantics of gradability, vagueness, and scale structure found in the philosophical and linguistics literature, as well as in the study of cognition and categorization. We then offer an overview of each chapter.

2 Perspectives

2.1 *The Perspective from Philosophy and Logic*

The philosophical literature has focused in large part on the challenge of understanding how reasoning with vague concepts is possible at all. Classical logic embraces the law of the excluded middle, making it impossible for an entity to be neither P nor not P , or to be both P and not P simultaneously. But precisely this seems possible and sometimes even rather natural for vague predicates like *tall* or *long*. Given our world knowledge concerning pencils and their varying sizes, an 18-cm long pencil is obviously long, and a 5-cm long pencil is obviously not long. But it is not clear whether a 10- or 12-cm long pencil is long or not. When asked about it, speakers

may well accept that such a pencil is “neither long nor not long,” or rather that such a pencil is “long and not long” (cf. Alxatib and Pelletier 2011; Égré et al. 2013).

Like contradictions, paradoxes also put the foundations of logic at risk, and therefore philosophers have long been engaged in a search for a solution to famous examples such as the Sorites paradox, illustrated in (1), where starting with natural premises such as **P1** and **P2**, the absurd conclusion **C** is reached through iteration. The question is whether a logic that derives this paradox is sound.

- (1) **P1**: A 20-cm long pencil is long.
P2: Any pencil that is 1 mm shorter than a long pencil is long.
C: A 1-cm long pencil is long.

These manifestations of vagueness have triggered the development of rich and diverse theories of the logic underlying the use of vague predicates. These theories offer alternative theoretical mechanisms for the representation of borderline cases, imprecise boundaries, the Sorites paradox, and the context dependency of vague predicates. Epistemicists (Williamson 1994; Fara 2000) urge us to conserve classical logic, but it is hard to resist accepting that vague predicates have no true cut-off points for applicability, as supervaluationists maintain (Fine 1975; Kamp 2013; Keefe 2000). It is also hard to disallow contradictions, as subvaluationists do (Hyde 1997; Cobreros 2011a) or to resist accepting the tolerance of vague predicates to small differences, a property that gives rise to the intuition that premises like **P2** in (1) are true. Moreover, the acceptance of tolerance takes us further away from classical logic (Kamp 2013; Shapiro 2006; Soames 1999; Stanley 2003; van Rooij 2011; Cobreros 2011b; Cobreros et al. 2012; Ripley 2011).

In an effort to determine which of the competing theoretical accounts of vagueness is more successful at explaining natural language data, experimentation has become not only welcome but also necessary. Philosophical experiments on vague adjectives most commonly investigate classifications on Sorites-like distributions, where the set of entities covers a full range of values for a property in a certain range with no gaps in between any two values (cf. Alxatib and Pelletier 2011; Égré et al. 2013; Raffman 1994, 2005; for a more recent study, see Verheyen et al. 2016). These experiments usually assess judgments of the truth of vague sentences (e.g. *X is tall*) or of the acceptability of contradictory or tautological predicates (e.g. *P and/or not P*; see Égré and Zehr, this volume). The related problem of faultless disagreement has also been explored recently (see the study by Solt in this volume and references cited there). Contradictory sentences and cases of faultless disagreement share the fact that they put the consistency of natural language at risk; they differ in that in the former case, one and the same speaker asserts or accepts the truth of a proposition and its negation (e.g. *John is tall and not tall*), while in the latter, the inconsistent propositions are ascribed to two different speakers, neither of which can be considered wrong in their belief.

Philosophers have typically been less concerned with other manifestations of vagueness in language and thought, such as the connection between vagueness and gradability, or the manifestation of vagueness in degree modification and comparison, issues to which we now turn.

2.2 *The Perspective from Linguistics*

We have seen that vagueness poses a problem for a theory of truth and for the foundations of logic: Is a binary system yielding truth and falsity enough to cover borderline cases, and can such a system be consistent? The linguist's take on vagueness is quite different. The starting point is rather the question of how vagueness is linguistically realized. One straightforward answer is that it is encoded as gradability, and yet this is just the beginning of the story.

Let us start by noting that gradability is naturally realized on adjectives and adverbs. The grammatical diagnostic to determine whether an adjective is gradable consists in placing it in a comparative construction, as in (2).

- (2) a. Maria is taller than John.
 b. Your shirt is wetter than mine.

Gradable adjectives like *tall* and *wet* pose an interesting problem for linguists because they seem to have many possible interpretations depending on the phrases in which they occur (for example, which degree modifiers they combine with) and the situations in which these phrases are used. For instance, adjectives like *tall* cannot easily be modified by *slightly*, while adjectives like *wet* can (Rotstein and Winter 2004; Kennedy and McNally 2005; Kennedy 2007b; Sassoon 2012b). This and other grammatical distinctions reveal a particular way of integrating contextual information in the lexical meaning of such adjectives. While some adjectives—those known as *relative* or *open scale*—resort to a contextual standard (to assert that Maria is tall, we need to consider similar individuals and average over their heights), for others—the *absolute* or *closed scale* adjectives—it is enough to attend to their conventional meaning to be certain that they can be truthfully predicated of an individual (but see Sassoon and Toledo 2011; McNally 2011 for qualifications).

Linguists are interested in determining what the core meaning components of gradable adjectives are that enable speakers to use them in different ways, and what information from the context of use speakers find relevant for interpretation (e.g. for standard selection), and what they ignore. For example, Aparicio et al. (2015) and Aparicio et al., this volume, report on a series of eye-tracking experiments which allow us to better understand the correlation between integration of contextual information and the scale structure of adjectives. A different line of research concerns how degree modifiers can only be used not only as indicators of lexical semantics, but also as sociolinguistic markers. Beltrama, this volume, presents a study that aims to characterize the social meanings associated with putatively illicit cases of modification such as *totally tall* (on this, see also Beltrama 2016, 2018a; Beltrama and Staum Casasanto 2017).

Linguists are also interested in formally representing the lexical meaning of vague predicates and the way they combine with their arguments and modifiers. On one view, let's call it the *vagueness approach*, gradable predicates such as *tall* should have the same semantic type as non-vague predicates such as *four-legged*; the difference is simply that the former have a gap in their extension, corresponding to the set of

individuals that cannot be said to be in either the positive or the negative extension of the predicate. There are different versions of the vagueness approach, most of which rely on the notion of an ordering to derive the degree-like properties of such predicates (Cresswell 1976; Klein 1980; van Rooij 2011). This can be illustrated as follows.

Assume that an adjective like *tall* denotes a partial function with a positive extension, a negative extension, and an extension gap. To judge whether a sentence including a vague adjective is true, context provides a comparison class (a collection of relevant similar objects); the choice of comparison class influences which entities fall into the positive and negative extensions, and which fall into the gap. For example, in (3), our judgment will partly depend on the physical properties of Peter and on whether the comparison class is the set of boys, male adults, jockeys or basketball players.

(3) Peter is tall.

For some comparison classes, Peter will fall into the positive extension, while in others he will fall into the negative extension, and in still others he might fall into the gap. The variation in the positive versus negative extensions across different comparison classes determines the ordering of entities relative to the given predicate: An entity x will be ordered higher than another entity y iff we can find a comparison class (most relevantly, the one consisting just of x and y) for which x falls into the positive extension and y falls into the negative extension—the assumption being that for no choice of comparison class can either of the two be empty. Thus, entity orderings (or even degrees, where those are needed) are derived from the entity sets forming the basic denotations of predicates (see especially van Benthem 1982; Bale 2008; Burnett 2016).

A more popular view in recent years is one on which entity sets are derived, and degrees are the primitives from which these sets are derived. Gradable predicates include a degree argument which is bound after modification by certain degree expressions such as *very*, *completely*, etc. This *degree approach* has at least two implementations. Gradable adjectives are either analyzed as relations between degrees and individuals (type $\langle d, et \rangle$), as proposed by e.g. Seuren (1973), von Stechow (1984), Heim (1985), or Bierwisch (1989), or as measure functions of type $\langle e, d \rangle$, as proposed by e.g. Bartsch and Vennemann (1972), Kennedy (1999, 2007b). On this view, to avoid a type mismatch in the composition of (3), and to introduce a standard for the truthful application of the predicate, it is common practice to assume the presence of a null positive degree morpheme—*pos*—that combines with the gradable adjective and yields a predicate of individuals (von Stechow 1984; Kennedy 2007b, a.m.o.). See Solt and Gotzner (2012) for an experimental study that bears on the question of whether degrees are among the primitives of natural language ontology or whether it is sufficient to posit ordering relations.

Up to this point in the discussion, we have made the maximally simple assumption that gradable predicates are, crucially, unidimensional. An adjective like *tall* lexicalizes a scale related to height, which can be associated with a set of degrees,

a dimension and an ordering. However, the empirical domain is much broader and more complex. For example, beyond dimensional adjectives, there are evaluative adjectives in the sense of Bierwisch (1989), such as *industrious*, for which there is no obvious measuring system. Should such adjectives include a degree argument? Is there a simple scale of industriousness, or rather are there various dimensions that can be evaluated when we assert that someone is or is not industrious?

Multidimensionality is precisely the focus of Sassoon (2012a, 2013a); Sassoon and Fadlon (2017), who build on Bartsch and Vennemann (1972) and Bartsch (1984, 1986) to develop and empirically test a quantificational analysis for such adjectives, examples of which also include *healthy* and *intelligent*. Quantification plays a role in the sense that, in some cases, in order for an entity to count as bearing a multidimensional property, *all* of its dimensions must be satisfied, while in other cases, it is argued, only *some* dimensions must be satisfied. Thus, for example, someone is healthy if she is healthy in every way, and sick if she is sick in some way.

The contribution by Solt to this volume looks at a different aspect of multidimensionality, namely its role in disagreements about orderings (Kennedy 2013; Bylinina 2014; Umbach 2016; McNally and Stojanovic 2017). Solt notes that even some physical properties are arguably multidimensional. For example, two individuals looking at photographs of roads might disagree about whether one road is bumpier than another because one places more weight on the number of bumps, while the other focuses on their size or severity. Solt presents evidence that the disagreement that can arise in such cases is distinct from the disagreement found with adjectives like *fun*, which arises due to variation in our subjective experience.

Another focus of study in the linguistic research on vagueness regards the categories in which gradability is realized. Gradability has been observed not only in adjectives and adverbs, but in verbs and nouns, as well (see Bolinger 1972; Hay et al. 1999; Vanden Wyngaerd 2001; Kennedy and Levin 2008; Piñón 2008; Kennedy 2012; Bochnak 2013b; Rappaport-Hovav 2014; Fleischhauer 2016, a.m.o.; see Wellwood 2014 for a recent attempt to unify gradability across categories under a single analysis). Consider, for example, the following data from Morzycki (2009) and Constantinescu (2011), respectively.

- (4) a. a big idiot/stamp collector
 b. that idiot/*stamp collector of a doctor.

As in the case of gradable adjectives, to determine whether a noun is gradable, linguists pay attention to the characteristics of the syntactic constructions in which they appear (and the lexical semantics of the categories with which they are composed). In (4a), we are inclined to consider *idiot* and *stamp collector* as gradable because they are modified by *big*, which does not restrict the size of the referent of the noun, but rather the degree of idiocy or of stamp collectorhood. In (4b), on the other hand, the *N-of-an-N* construction, analyzed by Bolinger (1972) and Matushansky (2002), among others, as a diagnostic for gradability, gives different results for *idiot* and *stamp collector*, whereby only the former but not the latter would be analyzed as such. These and other tests have been used by Constantinescu (2011) to conclude that gradability in the nominal domain is not a uniform phenomenon.

Data such as (4a) have led Morzycki (2009) and others to assign nouns a degree argument as part of their lexical entry. For a set of different constructions involving natural kind nouns such as *duck* and social nouns like *philosopher* (as in “more a duck than a goose” and “more a linguist than a philosopher”), Sassoon (2017) also advocates a degree analysis whereby the conceptual structure of nouns and closeness to prototypes or stereotypes can be modeled in the same way as for gradable predicates. The contribution to this volume by de Vries tests the viability of this sort of analysis. Others, such as Constantinescu (2011) and Beltrama and Bochnak (2015), defend a degreeless analysis for intensifying constructions that builds on quantification over contexts or worlds in an epistemic modal base. In a similar vein are those analyses that completely rely on degrees of precision, as in Lasersohn’s (1999) use of slack regulation and Morzycki’s (2011) more recent implementation.

The comparability of properties which otherwise do not appear to be gradable raises a final theoretical question that has concerned linguists interested in natural language metaphysics, namely: What is a degree? The advocates of the degree approach to gradability assume that degrees are primitive parts of the ontology of natural language. Nouwen and Dotlačil, this volume, go further and present experimental evidence that degrees are not only atomic entities, but also, just like regular individual objects, can cumulate as pluralities. Other options for the treatment of degrees include Grosu and Landman’s (1998) account, on which degrees consist of a measure value, a measure domain and the object measured. In contrast, among those who do not treat degrees as primitives, we find degrees being equated to equivalence classes of individuals (Cresswell 1976; Klein 1980) and to state kinds (Anderson and Morzycki 2015). Finally, in various works, Moltmann (2004, 2007, 2009) claims that tropes, roughly realizations of properties, are enough to account for gradability phenomena. See also Scontras (2014) for a recent in-depth development of the semantics of amounts and degrees.

To understand the linguistic realization of gradability, analyzing crosslinguistic data is essential. Beck et al. (2004, 2009), Kennedy (2007a); Bobaljik (2012), and Bochnak (2013a) are examples of theoretical works that either compare the syntax-semantics mappings of linguistic constructions, such as the comparative, in parametrically distinct languages, or else provide a formal account of these constructions in under-represented and under-studied languages. Experimental crosslinguistic research on gradability phenomena is also ongoing: Examples include Pancheva and Tomaszewicz (2011), O’Connor et al. (2012) and Tucker et al., this volume.

Let us conclude this subsection by zooming out to the level of the utterance. As noted in e.g. Doran et al. (2009) and Beltrama and Xiang (2012), statements including gradable predicates can give rise to scalar implicatures. That is, gradable predicates are sometimes part of a scale of lexical items of varying strengths, and thus asserting that something is, for example, good may give rise to the (cancellable) implicature that it is not excellent. However, gradable adjectives do not behave exactly like other sorts of expressions that can be ordered on scales (such as numerals or quantifiers), in the consistency with which they yield scalar implicatures (van Tiel et al. 2016). The question is why. Building on Krifka’s (2002) observation that scalar implicature with numerals is sensitive to contextually-relevant granularity (for example, whether exact

numbers or e.g. multiples of 100 are under discussion), McNally (2017) suggests that adjectival scales might be much more sensitive to granularity than numerals. Cummins' contribution to this volume supports the view that, rather than granularity as such, what matters is the set of salient scalar alternatives in the context.

The linguistic and logical complexity posed by vague and gradable expressions raises questions concerning the mechanisms required for their processing and the manifestation of these in the online processing of language. Such questions, including those regarding the stages at which different types of information (e.g. grammatical vs. visual context) intervene and assist or affect processing, are typically addressed by psychologists and psycholinguists. Linguists and philosophers theorizing on language have historically drawn a static picture focusing on grammatical rules as opposed to online processes of structure building or parsing and the dynamics of semantic composition as a sentence unfolds. It is only recently that they have begun testing theories of gradability, vagueness, and scale structure using online processing measurements of reaction time (RT), eye tracking or electrophysiological activity in the form of event-related potentials (ERP). For more on this, see the next subsection.

2.3 *The Perspective from the Study of Cognition and Categorization*

A rich psychological and psycholinguistic tradition has studied the structure of concepts and the processes of online categorization and their connections to the lexical semantics of nouns and verbs (see e.g. Rosch and Mervis 1975; McCloskey and Glucksberg 1978; Hampton 1979; Osherson and Smith 1981; Barsalou 1993; Hampton 1998, 2007; Verheyen et al. 2010). Two chapters in this volume fall into this tradition: The study by Verheyen and Storms, which focuses on intersubjective variation in classification under vague concepts; and that by de Vries, which investigates subtle distinctions in the way subjects categorize different types of nouns, inspired by influential accounts of vagueness and gradability in the nominal domain (Kamp and Partee 1995; Morzycki 2009). The contribution by Schumacher, et al., addresses the online processing of categories in the context of modifiers such as *real* and *fake*, using ERPs. A related line of psycholinguistic work on dimension indeterminacy and multidimensionality is also developing (see Sassoon and Fadlon 2017 for a judgment study and Sassoon et al. (t.a.) for an ongoing ERP project).

In contrast, the exploration of adjectives—the prime examples of vagueness and gradability—has generally lagged in this tradition behind that of nouns or verbs (Cappelletti et al. 2008, note 1). A notable exception is the early study by Rips and Turnbull (1980), who reported experimental evidence to the effect that relative standards for property ascription are not stored. Rather, in each context of use of an adjective like *big*, the characteristic size of one of the categories of the entity argument (e.g. child or female) determines the standard; when no such category is

particularly salient, additional factors are likely to affect the standard choice, such as the size of the speaker. Interestingly, they also found that this dependency of the standard on a comparison class or speaker was not manifest in absolute adjectives. Further studies exploring adjectival property ascription (including in first language acquisition), especially what sorts of factors figured into judging objects as e.g. big versus little, were carried out later in the 1980s by, e.g. Smith and colleagues (Smith et al. 1986, 1988, and references therein).

The linguistic work on scale structure in the first half of the 2000s triggered a burst of experimental research into the cognitive basis for scale structure and its connections to vagueness. For example, Frazier et al. (2008) considered adjectives like *dirty*, whose standard is identified with the minimum on their scale (a minimal amount of dirt suffices for something to count as dirty) and adjectives like *clean*, whose standard is identified with the maximum on their scale (maximal cleanliness is required for something to count as clean). Frazier et al. (2008) substantiated reported intuitions that minimum standard adjectives like *dirty* are more acceptable with minimizers like *slightly* and *a little* than are maximum standard adjectives like *clean*. The former were regarded as acceptable in 85% of the cases, like the non-modified forms, as opposed to 60% acceptance of modified upper-closed total adjectives (see also Bogal-Allbritten 2012). But more important, an eye-tracking study investigated the online processing of scale structure. This study further showed that the total first pass reading times of regions of sentences with maximum-standard adjectives modified by *slightly* were longer than those of similar regions with minimum-standard adjectives modified by *slightly*, suggesting that the processing of scales and endpoint standards is an obligatory part of semantic composition of phrases containing absolute adjectives, rather than a non-compulsory sort of late pragmatic processing.

Results reported in Syrett (2007) showed that adults treat absolute adjectives differently than relative adjectives. When presented with two rods of different lengths, but neither of which was obviously long, participants easily complied with a request to hand the experimenter “the long one,” suggesting that they can readily appeal to a contextually given comparison class to find a standard for membership that satisfies the existence presupposition associated with *the*. In contrast, when faced with two containers which were not full, participants tended to reject a request to hand the experimenter “the full one,” suggesting that they use a maximum standard of membership and are reluctant to shift this standard even in the presence of a presuppositional demand to this end. This and other work by Syrett and colleagues (e.g. Syrett et al. 2010; Syrett and Lidz 2010) further explored the differences between relative and absolute adjectives in the course of first language acquisition (for additional acquisition studies on related questions, see also e.g. Barner and Snedeker 2008; Tribushinina and Gillis 2012).

In the intervening years, additional experimental research—not only by psycholinguists but also by theoretical linguists and philosophers—has focused on the understanding of relative and absolute adjectives, including their interaction with degree modification (e.g. Schmidt et al. 2009; Solt 2011; McNabb 2012; Solt and Gotzner 2012; Liao and Meskin 2017; Liao et al. 2016; Solt 2016; Hansen and Chemla 2017), and their online processing (Aparicio et al. 2015). This research supports various

theoretical distinctions between relative and absolute adjectives that also manifest themselves in their online processing. Aparicio et al., this volume, contribute to this ongoing endeavor via an eye tracking study of minimum-standard adjectives and an offline study of the pragmatics of their use.

Finally, beyond psycholinguistic analyses that involve the notion of comparison (e.g. Scontras et al. 2012), research on the interpretation and processing of comparative constructions is starting to develop, especially in the domain of so-called “comparative illusions” (Wellwood et al. 2009, 2017; O’Connor 2015), but also more generally, as in Grant (2013). Tucker et al., this volume, continues this line of research.

3 The Chapters in This Volume

With this brief introduction in hand we now turn to the contributions of this volume to the state of the art concerning the offline and online study of gradability, scale structure and vagueness. We have not divided the volume into different thematically unified sections because, as the following overview of the book chapters reveals, we do not think they lend themselves easily to such a classical division. Rather, the volume has a family resemblance structure where each paper shares a slightly different set of properties with each other paper. Nonetheless, for the reader’s convenience we would like to indicate several different ways in which the papers may be grouped so as to help readers navigate in the volume according to the topics that most interest them.

First, the papers can be grouped according the lexical category of the items under investigation. Chapters 2, 5 and 6 are concerned with nouns, while Chaps. 3, 4, 7, 8, 10 and 11 investigate adjectives; Chap. 9 focuses on fractions which, while structurally nominal, functionally behave more like quantifiers.¹

Second, the papers can also be divided up according to the linguistic constructions that they address. While Chaps. 2, 5, 6 and 8 specifically examine the positive forms of adjectives and nouns, Chaps. 3, 10 and 11 discuss comparative constructions, and Chap. 9 deals with modifiers of fractions that are formally close to comparatives (e.g. *more than half*). Chapter 7 is concerned with (non-comparative) degree-modification of adjectives. Finally, Chap. 4 examines the effect of certain sorts of (positive form) adjectival modification on noun-based categorization.

Third, various subgroupings are possible based on the specific aspects of interest in the given items or constructions. While Chaps. 3–5, 10 and 11 address formal semantic aspects of words and constructions, Chaps. 2, 8 and 9 are concerned with general pragmatic factors, and Chaps. 6 and 7 specifically examine the effects of speaker group on linguistic behavior. Chapters 2–6 constitute an additional subgroup, as all are centered on issues related to the multidimensionality of concepts.

¹For research of the verbal domain, see, for example, Bochnak (2013b), Rappaport-Hovav (2014), and Fleischhauer (2016).

Yet another subgrouping is possible according to the language(s) studied: Most of the papers report studies of English, but Chap. 4 reports on German data, Chaps. 5, 6 and 11 present studies of Dutch (including Flemish), and Chap. 10 provides a contrastive study of Polish and English.

Finally, the chapters can be divided according to the methodologies used. Chapters 2, 3, 5–7, 9 and 11 use offline methods, while the remaining chapters use online methods, including measurement of response times (Chap. 10), eye movements (Chap. 8) and neural activity along the scalp (Chap. 4).

The volume begins with a study by **Paul Égré** and **Jérémy Zehr** (“Are gaps preferred to gluts? A closer look at borderline contradictions”) of seemingly contradictory assertions involving vague predicates. Vague adjectives, as noted above, admit borderline cases. One manifestation of borderline cases is that individuals can be ascribed both an adjective and its negation at the same time, as in e.g. *X is tall and not tall*, as well as neither the adjective nor its negation, as in *X is neither tall nor not tall* (Ripley 2011; Alxatib and Pelletier 2011; Serchuk et al. 2011; Égré et al. 2013). Égré and Zehr hypothesize that there is a preference for “gappy” descriptions (*neither A nor not A*) over “glutty” descriptions of the form *A and not A*. Though this hypothesis is supported by the results, they show that both kinds of descriptions are acceptable.

The analysis they propose for the data adopts the distinction offered by Cobreros et al. (2012) between strict and tolerant meanings for vague adjectives, and a specific implementation of the Strongest Meaning Hypothesis, in line with Alxatib and Pelletier (2011). However, in contrast to previous literature, Égré and Zehr argue in favor of local, rather than global, pragmatic accommodation of strict and tolerant truth operators. Assuming the strongest meaning of, e.g., *tall* to convey “tall by every standard,” *neither*-descriptions are consistent, while *and*-descriptions are not. A penalty is exerted on the acceptability of the latter due to the need to resort to a more tolerant interpretation such as “tall by some standard.” Given that evaluative multidimensional adjectives are associated with context dependent sets of dimensions (e.g. the adjectives *conservative* and *liberal* can relate to politics, religion, sex, family structure, dress code, music, and/or theoretical views), Égré and Zehr point out, in agreement with earlier observations by Kamp and Partee, that such adjectives are, intuitively, more acceptable in forms like *X is A and not A* than dimensional adjectives are, because they can be interpreted as, e.g., “X is A in some respects and not A in other respects.” More generally, they question whether the preference for *neither*-descriptions over *and*-descriptions is systematic, or whether it is likely to vary depending on the adjectival type (relative vs. absolute, or unidimensional vs. multidimensional).

The increased level of indeterminacy and contextual variance that multidimensional adjectives manifest is also exhibited in a higher acceptability of so-called faultless disagreements concerning their application. **Stephanie Solt**’s contribution (“Multidimensionality, subjectivity and scales: Experimental evidence”) reports on a study of precisely this phenomenon. Solt asked participants to decide whether only one of two speakers in disagreements such as (5) can be right, or whether both can be right, that is, whether the disagreement was a matter of fact or opinion.

- (5) A: Look Tommy’s shirt is dirtier than the one his little brother Billy is wearing.
 B: No, Billy’s shirt is dirtier than Tommy’s.

The results indicate that multidimensionality is indeed a source of subjectivity in comparative forms. However, Solt’s study also reveals that what it means to be multidimensional and what sorts of factors underlie disagreement is more complex than suggested by earlier work such as e.g. Sassoon (2013b), Lasersohn (2005), or Bylinina (2014). Multidimensional adjectives such as *good*, *intelligent*, or *beautiful* yielded different results not only from unidimensional adjectives such as *tall*, *old*, *expensive*, or *empty*, but also from adjectives such as *dirty*, *smooth*, *light*, or *sharp*. The first group clearly permitted faultless disagreements; the second group tends not to permit faultless disagreement at all. However, the third group yielded clearly mixed judgments as to whether a disagreement involving them would be a matter of fact or opinion—for example, we might disagree as to whether Tommy’s shirt is dirtier than Billy’s because we choose different ways of measuring dirtiness (e.g. overall presence of dirt vs. presence of a small amount of highly noticeable dirt), and depending on the choice of measure, the ordering of the shirts might be different. However, in some cases it may happen that all of the different choices of measure result in the same ordering, leading to the intuition that the disagreement is a matter of fact. Solt concludes that judge dependence is crucial to the first group (indeed, she argues that while the properties in question are clearly conceptually multidimensional, they do not always behave grammatically as if they were multidimensional), while in the case of the third group different dimensions can be distinguished, selected, and integrated in a contextually determined manner for the purposes of comparison.

The processing of adjectivally modified nouns such as *real diamond* or *fake diamond* also seems to depend on the identification and highlighting of a set of dimensions. However, in this case the dimensions consist of prototypical or stereotypical features of the head noun. Crucially, with *fake*-type adjectives, in contrast to *real*-type adjectives, these features of the meaning of the head noun are negated. Thus, the meaning of a nominal containing *fake* can be understood as, e.g., “in some sense x is N and in some sense x is not N;” that is, it seems to involve quantification over dimensions. Following Peirce (1910), and very much in line with Solt’s and Égré and Zehr’s observations, **Petra Schumacher, Patrick Brandt and Hanna Weiland-Breckle** propose in their contribution (“Online processing of *real* and *fake*: The cost of being too strong”) that this hidden meaning is the result of a repair that circumvents contradiction.

To test this hypothesis and its consequences for the neural signature of the processing of such modified nouns, Schumacher et al. measured ERPs of participants during the processing of *fake*-modified nouns, as compared to baselines formed by nouns modified by ordinary negative adjectives like *flawed*. They observed a Late Positivity, which is characteristic of referential shifts or reconceptualization—for example, it has been observed during processing of metonymic uses of noun phrases, as in *The ham sandwich paid*. Schumacher et al. argue that since *fake*-type modification involves an intermediate representation that is semantically contradictory, the Late

Positivity reflects an interface repair mechanism that deals with the contradiction. In contrast, processing of *real*-type adjectives, as compared with simpler baselines formed by ordinary positive adjectives like *white*, evoked no comparable processing costs. This finding aligns with the cost-free processing of e.g. *She read Dickens before she met him*, where different aspects of the meaning of *Dickens* are highlighted, but no reconceptualization occurs. Thus, the chapter locates the processing of *fake*- and *real*-type adjectives within a typology of the neural signatures of different types of semantic and pragmatic operations, explaining the processing differences as effects of recovery from inconsistent interpretations through dimension shifting.

Like Schumacher et al., **Hanna de Vries** (“Gradable nouns as concepts without prototypes”) is concerned with concepts expressed by nouns. However, her paper addresses a different issue. In a foundational paper, Kamp and Partee (1995) argued against the association of certain predicates with a prototype. For example, adjectives like *tall* or *intelligent* are associated with upper-open scales, and their meanings do not seem to be represented correctly by means of any prototypical values on those scales: There is no upper bound such that higher degrees of height decrease an entity’s tallness. Rather, the taller one is, the better. These adjectives are also vague and gradable. In contrast, nouns such as *bird* are associated with a prototype, but are not gradable in the same way that adjectives and adverbs are.

Considering this typology, de Vries argues that nouns like *genius*, which are also associated with upper-open scales (e.g. intelligence), do not have a prototype representation, either. She starts by characterizing prototypicality in terms of “maximal embodiment,” that is, in terms of manifesting, in the case of gradable properties, ideal values for those properties. Maximal embodiment cannot be satisfied in cases where (1) having more of a property is considered better, and (2) there is no maximal value of the property in question. Thus, if subjects appear to have prototypicality judgments for concepts that do not satisfy maximal embodiment, like “genius,” these judgments must reflect other factors.

De Vries tests the hypothesized difference between nouns like *genius*, which are associated with upper open scales, and nouns like *bird*, which are associated with upper closed prototypicality scales, using classical methodologies developed within the cognitive psychological research on conceptual structure. One experiment shows that factors like familiarity and, especially, attitude largely explain the prototypicality judgments in *genius*-type nouns, but are unrelated to prototypicality in *bird*-type nouns. The other experiments look at a related question: Do unbounded properties in fact play a greater role than bounded properties in subjects’ decisions about categorization with nouns like *genius*, as opposed to nouns like *bird*? Subjects were asked to generate properties for a range of *genius*- and *bird*-type nouns. De Vries then measures to what extent membership in the class described by each type of noun is linked with unbounded property dimensions (for example, the more intelligent one is, the more likely they are to be classified as a genius). She finds a strong tendency towards the use of unbounded dimensions for nouns like *genius*, but not for nouns like *bird*.

The contribution by **Steven Verheyen** and **Gert Storms** (“Education as a source of vagueness in criteria and degree”) examines yet another factor that plays a role

in how we understand and navigate the dimensions and boundaries associated with nominal predicates. Building on previous studies that suggest that upbringing plays a role in categorization behavior, Verheyen and Storms look specifically at level of education—i.e. whether individuals have completed *only* compulsory education or *also* higher education—as a factor in the classification behavior of individuals with different levels of education for categories such as fruits, vegetables, fish, insects, sports, sciences, tools, and furniture.

Verheyen and Storms’s study starts with a distinction proposed in Devos (1995, 2003) between vagueness in criteria and vagueness in degree. In the former, there is indeterminacy with respect to the conditions of application of the predicate to the noun. For instance, is chess a sport? There could be disagreement depending on whether the relevant criteria include physical activity or competition. In the latter, there is indeterminacy with respect to the extent of application given fixed conditions. Consider hiking, for example. While we can be certain that it meets the criterion of physical activity, we could argue about whether it meets this criterion sufficiently. Devos suggested that vagueness in criteria is primarily involved in categorization involving nouns, while vagueness in degree is mainly involved in categorization involving adjectives. Verheyen and Storms challenge this idea and implement a mathematical model capable of measuring both factors. The results of this study show that (1) both vagueness in criteria and vagueness in degree are found in nouns, and (2) criteria and degree differences are systematically related to subjects’ properties, such as their level of education. Compared to subjects with only compulsory education, subjects with higher education endorse fewer items and use different conditions of application, especially in nouns they are more familiar with, such as those related to science categories.

Social criteria also play a role in the data explored by **Andrea Beltrama** (“Intensification, gradability and social perception: The case of *totally*”). Grammatical distinctions in gradable adjectives—i.e. whether the standard for ascribing the adjective is a minimum, a maximum, or is contextually determined—typically condition the acceptability of degree modifiers. For instance, whereas *very* can modify relative adjectives such as *tall*, *completely* only targets absolute adjectives such as *full* or *empty*. However, such restrictions cease to apply when such modifiers target scales that are grounded in the attitude of the speaker, rather than in the lexical meaning of the subsequent adjective. It is in these contexts that such intensifiers are perceived as having an especially salient social meaning. Beltrama’s chapter explores precisely this phenomenon. It falls within the new domain of what we might call experimental socio-semantics, in which social meaning is assumed to be amenable to systematic and formal scrutiny, and compositional meaning is taken to affect an “expression’s suitability to serve as a vehicle for social meaning.”

More specifically, the chapter presents a study of the social meaning conveyed by the degree expression *totally*, whose interpretation varies depending on, roughly, whether the constituent it modifies is a closed-scale adjective or not. If the adjective is associated with a closed scale, *totally* is understood lexically, i.e. as entailing a maximal degree. If the adjective is not closed-scale, *totally* is understood as a “speaker-oriented” intensifier that does not entail maximal degree. Beltrama tested

the reactions of participants (concerning factors such as solidarity and status) when presented with the two types of *totally* as modifiers of different constituents, and also in comparison to other intensifiers (*completely*, *really*) and unmodified adjectives. The results support the author's hypothesis that *totally* is more likely to be interpreted as a carrier of social meaning on its speaker-oriented interpretation than on its lexical interpretation. Beltrama further discusses the social meaning carried by *totally* when modifying extreme adjectives such as *awesome*. The inherent emotive meaning conveyed by such adjectives is considered as a potential factor in explaining their unpredicted behavior, as reflected in the experimental results.

Beltrama's study points to the salience of the distinction between relative and absolute adjectives for (socio) linguistic phenomena. **Helena Aparicio**, **Ming Xiang** and **Christopher Kennedy** ("Informativity and grammar in referential effects of contrast involving adjectivally modified NPs") consider this distinction in relation to language processing. Their work builds on seminal psycholinguistic research on the interpretation of gradable adjectives by Julie Sedivy and collaborators (Sedivy et al. 1999; Sedivy 2003, 2005) which, in a series of Visual World eye-tracking studies, investigated the effect that contextual information has on incremental semantic processing. Participants in their experiments received verbal instructions such as "pick up the tall glass" while looking at displays of four objects. The instructions contained a restrictive pronominal adjective, which can trigger quantity-based pragmatic reasoning about a set of referents that contrast along the adjectival dimension. Each instruction was tested against two types of displays that either supported a contrastive interpretation of the adjective by including a contrastive element in the display (i.e. an object that could be described by the head noun in the instruction but not the adjective, e.g., a short glass), or lacked such contrastive object, rendering the use of the adjective redundant. Their results showed that the presence of a contrasting object in the visual display facilitated the lexical processing of the adjective, as revealed by the fact that the target object was identified significantly faster (even before information about the head noun was available to the participants) in those displays that contained a contrastive set of objects compared to those that did not. These results suggest that semantic processing is incremental and that the processing of attributive relative adjectives like *tall* is facilitated when the visual context supports a restrictive interpretation of the predicate—what is often called the Referential Effect of Contrast (REC).

Aparicio et al. aim to test whether pragmatic reasoning alone is sufficient to explain RECs, or whether lexical properties of the different classes of adjectives (essentially, whether the threshold that establishes the positive form of the adjective is an endpoint of the scale vs. determined through an extensional comparison class of individuals) also contribute to these effects. With this goal in mind, they carried out two experiments. The first is an extension of the Visual World study reported in Aparicio et al. (2015), in which the authors tested color, relative and Maximum Standard Absolute adjectives, to Minimum Standard Absolute Adjectives (MinAAs, such as *spotted*, *bent*, *striped*). The results show that unlike relative, color or Maximum Standard Absolute Adjectives (see Aparicio et al. 2015), MinAAs do not exhibit RECs.

The second experiment addresses the question of how informative the different classes of adjectives tested by Aparicio et al. (2015) and the eye-tracking study reported in this volume are perceived to be when used restrictively versus redundantly. In this study, participants rated the informativity of the instructions used in the eyetracking studies given the two displays (i.e. contrastive and non-contrastive) tested. The results indicate that color adjectives, relative adjectives and Maximum Standard Absolute Adjectives show a difference in ratings such that redundant uses were perceived as too informative compared to restrictive uses, whereas ratings pertaining to MinAAs did not show any difference between conditions. Putting together these experimental results and the work in Aparicio et al. (2015), a correlation is found between giving rise to RECs and penalizing overspecification. However, it is argued that informativity alone cannot account for the different properties of all the RECs reported in Aparicio et al. (2015). The authors conclude that even though informativity is clearly an important driver of RECs, the lexical semantics of the adjective classes also contributes to further shape RECs, a result that reveals interesting connections between scale structure, contrast effects and informativity that are worthy of future investigation and theorizing.

The remaining three papers in the volume take us away from lexical semantics, each presenting studies that focus on a different formal semantic and pragmatic property that has been associated with gradability. In “Modified fractions, granularity and scale structure,” **Chris Cummins** further explores Krifka’s (2002) hypothesis that the granularity of the scalar alternatives associated with an expression influence the pragmatic inferences that hearers draw. Krifka observed that when one hears, for example, *There were 81 people at the meeting*, one is likely to infer that exactly 81 people were there (arguably via a standard scalar implicature), but if one hears that there were 80 people at the meeting, the coarser granularity of a scale in tens rather than in units facilitates more approximative interpretations (and, arguably, influences related scalar inferences). Cummins et al. (2012) tested this hypothesis on the domain of modified numerals, showing that a sentence like *There’s room for more than 80 people* yields an upper bound inference of something like *not more than 100*, even though such upper bound inferences do not arise with modified numerals of finer granularity (for instance, *There’s room for more than 81 people* does not yield the inference that there is not room for more than 82). Cummins’ study in this volume extends the exploration of this phenomenon to modified fractions such as *more than one third* or *less than five sixths*.

The chapter reports on a series of judgment studies in which subjects were presented with sentences describing quantities in terms of modified fractions with different numerators and denominators (thirds, fourths, fifths, and so on). For each modified fraction, subjects had to freely supply the numerical percentage range they thought the fraction corresponded to (e.g. they might provide 35–60% for *more than one third*). Different experiments tested specific questions, such as whether making particularly salient a particular fractional scale (e.g. fifths), and thus a particular set of scalar alternatives, influenced the upper bound that subjects offered. Cummins’ results suggest that, rather than granularity being an explanation for inference patterns in and of itself, it is the pragmatically salient alternatives—no matter what their

granularity—that influence scalar inference. Though numerical expressions at a finer level of granularity generally make salient a different set of alternatives than do those at a coarser level of granularity, in the case of fractions, the situation proves to be more complex: some fractions (such as quarters and tenths) are highly salient even in cases that are unexpected based on a direct extension of the analysis of granularity effects for non-fractional expressions. Cummins’ work underscores not only the key role of scalar alternatives in interpretation, but also the complexity involved in determining exactly what these alternatives are in any given case.

Another ongoing issue in the formal semantic representation of gradable predicates involves the relation between polar opposites, including comparative morphology itself (*more A/A-er...than...*, *less A...than...*). Büring (2007) argues that the semantics of adjectives like *short* should be characterized in terms of decomposition into the semantics for *tall* plus the semantics for an abstract morpheme paraphraseable as *little*, but e.g. Heim (2008) has pointed out problems for such an analysis. In “Decomposition and processing of negative adjectival comparatives,” **Daniel Tucker, Barbara Tomaszewicz, and Alexis Wellwood** contribute to this debate by bringing processing data to bear on the question. Their experiments take as a premise the Interface Transparency Thesis (Lidz et al. 2011), according to which “the verification procedures employed in understanding a declarative sentence are biased towards algorithms that directly compute the relations and operations expressed by the semantic representation of that sentence.” This thesis leads to the prediction that if the semantic representation of an adjective or a comparative involves decomposition (as would be the case of *short(er)* on Büring’s analysis), it should take longer to process than the base (e.g. *tall(er)*). The authors were also interested in testing whether there were any differences in the processing of comparative statements if the statement was presented in mathematical notation (e.g. $A < B$), in which the comparative morpheme is arguably not decomposable, instead of natural language (e.g. *A is shorter than B*).

Tucker et al. designed a picture-matching task in which subjects were given a comparative statement in either natural language or mathematical notation and had to decide whether the statement truthfully described an image presenting the two compared objects, or not. Variants of the experiment were done in English and in Polish. The reaction times recorded on the task reveal that processing sentences with negative polar comparatives (e.g. *shorter than*) was systematically slower than the processing of positive comparatives, lending initial support to the decomposition hypothesis. However, the same effect was registered when the comparative was expressed in mathematical notation (i.e., $A < B$ took longer to process than $A > B$), in contrast to the results reported in Deschamps et al. (2015), where no such difference was found. Tucker et al.’s chapter thus leaves open the question as to whether decomposition is supported. They hypothesize that mathematical notation might show the same effect due to translation into natural language during processing; if that is not the case, then an alternative explanation for the slower processing of negative polar adjectives is probably called for.

Perhaps one of the biggest unresolved questions in the analysis of comparatives and, indeed, all gradable expressions, involves the status of degrees. If degrees are

crucial to an analysis of gradability in language and thus have a place in natural language metaphysics, what are they like? **Rick Nouwen** and **Jakub Dotlačil** (“Plural comparison?”) focus specifically on the possibility that not only do degrees constitute a subsort of entity, but moreover the domain of degrees has the same mereological structure as other (sub)domains of entities. On this view, thus, there are pluralities of degrees.

Nouwen and Dotlačil suggest that positing pluralities of degrees could help overcome problems faced by the analysis of sentences like *The participants typed faster than each of them wrote*, where a plural or quantified expression appears in the *than*-clause of a comparative. Specifically, in some cases such sentences appear to involve quantifier raising out of the *than*-clause (so that, for example, the just mentioned sentence could be understood as equivalent to “Each participant’s writing speed is such that all of the participants typed faster than that speed”), despite the fact that the general theory of quantifier scope does not predict *than*-clause-internal quantifiers to be able to raise. As an alternative, they suggest that the reading can be explained if the *than*-clause introduces a plurality of degrees, and comparison is cumulative. To test this hypothesis, they carried out a judgment task in which subjects had to decide whether a given sentence truthfully described a particular situation of plural comparison, presented in the form of a graph. Their results are better explained by a semantics that includes the possibility of pluralities of degrees and cumulative comparison than one in which such an option is not available.

Acknowledgements Most of the chapters in this volume are based on work presented at the workshop “Gradability, Scale Structure, and Vagueness: Experimental Perspectives,” held at CSIC’s Centro de Ciencias Humanas y Sociales (CCHS) in Madrid, in May 2015; two additional contributions were invited. We thank the members of the Linguistics and Cognitive Science Group (LyCC), led by Violeta Demonte, for their support in the organization of the workshop. We are also grateful to our contributors and to the other participants in the workshop for the lively discussion we had there. We also acknowledge the various sources of funding that helped underwrite the workshop and the preparation of this volume. Castroviejo and Sassoon were partially supported by grant FFI2012-34170, from the Ministry of Economy and Competitiveness (MINECO). Castroviejo also acknowledges grant FFI2015-66732-P (Ministry of Economy, Industry and Competitiveness, and the European Regional Development Fund, FEDER, EU), the IT769-13 Research Group (Basque Government), and UFI11/14 (University of the Basque Country). McNally acknowledges Spanish grants FFI2013-41301-P (MINECO) and FFI2016-76045-P (AEI/MINEICO/FEDER, UE), AGAUR (Catalan government) grant 2014SGR698, an ICREA Foundation Acadèmia award, and the Department of Translation and Language Sciences at Universitat Pompeu Fabra. Sassoon acknowledges the support of the German-Israeli Foundation for Scientific Research and Development (GIF). Thanks also to our abstract and manuscript reviewers, to Helen van der Stelt and Jolanda Voogd at Springer, and especially to Chungmin Lee for his initial encouragement to produce this volume.

References

- Alxatib, S., & Pelletier, F. J. (2011). The psychology of vagueness: Borderline cases and contradictions. *Mind and Language*, 26(3), 287–326.
- Anderson, C., & Morzycki, M. (2015). Degrees as kinds. *Natural Language and Linguistic Theory*, 33(3), 791–828.
- Aparicio, H., Xiang, M., & Kennedy, C. (2015). Processing gradable adjectives in context: A visual world study. In S. D'Antonio, M. Moroney, & C. Little (Eds.), *Proceedings of Semantics and Linguistic Theory* (Vol. 25, pp. 413–432).
- Aparicio, H., Xiang, M., & Kennedy, C. (2018). Informativity and grammar in referential effects of contrast involving adjectivally modified NPs. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Bale, A. (2008). A universal scale of comparison. *Linguistics and Philosophy*, 31(1), 1–55.
- Barner, D., & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret *tall* and *short* based on the size distributions of novel noun referents. *Child Development*, 79(3), 594–608.
- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 29–101). East Sussex, UK: Lawrence Erlbaum Associates.
- Bartsch, R. (1984). The structure of word meanings: Polysemy, metaphor, metonymy. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 25–54). Dordrecht: Foris.
- Bartsch, R. (1986). Context-dependent interpretations of lexical items. In J. Groenendijk, D. de Jongh, & M. Stokhof (Eds.), *Foundations of pragmatics and lexical semantics*, GRASS 7 (pp. 1–26). Dordrecht: Foris.
- Bartsch, R., & Vennemann, T. (1972). The grammar of relative adjectives and comparison. *Linguistische Berichte*, 20, 19–32.
- Beck, S., Krasnikova, S., Fleischer, D., Gergel, R., Hofstetter, S., Savelsberg, C., et al. (2009). Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook*, 9(1), 1–66.
- Beck, S., Oda, T., & Sugisaki, K. (2004). Parametric variation in the semantics of comparison: Japanese versus English. *Journal of East Asian Linguistics*, 13(4), 289–344.
- Beltrama, A. (2016). *Bridging the gap: Intensifiers between semantic and social meaning*. Dissertation, University of Chicago.
- Beltrama, A. (2018a). Totally between discourse and subjectivity: Exploring the pragmatic side of intensification. *Journal of Semantics* 35(2), 219–267.
- Beltrama, A. (2018b). Intensification, gradability and social perception: The case of totally. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Beltrama, A., & Bochnak, M. (2015). Intensification without degrees cross-linguistically. *Natural Language and Linguistic Theory*, 33(3), 843–879.
- Beltrama, A., & Staum Casasanto, L. (2017). *Totally* tall sounds *totally* younger: Intensification at the socio-semantics interface. *Journal of Sociolinguistics*, 21(2), 154–182.
- Beltrama, A., & Xiang, M. (2012). Is good better than excellent? An experimental investigation on scalar implicatures and gradable adjectives. In *Proceedings of Sinn und Bedeutung* (Vol. 17, pp. 81–98).
- van Benthem, J. (1982). Later than late: On the logical origin of the temporal order. *Pacific Philosophical Quarterly*, 63, 193–203.
- Bierwisch, M. (1989). The semantics of gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional adjectives* (pp. 71–261). Berlin: Springer.
- Bobaljik, J. D. (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.
- Bochnak, M. (2013a). *Cross-linguistic variation in the semantics of comparatives*. Dissertation, University of Chicago.

- Bochnak, M. (2013b). Two sources of scalarity within the verb phrase. In B. Arsenijević, B. Gehrke, & R. Marín, (Eds.), *Studies in the composition and decomposition of event predicates* (pp. 99–123). Dordrecht: Springer.
- Bogal-Allbritten, E. (2012). Slightly coerced: Processing evidence for adjectival coercion by minimizers. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 48.1, pp. 77–92). Chicago: Chicago Linguistic Society.
- Bolinger, D. (1972). *Degree words*. The Hague: Mouton.
- Büring, D. (2007). Cross-polar nomalies. In T. Friedman & M. Gibson (Eds.), *Proceedings of Semantics and Linguistic Theory* (Vol. 17, pp. 37–52). Ithaca, NY: Cornell University.
- Burnett, H. (2016). *Gradability in natural language: Logical and grammatical foundations*. Oxford: Oxford University Press.
- Bylina, L. (2014). *The grammar of standards: Judge-dependence, purpose-relativity, and comparison classes in degree constructions*. Dissertation, Utrecht University.
- Cappelletti, M., Fregni, F., Shapiro, K., Pascual-Leone, A., & Caramazza, A. (2008). Processing nouns and verbs in the left frontal cortex: A transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 20(4), 707–720.
- Cobrerros, P. (2011a). Paraconsistent vagueness: A positive argument. *Synthese*, 183(2), 211–227.
- Cobrerros, P. (2011b). Supervaluationism and classical logic. In R. Nouwen, R. van Rooij, H.-C. Schmitz, & U. Sauerland (Eds.), *Vagueness in communication*. Lecture Notes in Computer Science (Vol. 6517, pp. 51–63). Berlin: Springer.
- Cobrerros, P., Égré, P., Ripley, D., & van Rooij, R. (2012). Tolerance and mixed consequence in the supervaluationist setting. *Studia Logica*, 100(4), 855–877.
- Constantinescu, C. (2011). *Gradability in the nominal domain*. Dissertation, Leiden University.
- Cresswell, M. (1976). The semantics of degree. In B. Partee (Ed.), *Montague Grammar* (pp. 261–292). New York: Academic Press.
- Cummins, C. (2018). Modified fractions, granularity and scale structure. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Cummins, C., Sauerland, U., & Solt, S. (2012). Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy*, 35(2), 135–169.
- Deschamps, I., Agmon, G., Lewenstein, Y., & Grodzinsky, Y. (2015). The processing of polar quantifiers, and numerosity perception. *Cognition*, 143(2), 115–128.
- Devos, F. (1995). Still fuzzy after all these years. A linguistic evaluation of the fuzzy set approach to semantic vagueness. *Quaderni di Semantica*, 16(31–32), 47–82.
- Devos, F. (2003). Semantic vagueness and lexical polyvalence. *Studia Linguistica*, 57(3), 121–141.
- Doran, R., Baker, R. E., McNabb, Y., Larson, M., & Ward, G. (2009). On the non-unified nature of scalar implicature: An empirical investigation. *International Review of Pragmatics*, 1(2), 211–248.
- Égré, P., & Zehr, J. (2018). Are gaps preferred to gluts? A closer look at borderline contradictions. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Égré, P., De Gardelle, V., & Ripley, D. (2013). Vagueness and order effects in color categorization. *Journal of Logic, Language and Information*, 22(4), 391–420.
- Fara, D. G. (2000). Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28(1), 45–81.
- Fine, K. (1975). Vagueness, truth and logic. *Synthese*, 30(3), 265–300.
- Fleischhauer, J. (2016). *Degree gradation of verbs*. Düsseldorf: Düsseldorf University Press.
- Frazier, L., Clifton, C., & Stolterfoht, B. (2008). Scale structure: Processing minimum standard and maximum standard scalar adjectives. *Cognition*, 106(1), 299–324.
- Grant, M. A. (2013). *The parsing and interpretation of comparatives: More than meets the eye*. Dissertation, University of Massachusetts, Amherst.
- Grosu, A., & Landman, F. (1998). Strange relatives of the third kind. *Natural Language Semantics*, 6(2), 125–170.

- Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 441–461.
- Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition*, 65(2), 137–165.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31(3), 355–384.
- Hansen, N., & Chemla, E. (2017). Color adjectives, standards, and thresholds: An experimental investigation. *Linguistics and Philosophy*, 40(3), 239–278.
- Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in degree achievements. In T. Matthews & D. Strolovitch (Eds.), *Proceedings of Semantics and Linguistic Theory* (Vol. 9, pp. 127–144). Ithaca, NY: Cornell University.
- Heim, I. (1985). *Notes on comparatives and related matters*. Unpublished manuscript, University of Texas.
- Heim, I. (2008). Decomposing antonyms. In *Proceedings of Sinn und Bedeutung* (Vol. 12, pp. 212–225).
- Hyde, D. (1997). From heaps and gaps to heaps of gluts. *Mind*, 106(424), 641–660.
- Kamp, H. (2013). The paradox of the heap. In K. von Heusinger & A. G. B. Ter Meulen (Eds.), *Meaning and the dynamics of interpretation. Selected writings of Hans Kamp* (pp. 263–319). Leiden: Brill.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57(2), 129–191.
- Keefe, R. (2000). *Theories of vagueness*. Cambridge: Cambridge University Press.
- Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland Press.
- Kennedy, C. (2007a). Modes of comparison. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 43, pp. 141–165). Chicago: Chicago Linguistic Society.
- Kennedy, C. (2007b). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1), 1–45.
- Kennedy, C. (2012). The composition of incremental change. In V. Demonte & L. McNally (Eds.), *Telicity, change, state: A cross-categorical view of event structure* (pp. 103–121). Oxford: Oxford University Press.
- Kennedy, C. (2013). Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry*, 56(2–3), 258–277.
- Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics and discourse* (pp. 156–182). Oxford: Oxford University Press.
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2), 345–381.
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4(1), 1–45.
- Krifka, M. (2002). Be brief and vague! And how Bidirectional Optimality Theory allows for verbosity and precision. In D. Restle & D. Zaefferer (Eds.), *Sounds and systems: Studies in structure and change: A Festschrift for Theo Vennemann* (pp. 439–458). Berlin: Mouton de Gruyter.
- Laserson, P. (1999). Pragmatic halos. *Language*, 75(3), 522–551.
- Laserson, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, 28(6), 643–686.
- Liao, S.-Y., & Meskin, A. (2017). Aesthetic adjectives: Experimental semantics and context-sensitivity. *Philosophy and Phenomenological Research*, 94(2), 371–398.
- Liao, S.-Y., McNally, L., & Meskin, A. (2016). Aesthetic adjectives lack uniform behavior. *Inquiry*, 59(6), 618–631.
- Lidz, J., Halberda, J., Pietroski, P., & Hunter, T. (2011). Interface transparency and the psychosemantics of *most*. *Natural Language Semantics*, 6(3), 227–256.

- Matushansky, O. (2002). A beauty of a construction. In L. Mikkelsen & C. Potts (Eds.), *Proceedings of the 21st West Coast Conference on Formal Linguistics* (pp. 264–277), Somerville, MA: Cascadilla Press.
- McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition*, 6(4), 462–472.
- McNabb, Y. (2012). *The syntax and semantics of degree modification*. Dissertation, University of Chicago.
- McNally, L. (2011). The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In R. Nouwen, R. van Rooij, U. Sauerland, & H.-C. Schmitz (Eds.), *Vagueness in communication*. Lecture Notes in Computer Science (Vol. 6517, pp. 151–168). Berlin, Heidelberg: Springer.
- McNally, L. (2017). Scalar alternatives and scalar inference involving adjectives: A comment on van Tiel, et al. 2016. In J. Ostrove, R. Kramer, & J. Sabbagh (Eds.), *Asking the right questions: Essays in honor of Sandra Chung* (pp. 17–28). <https://escholarship.org/uc/item/8255v8sc>.
- McNally, L., & Stojanovic, I. (2017). Aesthetic adjectives. In J. Young (Ed.), *The semantics of aesthetic judgments* (pp. 17–37). Oxford: Oxford University Press.
- Moltmann, F. (2004). Properties and kinds of tropes: New linguistic facts and old philosophical insights. *Mind*, 113(449), 1–41.
- Moltmann, F. (2007). Events, tropes, and truthmaking. *Philosophical Studies*, 134(3), 363–403.
- Moltmann, F. (2009). Degree structure as trope structure: A trope-based analysis of positive and comparative adjectives. *Linguistics and Philosophy*, 32(1), 51–94.
- Morzycki, M. (2009). Degree modification of gradable nouns: Size adjectives and adnominal degree morphemes. *Natural Language Semantics*, 17(2), 175–203.
- Morzycki, M. (2011). Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics*, 19(1), 39–86.
- Nouwen, R., & Dotlačil, J. (2018). Cumulative comparison: Experimental evidence for degree cumulation. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- O'Connor, E. (2015). *Comparative illusions at the syntax-semantics interface*. Dissertation, University of Southern California.
- O'Connor, E., Pancheva, R., & Kaiser, E. (2012). Evidence for online repair of Escher sentences. In *Proceedings of Sinn und Bedeutung* (Vol. 17, pp. 363–380).
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1), 35–58.
- Pancheva, R., & Tomaszewicz, B. (2011). *Experimental evidence for the syntax of phrasal comparatives in Polish* (Vol. 17.1, pp. 185–194). University of Pennsylvania Working Papers in Linguistics.
- Peirce, C. S. (1910). *The Charles S. Peirce papers*. Cambridge, MA: Harvard University Press.
- Piñón, C. (2008). Aspectual composition with degrees. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics and discourse* (pp. 183–219). Oxford: Oxford University Press.
- Raffman, D. (1994). Vagueness without paradox. *The Philosophical Review*, 103(1), 41–74.
- Raffman, D. (2005). How to understand contextualism about vagueness: Reply to Stanley. *Analysis*, 65(3), 244–248.
- Rappaport-Hovav, M. (2014). Building scalar changes. In A. Alexiadou, H. Borer, & F. Schäfer, (Eds.), *The syntax of roots and the roots of syntax* (pp. 259–281). Oxford: Oxford University Press.
- Ripley, D. (2011). Contradictions at the borders. In R. Nouwen, R. van Rooij, U. Sauerland, & H.-C. Schmitz (Eds.), *Vagueness in communication*. Lecture Notes in Computer Science (Vol. 6517, pp. 169–188). Berlin: Springer.
- Rips, L. J., & Turnbull, W. (1980). How big is big? Relative and absolute properties in memory. *Cognition*, 8(2), 145–174.

- van Rooij, R. (2011). Vagueness and linguistics. In G. Ronzitti (Ed.), *Vagueness: A guide* (pp. 123–170). Berlin: Springer.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605.
- Rotstein, C., & Winter, Y. (2004). Total adjectives versus partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, 12(3), 259–288.
- Sassoon, G. W. (2012a). The double nature of negative antonymy. In A. Aguilar-Guevara, A. Chernilovskaya, & R. Nouwen (Eds.), *Proceedings of Sinn und Bedeutung* (Vol. 16.2, MIT Working Papers in Linguistics, pp. 543–556). Cambridge, MA: MIT Press.
- Sassoon, G. W. (2012b). A slightly modified economy principle: Stable properties have non-stable standards. In *Proceedings of the Israel Association of Theoretical Linguistics (IATL)* (Vol. 27, MIT Working Papers in Linguistics, pp. 163–181). Cambridge, MA: MIT Press.
- Sassoon, G. W. (2013a). A typology of multidimensional adjectives. *Journal of Semantics*, 30(3), 335–380.
- Sassoon, G. W. (2013b). *Vagueness, gradability and typicality: The interpretation of adjectives and nouns*. Leiden: Brill.
- Sassoon, G. W. (2017). Comparisons of nominal degrees. *Language*, 93(1), 153–188.
- Sassoon, G. W., & Fadlon, J. (2017). The role of dimensions in classification under predicates predicts their status in degree constructions. *Glossa: A Journal of General Linguistics*, 2(1), 42, 1–40.
- Sassoon, G. W., & Toledo, A. (2011). *Absolute and relative adjectives and their comparison classes*. Unpublished manuscript, University of Amsterdam and Utrecht University.
- Sassoon, G. W., Meir, N., Fadlon, J., Anaki, D., & Schumacher, P. (t.a.). *The acceptability, processing and neural signature of nominal gradability*. Unpublished manuscript, Bar Ilan University, The Hebrew University, Jerusalem, and University of Cologne.
- Schmidt, L. A., Goodman, N. D., Barner, D., & Tenenbaum, J. B. (2009). How tall is *tall*? Compositionality, statistics, and gradable adjectives. In *Proceedings of the Annual Conference of the Cognitive Science Society* (Vol. 31, pp. 2759–2764).
- Schumacher, P. B., Brandt, P., & Weiland-Breckle, H. (2018). Online processing of *real* and *fake*: The cost of being too strong. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Scontras, G. (2014). *The semantics of measurement*. Dissertation, Harvard University.
- Scontras, G., Graff, P., & Goodman, N. D. (2012). Comparing pluralities. *Cognition*, 123(1), 190–197.
- Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, 32(1), 3–23.
- Sedivy, J. C. (2005). Evaluating explanations for referential context effects: Evidence for Gricean mechanisms in online language interpretation. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 345–364). Cambridge, MA: MIT Press.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147.
- Serchuk, P., Hargreaves, I., & Zach, R. (2011). Vagueness, logic and use: Four experimental studies on vagueness. *Mind and Language*, 26(5), 540–573.
- Seuren, P. (1973). The comparative. In F. Kiefer & N. Ruwet (Eds.), *Generative Grammar in Europe* (pp. 528–564). Dordrecht: Reidel.
- Shapiro, S. (2006). *Vagueness in context*. Oxford: Oxford University Press.
- Smith, L., Cooney, N. J., & McCord, C. (1986). What is “high”? The development of reference points for “high” and “low”. *Child Development*, 57, 583–602.
- Smith, L., Rattermann, M. J., & Sera, M. (1988). “Higher” and “lower”: Comparative and categorical interpretations by children. *Cognitive Development*, 6, 131–145.
- Soames, S. (1999). *Understanding truth*. Oxford: Oxford University Press.
- Solt, S. (2011). How many *Mosts*? In *Proceedings of Sinn und Bedeutung* (Vol. 15, pp. 565–579).

- Solt, S. (2016). On measurement and quantification: The case of *most* and *more than half*. *Language*, 92(1), 65–100.
- Solt, S. (2018). Multidimensionality, subjectivity and scales: Experimental evidence. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Solt, S., & Gotzner, N. (2012). Who here is tall? Comparison classes, standards and scales. In *Pre-Proceedings of the International Conference Linguistic Evidence 2012* (pp. 79–83). Tübingen: Eberhard Karls Universität Tübingen.
- Stanley, J. (2003). Context, interest relativity and the Sorites. *Analysis*, 63(280), 269–281.
- von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics*, 3(1), 1–77.
- Syrett, K. (2007). *Learning about the structure of scales: Adverbial modification and the acquisition of the semantics of gradable adjectives*. Dissertation, Northwestern University.
- Syrett, K., & Lidz, J. (2010). 30-month-olds use the distribution and meaning of adverbs to interpret novel adjectives. *Language Learning and Development*, 6(4), 258–282.
- Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and context in children’s understanding of gradable adjectives. *Journal of Semantics*, 27(1), 1–35.
- van Tiel, B., van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics*, 33(1), 137–175.
- Tribushinina, E., & Gillis, S. (2012). The acquisition of scalar structures: Production of adjectives and degree markers by Dutch-speaking children and their caregivers. *Linguistics*, 50(2), 241–268.
- Tucker, D., Tomaszewicz, B., & Wellwood, A. (2018). Decomposition and processing of negative adjectival comparatives. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Umbach, C. (2016). Evaluative propositions and subjective judgments. In C. Meier & J. van Wijnbergen-Huitink (Eds.), *Subjective meaning: Alternatives to relativism* (pp. 127–168). Berlin: De Gruyter Mouton.
- Vanden Wyngaerd, G. (2001). Measuring events. *Language*, 77(1), 61–90.
- Verheyen, S., & Storms, G. (2018). Education as a source of vagueness in criteria and degree. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Verheyen, S., Hampton, J. A., & Storms, G. (2010). A probabilistic threshold model: Analyzing semantic categorization data with the Rasch model. *Acta Psychologica*, 135(2), 216–225.
- Verheyen, S., Dewil, S., & Égré, P. (2016). Subjective meaning in gradable adjectives: The case of ‘tall’ and ‘heavy’. Submitted.
- de Vries, H. (2018). Gradable nouns as concepts without prototypes. In E. Castroviejo, L. McNally, & G. W. Sassoon (Eds.), *Gradability, vagueness, and scale structure: Experimental perspectives*. Berlin: Springer.
- Wellwood, A. (2014). *Measuring predicates*. Dissertation, University of Maryland.
- Wellwood, A., Pancheva, R., Hacquard, V., Fults, S., & Phillips, C. (2009). The role of event comparison in comparative illusions. Poster Presented at the 22nd Annual CUNY Conference on Human Sentence Processing, Davis, CA.
- Wellwood, A., Pancheva, R., Hacquard, V., & Phillips, C. (2017). The anatomy of a comparative illusion. Submitted.
- Williamson, T. (1994). *Vagueness*. London: Routledge.