Glossa a journal of general linguistics

Sassoon, Galit Weidman and Julie Fadlon. 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, DOI: https://doi.org/10.5334/gjgl.155

RESEARCH

The role of dimensions in classification under predicates predicts their status in degree constructions

Galit Weidman Sassoon[1] and Julie Fadlon[2]

[1] Department of English Literature and Linguistics, Bar Ilan University, IL

[2] Department of Psychology, University of California, San Diego, US

Corresponding author: Galit Weidman Sassoon (galit.weidman-sassoon@biu.ac.il)

This paper presents an experimental study of multidimensional gradability across categories. The study tests whether and to what extent the naturalness of multidimensional adjectives and nouns in degree constructions is predictable from their conceptual-semantic properties – the way their dimensions are typically bound to create a unified interpretation. Past research suggests that binding by counting operations is common in multidimensional adjectives, possible in certain nouns (labeled *additive nouns*), and rare in others (dubbed *multiplicative nouns*; Hampton et al. 2009). We hypothesize that the higher naturalness of adjectives in degree constructions (Kennedy 1999) stems from a preference for dimension-binding by counting operations. Accordingly, we predict that additive nouns would be judged more natural than multiplicative nouns in any construction whose interpretation involves dimension-counting, e.g., dimensional quantifiers (as in: *in {some, most, every} respect}*) and degree modifiers (as in *more P {than y, than Q, than y is Q}*). The results of a naturalness survey involving 139 English speakers confirm our predictions. Moreover, our results indicate that the naturalness of a predicate in degree constructions and dimensional-quantifier constructions are tightly correlated, suggesting that dimension accessibility for counting is indeed an important predictor of morphological gradability.

Keywords: degree; adjective; noun; comparison; dimension; similarity

1 Dimension-binding as a predictor of morphological gradability

The interpretation of *multidimensional predicates* (e.g., *optimistic, linguist, chair, duck*) can be broken down into multiple building blocks called *dimensions*. For example, (1a) lists different ways in which individuals can be *optimistic*, and (1b, c) list characteristics that raise the degree to which entities exemplify concepts like *linguist* or *duck*, respectively (cf., Hampton et al. 2009).

(1) a. optimistic about their careers, finding love, friends, well-being, health, economical condition, studies, weather, peace …
b. affiliation in a linguistic department, familiarity with languages and language theories, language research, sensitivity to language, publications in linguistic journals, presentations in linguistic conferences, subscriptions to linguistic mailing lists …
c. duck's shape, color, behavior, genetic layout, inner biological function, offspring characterization, movement type, voice, environment …

The cognitive role of multiple dimensions was acknowledged by the founders of formal semantics (Wittgenstein 1953; Searle 1958; Kamp 1975; Bartsch 1986), but the idea that they affect truth conditions has been rejected (Lewis 1970; Kripke 1972; Putnam 1975; Kamp & Partee 1995). As a consequence, formal-semantic accounts of morphological gradability typically model all predicates in terms of a single dimension as is appropriate for *dimensional predicates* such as *tall* or *long* (Kennedy 1999; 2007; Rothstein & Winter 2005); e.g., *healthy* is modeled by a scale ordered along a **unique unanalyzed health dimension**. Entities' degrees on this scale are intuitively thought to be based on averaging over their degrees in multiple health measurements, but this fact is assumed not to be encoded in grammar.

Contrary to this view, theoretical and experimental research suggests that **sets of dimensions**, and **operations that bind them**, do play a role in the lexical representations of predicates and have an effect on the meaning and grammatical status of utterances (Kamp 1975; Klein 1980; Bartsch 1986; Landman 1989; Hoeksema 1995; van Rooij 2010; Grinsell 2013; Shamir 2013; Sassoon 2013a; b; Bylinina 2014; Solt 2018; Hampton & Winter in progress). This research aims to link formal semantics and cognitive psychological theories of classification and grading. The present paper reports an experimental examination of a theory that represents the outgrowth of this research program, namely *the Dimension-Accessibility theory*. According to this theory, multiple dimensions and their binding operations affect the interpretation and naturalness of the constructions listed in (2a, b).

(2) a. **Dimensional constructions**, which include modifiers of access to, and quantification over, dimensions of a predicate, e.g., *with respect to Q, in some respects*, or *in every way*. These modifiers are headed by respect-accessing prepositions such as *with respect to, by way of, in terms of, about,* or *in*.
b. **Degree constructions**, which include degree-heads like *more, very* and *perfectly*.

The core of the theory is *the Accessibility Hypothesis* in (3) (Sassoon 2017 in progress).

(3) a. **Dimension accessibility** is a high tendency to base exemplariness judgments and classification under a predicate on dimension-counting.
b. **The Accessibility Hypothesis:** Dimension accessibility is a strong predictor of the naturalness of a predicate in dimensional- and degree-constructions.

The paper is structured as follows. Section 1 motivates the Accessibility Hypothesis and presents its predictions with regard to the naturalness of quantification and degree constructions featuring multi-dimensional adjectives and nouns. Section 2 reports an extensive naturalness study involving 139 native English speakers designed to test these predictions. Finally, section 3 discusses the implications of our results for formal semantic models of dimensional- and degree-constructions and possible directions for future research.

1.1 The Accessibility Hypothesis: Rationale and motivation

Section 1.1.1 introduces readings of classification- and comparative forms of adjectives (e.g., *is optimistic; is more optimistic*) which involve counting and quantification over dimensions, suggesting that classification under adjectives can be based on dimension-counting. By contrast, typically, classification-forms of nouns do not have such readings, as classification in nouns is not usually based on dimension-

counting. Moreover, intuitively, nouns do not seem to be as natural as adjectives are in dimensional and comparative constructions, suggesting that dimension-counting is the factor that enables licensing in these constructions. This is the basic hypothesis that our study tests.

Section 1.1.2 introduces the concept of naturalness judgments in experimental syntax and its use in the present study to tap into dimensional and degree semantics. Then, section 1.1.3 introduces independent psychological evidence for the higher prevalence of dimension counting in classification under social nouns than natural-kind nouns. In section 1.2, these findings are used for the derivation of testable predictions of the Accessibility Hypothesis.

1.1.1 Dimension-counting as enabling licensing in dimensional and degree constructions

According to the Accessibility Hypothesis the dimensions of a multidimensional predicate are *accessible* when classification decisions and exemplariness ratings in that predicate are based on dimension-counting. For example, in a prominent reading, the classification-form in (4a) and comparison-form in (4b) are true if and only if (henceforth *iff*) the woman talked about is optimistic in enough respects, i.e., in as many respects as context requires, or in more respects than Bill is optimistic about, respectively.

(4) a. This woman is optimistic.
b. This woman is more optimistic than Bill.

That is, *optimistic* in such uses is associated with a scale representing the **number** of optimistic dimensions whose norm entities exceed. As judgments of exemplariness and classification can be and often are based on dimension-counting, the dimensions are *accessible* for binding by counting operations.

Moreover, the forms in (4a, b) have additional derived readings, which can be paraphrased using (5a–c), namely readings based on access to and quantification over dimensions.

(5) a. This woman is optimistic **about** finding love.
b. This woman is optimistic **in** {many ways, all/most respects}.
c. This woman is {generally, overall, perfectly, surprisingly, very, more} optimistic.

When discourse highlights the dimension of finding love, (4a, b) are understood to be true iff the woman in question is more optimistic about love than the norm or than Bill is, respectively, similarly to (5a). But when discourse does not highlight any particular dimension, (4a, b) can be understood as true iff in each one of some, many, surprisingly many (cf., Nouwen 2011), most, or all of the dimensions of *optimistic*, the woman is more optimistic about it than the norm or than Bill is, similarly to (5b, c) (Sassoon 2013a; 2017; in progress).

While not necessarily more prominent, dimension-counting readings seem to enable the derivation of readings relating to a single dimension or quantification over dimensions, and the licensing of examples like (5a–c) (see a detailed discussion in section 3.1). Briefly, assuming that regular adjectives do not lexicalize a respect argument (only a degree and an entity argument; cf. von Stechow 1984; 2009), a special operation, WRT ("with respect to"), the denotation of, e.g., *with respect to*, *about* or *in* (as in (5a, b)), must be used to introduce a dimension argument into the derivation.

The WRT operation poses restrictions on the value of the argument slot that it introduces. The set of dimensions of, e.g., *optimistic* must include the dimension accessed by *about* in *optimistic about love*, the dimensions quantified over in *optimistic in all respects*, and the dimensions involved in the interpretation of *generally, perfectly, very,* or *more optimistic*.

Moreover, Sassoon (in progress) argues that the WRT operation only accesses dimensions whose role in interpretation is to be **counted**; e.g., in *optimistic about finding love*, finding love is presupposed to be one of the dimensions that speakers **count** when deciding whether and to what extent entities are optimistic, one of the dimensions generating the **dimension-counting measurement** associated with *optimistic*. And since a quantifier can bind a respect argument only when WRT is licensed, as in *in some/all respects*, dimensional quantifiers seem to only bind the dimensions of predicates whose interpretation is based on dimension-counting.

Evidence for these claims (namely for the role of dimension-counting in classification and comparative forms like (4a, b) and in dimensional constructions like (5a–c)) comes from past corpus research and judgment studies. Evidence suggests that readings based on dimension counting and quantification are readily available in adjectives, as we just illustrated, but they are harder to get in classification forms with nouns, where psychological research points against classification based on dimension-counting.

Starting with dimension-counting in adjectives, Sassoon (2012), for example, presented 50 participants with contexts in which the health of two individuals, Dan and Sam, was measured by three medical blood tests of health dimensions like pneumonia, flu and chickenpox or blood pressure, cholesterol, and sugar level indicating diabetics. Dan's average score on the tests was higher than Sam's, but Sam was healthier in more respects than Dan. Sam's levels in all tests were within the norm, while Dan's level in one test was not.

In about 75% of the cases, participants judged (6a) to be true and (6b) to be false. Thus, for the most part, in judging the truth of these comparisons, participants seemed to disregard averaged health levels and to merely count dimensions. In addition, participants considered (6c) true and (6d) false in about 57% of the contexts, considering *healthy*'s membership criterion to require having a maximal amount of respects (having no disease), but not a maximal or high averaged health. Other participants, who considered (6d) true, often indicated that being healthy in most (two out of three) respects is sufficient, suggesting that their judgments involved dimension counting as well.

(6) a. Sam is healthier than Dan.
b. Dan is healthier than Sam.
c. Sam is healthy.
d. Dan is healthy.

The following naturally occurring example invites a dimension-counting interpretation (Sassoon in progress). It concludes the discussion of multiple positive effects of technology on high education (e.g., on its accessibility, interaction, curriculum, instruction, and university ranking) and negative effects (on inequality, inequity, minorities' representation, and financing). A dimension-counting interpretation renders the use of *sum up* in (7) quite literal (for theoretical arguments for dimension-counting comparisons see Sassoon in progress).

(7) To sum up, my vision of higher education in 2020 is more optimistic than pessimistic.[1]

[1] http://lrs.ed.uiuc.edu/students/cwang2/vision.html.

Additional studies of dimension counting and quantification in adjectival antonyms as opposed to nouns used an indirect methodology, exploiting the fact that exception phrases indicate universal generalizations (which, unlike existence statements, can have exceptions). This fact is manifested in the higher naturalness of (8a) compared to (8c) (von Fintel 1994; Hoeksema 1995; Moltmann 1995; Fox & Hackl 2006). Moreover, as logical theories predict, negated existential quantifiers are universal, thus (8b) is natural, whereas negated universal quantifiers are existential, thus (8d) is not natural.

(8) a. Everyone is happy except for Dan.
b. No one is happy except for Dan.
c. #Someone is happy except for Dan.
d. #Not everyone is happy except for Dan.

Accordingly, the exception phrases in (9) seem to operate on implicit quantifiers over adjectival respects, thus revealing the involvement of counting and quantification in classification under adjectives. Moreover, judgment surveys (Sassoon 2012; Shamir 2013) reveal that participants tend to rate exception phrases as more natural with adjectives like *healthy* whose dimensions are intuitively combined by an implicit universal quantifier ALL (*healthy* intuitively means healthy in all respects) as in (9a), than with their antonyms, whose dimensions are combined by an existential quantifier SOME, as in in (9c), where *sick* intuitively means sick in some respect (Hoeksema 1995). And the effect is reversed in negated adjectives; exception phrases are rated as more natural with negated existential adjectives than universal ones, as in (9b, d).

(9) a. I am **healthy** except for high blood pressure (bp) (∀F≠bp, I'm healthy in F).
b. He is **not sick** except for the flu (¬∃F≠flu, He's sick in F/∀F≠flu, He's healthy in F).
c. #I am **sick** except for normative blood pressure (#∃F≠bp, I'm sick in F).
d. #I am **not healthy** except for (normal) cholesterol (#∃F≠ch, ¬(I'm healthy in F)).

These judgments are also supported by distributional patterns observed in corpora. One study examined a corpus of 1300 naturally occurring examples of the form "Adj. except" with 8 antonym pairs in positive vs. negated contexts (Sassoon 2013b). The frequency of co-occurrence of adjectives with dimensional exception phrases depended both on adjective polarity (positive vs. negative) and context (existence or absence of negation), which interacted as predicted.

These studies suggest that quantification over dimensions is general among adjectives, and that, moreover, the positive or negative polarity of antonyms is systematically related to the force of quantifier over their dimensions. The dimensions of positive adjectives often tend to be bound by a universal quantifier, ALL or NO, while those of negative antonyms tend to be bound by an existential quantifier, SOME. That is, multidimensional adjectives tend to convey generalizations about dimensions, while multidimensional antonyms tend to relate to counterexamples to such generalizations.

Turning to nouns, Shamir (2013) compared adjectival antonyms to concrete nouns like *duck* in (10a), showing that such nouns are less natural with dimensional exception phrases than either positive or negative adjectives in both negated and non-negated contexts (cf. *#bird/not a bird, except for flying/ size/ wings*), in line with dimension-binding by similarity operations, instead of quantifiers, as explained shortly.

Furthermore, intuitively, examples (10b)–(11) are odd.

(10) a. This bird is a duck.
b. #This bird is more a duck than that one.

(11) a. #This bird is a duck in the color of his feathers, but not in flying.
b. #This bird is a duck {in many ways, in all/most respects}.
c. #This bird is {generally, overall, perfectly, very, more} a duck.

Again, Sassoon (2017; in progress) has linked this fact with the independently motivated finding that, in ordinary contexts, classification and exemplariness judgments in predicates like *duck* are not based on the mere counting of dimensions. First, the dimensions are not equally important (e.g., *duck-beak* is more important for ducks than *flying*). Second, their membership norms are usually irrelevant for classification. For each dimension, even in contexts in which it is considered perfectly relevant, instances of the concept can be imagined that are not classified in it (Wittgenstein 1953).

Rather, in natural-kind nouns, such as *duck, oak* or *white bear*, an entity's classification depends on its degrees in the dimensions and their weights. Only its **averaged-similarity** to the concept's prototypical values must be high enough (higher than the classification norm; Rosch 1973; Tversky 1977; Rosch & Mervis 1975; Hampton 1995; 1998; Murphy 2002).

The averaged-similarity account is based on extensive experimental work within cognitive psychology. This work has shown that speakers systematically consider entities that score **highly on average in the dimensions** of a nominal concept as better examples than others of the concept (Rosch 1973). The exemplariness ratings of different participants are highly similar (McCloskey & Glucksberg 1978). They form a strong predictor of classification probability (Hampton 1998) and speed (Rosch 1973; Roth & Shoben 1983). Moreover, they systematically affect performance in any task involving nominal concepts (Murphy 2002: Chapter 1–2).

Thus, by the Accessibility Hypothesis, *optimistic* can **naturally** occur in dimensional and degree constructions, but *duck* cannot, because the interpretation of *optimistic* is based on dimension-counting, while that of *duck* is not. Before further examining this idea, however, a word about **naturalness** (or acceptability) judgments is needed.

1.1.2 Naturalness judgments about dimensional and degree constructions

Sprouse (2013) characterizes acceptability judgments as consciously reported perceptions of acceptability assessed by asking native speakers how **natural**, **acceptable**, or **grammatical** a spoken or written utterance is. The grammatical status of an utterance strongly but not uniquely influences its acceptability. Other factors include meaningfulness, plausibility, informativity, correctness, processing cost, familiarity and frequency of use. Moreover, acceptability judgments are often gradient (Sprouse 2013). However, whether the source of gradience is a gradient grammar (Sorace & Keller 2005; Bresnan & Ford 2010) or other factors such as, e.g., processing cost (Sprouse 2007) is under debate. Because of this complexity, different researchers use different instructions in assessing acceptability.

This paper reports an experimental assessment of **naturalness** judgments (see, e.g., Mahowald et al. 2016).[2] The notion of naturalness is readily accessible for non-linguists, and it is agnostic of the factors affecting the judgment (e.g., syntax, semantics, or

[2] This methodological paper evaluates how often published linguistic judgments in syntax and semantics research differ from judgments found in large-scale formal experiments with naive participants. Considering 100 English contrasts randomly sampled from *Linguistic Inquiry* 2001–2010, participants in an *acceptability rating task* were asked to rate how natural 606 sentences were on a Likert scale from 1–7.

processing). Attempts to use more specific notions such as *grammatical* or *meaningful* often fail to tackle a unique factor (Sprouse 1996: 157–160).

The predicates that can naturally occur in degree constructions such as the comparison construction in (10b) (*#more a duck than that one*) are called **morphologically gradable predicates** (Kennedy 1999; Kennedy & McNally 2005). Thus, assessment of naturalness of nominal predicates in such comparison constructions aims to directly tap into their morphological gradability.

Furthermore, morphological gradability is typically considered as evidence for an interpretation in terms of **a scale and a degree argument**; e.g., *optimistic* is typically assumed to relate to a scale (set of degrees ordered along the dimension optimism), and to denote a relation between degrees d and entities x who are optimistic to at least degree d. In positive forms like (4a), *This woman is optimistic*, the degree argument is implicitly saturated by the contextual membership norm (standard). In comparative constructions like (4b), *This woman is more optimistic than that one,* the degree argument is saturated explicitly by the than-clause to be, e.g., the maximal degree of optimism of the woman indicated by *that one*. Nouns like *duck* which are not natural in comparisons like, e.g., (10b), are assumed not to have a degree argument. They are rather assumed to directly link with entity-sets.

However, various hypotheses about the precise class of morphologically gradable predicates seem feasible in principle. One obvious hypothesis is syntactic. Adjectives are more natural than nouns in degree constructions, possibly because a formal feature such as +N prevents nouns from having a degree argument (Kennedy 1999; Baker 2003; Neeleman et al. 2004; Doetjes 2008). Another hypothesis, e.g., (3), is semantic or conceptual. Naturalness is affected by the conceptual structures of different types of nouns and adjectives. To tackle this alternative account, this paper addresses two conceptually distinct types of nouns.

The conceptual difference is hypothesized to yield different naturalness levels for nouns in various typically adjectival constructions, because it implies that a smaller interpretation shift is needed for some nouns than others to be licensed in these constructions. Thus, a differential processing cost (together with a penalty on naturalness) is suggested to be exerted due to conceptual differences. Since semantic or conceptual distinctions in naturalness are more subtle than syntactic ones, experimental data collection is important. However, the present study does not tap into processing data, it only sets out to refine our understanding of the naturalness levels of different nominal concepts in certain dimensional and degree constructions.

Moreover, while the reported study tests a number of predictions of the conceptual hypothesis and its results seem to support the role of conceptual factors, its design does not allow us to rule out other factors. Concrete suggestions for further research are therefore indicated.

Having set the study's goals, we are now ready to consider conceptual distinctions between different types of nouns.

1.1.3 The differential availability of dimension-counting in different types of nouns

Classification under **natural-kind nouns** (animals and plants) is often thought to be based on a causal model of the world (Wattenmaker 1995; Murphy 2002; Gelman 2003); e.g., a creature may be considered to be a duck because it has some essential property that causes it to look and behave like one. Since the dimensions are causally connected, the size of increase of an entity's classification probability given that it is similar to the prototype along one dimension (like *genetic makeup, appearance* or *offspring nature*) depends on the presence or absence of other dimensions. The absence of one dimension can nullify all the other dimensional contributions.

For example, Hampton et al. (2009) asked participants to help scientists classify hybrid creatures, e.g., an invented subkind with some features of lobsters and some features of crabs. In this task, the effect of one dimension varied as a function of the presence or absence of others. A big drop in classification probability occurred between entities whose values always matched the prototypical values and entities that manifested a mismatch in one dimension. Any additional mismatches incurred significantly smaller drops, suggesting that dimension-binding was non-additive. Moreover, dimensional effects on classification interacted significantly (a dimension had more effect in the presence of other dimensions), consistent with causal models. Finally, since the first mismatch was the most critical, hybrids of two kinds of creatures or plants were more likely to belong to neither kind than to both (Hampton et al. 2009).

Such findings have motivated a multiplicative representation, by which, e.g., (10a), *This bird is a duck*, is true iff the averaged degree of similarity to the prototype of the given bird in the dimensions of *duck*, **the weighted product** of its dimensional degrees, $\mathbf{f_1(x)^{w_1} \times \ldots \times f_n(x)^{w_n}}$, is above the membership norm (Medin & Schaffer 1978; for more complex models see Gelman 2003). For example, assume, for simplicity, equal dimensional weights ($w_1 = \ldots = w_n = 1$) and dimensional degrees between 0 and 1. Multiplicative averaging yields a low similarity degree even for entities that match the prototypical values in all the dimensions, except for one 0.5 score: $0.5 \times 1 \times \ldots \times 1 = 0.5$ (Murphy 2002).

Since the dimensions are bound by an averaged similarity function (e.g., weighted product) they are not accessible for binding by counting, accessing and quantifying operations. Thus, by the Accessibility Hypothesis in (3) natural-kind nouns should be judged unnatural in dimensional and degree constructions.

Interestingly, however, a different prediction concerns social nouns, namely nominal labels of human properties or human-made objects (e.g., *linguist, scarf, church*). The causal connections between the dimensions of social nouns (e.g., *intended function, actual use,* and *appearance*) are much looser, and accordingly the size of increase of an entity's classification probability given that it is similar to the prototype along one dimension does not depend on the degrees in the other dimensions (Wattenmaker 1995; Gelman 2003; Hampton et al. 2009). Therefore, the dimensions of social nouns are naturally modeled as bound by additive averaging operations (weighted sums).

For example, Hampton et al. (2009) also asked participants to classify artifacts which a secluded society habitually uses in ambiguous ways. For instance, some pieces of clothing had features of both a scarf and a tie. Such hybrids were often classified in both categories. Each dimension had a constant effect on classification, regardless of whether other dimensions were present or absent.

Such findings have motivated an additive representation (Rosch & Mervis 1975; Tversky 1977; Hampton et al. 2009). In this account, *This scholar is a linguist* is true if **the weighted sum** of degrees of the scholar in question in the various dimensions, $\mathbf{w_1f_1(x) + \ldots + w_nf_n(x)}$, is above the membership norm. Degrees are added, rather than multiplied, so each dimensional degree increases classification probability independently of the others. The decrease predicted by degree-addition in case of mismatches is much smaller compared with multiplication; e.g., $0.5 + 1 + \ldots + 1$ is almost the maximal sum possible, predicting category overlaps and high variability in dimensional values within categories.

In a series of additional studies, categorization under labels of social nouns (artifacts and human traits), unlike natural-kind nouns (animals and plants), appeared to often be based

on simple **counting**, i.e., on whether entities were within the norm in *some* (or *most*) of the dimensions (Wattenmaker 1995). This fact was modeled with **additive similarity, binary dimensional scales** (consisting of the degrees 1 and 0) and **equal dimensional weights**. Entities were classified under social concepts iff their *counting-based degree* – the number of dimensions whose norm they exceeded, **|{ f: f(x) ≥ norm(f) }|** – was above the concept's norm.

This case is unique in that the effect of averaging can be represented with quantifiers. A social noun N denotes the property that an entity has if it is classified under SOME (or MOST) of N's dimensions. This property discerns social nouns from natural-kind nouns, where classification is based on a weighted *product*, rather than *sum*, so even a flattened representation of their dimensions as binary and of equal weights does not reduce to counting. Counting is an additive process.

Thus, by default the dimensions of additive nouns are not accessible for binding by counting operations, but they can become accessible. When context triggers a shift to a flattened representation of the dimensions, the weighted sum of degrees of an entity is reduced to the number of dimensions under which it is classified, i.e., to counting (Wattenmaker 1995). Therefore, the status of social additive nouns in dimensional and degree constructions as in (12b)–(13c) is predicted to be improved.

(12) a. This scholar is a linguist.
b. ?This scholar is more a linguist than that one.

(13) a. ?This scholar is a linguist in her research interests, but not in affiliation.
b. ?This scholar is a linguist in {many ways, all/most respects}.
c. ?This scholar is {generally, overall, perfectly, very, more} a linguist.

1.1.4 Intermediate summary: The proposed conceptual account

To wrap up, past research suggests that multidimensional predicates divide into three types by the default way their dimensions are combined to form classification criteria. The dimensions of multidimensional adjectives are often bound by counting operations (Shamir 2013; Sassoon 2013a; b), whereas the dimensions of social and natural-kind nouns are usually bound by additive and non-additive (typically, multiplicative) averaged similarity operations, respectively:

(14) a. *Quantificational* adjectives like *optimistic* often involve counting of dimensions. As a default, entities fall under them iff they are classified under sufficiently many (e.g., some, most or all) dimensions.
b. *Multiplicative nouns* like *oak* or *duck* involve multiplicative averaging. As a default, entities fall under them iff the weighted product of their degrees in the dimensions is sufficiently high.
c. *Additive nouns* like *philosopher* or *chair* involve additive averaging. As a default, entities fall under them iff the weighted sum of their degrees in the dimensions is sufficiently high.

Moreover, the dominant readings of degree-constructions with multidimensional predicates are hypothesized to be based on dimension-counting, dimension-accessing, and quantification over dimensions, as in (15a–c).

(15) Hypothesized readings of comparisons of the form "x is more P than y":
  a. Dimension-counting: "x is P in more respects F than y is".
  b. Dimension-accessing: "x is P-er with respect to F", where F is a dimension of P.
  c. Quantification over dimensions: "x is P-er than y in every respect F".[3]

An interpretation based on dimension-counting is directly exploited in reading (15a) and indirectly exploited by the respect-accessing operations mediating readings (15b–c) (the operations denoted by *with respect to* and *in*). These operations seem to apply only to interpretations based on dimension-counting, namely adjectival interpretations. But additive nouns are more similar to adjectives than multiplicative nouns are in this respect. Given equally important dimensions with binary scales (consisting of 0 and 1 only), additive averaging reduces to counting. Classification depends on degree 1 in sufficiently many dimensions.

In hypothesis (3), repeated in (16), these conceptual distinctions affect the truth-conditions and thus naturalness of dimensional and comparative constructions, rendering, e.g., examples (4b)–(5) more natural than (10b)–(11), and the latter more natural than (12b)–(13). Naturalness mirrors the ease with which interpretation can shift to one based on dimension counting (adjectives > social additive nouns > natural multiplicative nouns).

(16) **The Accessibility Hypothesis:** The accessibility of the dimensions of a predicate for binding by counting operations predicts its naturalness in dimensional- and degree-constructions.

Moreover, the naturalness of predicates in the two constructions is expected to correlate as they share a main licensing condition: dimension accessibility (interpretation based on counting). Regarding degree constructions, this claim is especially non-trivial, as they also license adjectives which are clearly not multi-dimensional like *tall*. The question is whether there are, nonetheless, conceptual restrictions on the licensing of multidimensional predicates. The study reported in this paper aims to test precisely this. The following section presents the research questions and predictions in detail.

1.2 Research questions

1.2.1 Quantification over dimensions in classification forms

In the present proposal, the interpretation of adjectives, but not nouns, involves dimension-counting as a default. Hence, the study reported in this paper considers the naturalness of multidimensional adjectives, additive nouns and multiplicative nouns, respectively, both in classification forms as in (17)–(19), and dimensional-quantifier constructions as in (20)–(22).

(17) a. Ann is healthy.
  b. The neighborhood is safe.
  c. Bill is sick.
  d. The dish is dangerous.
  e. They are {clever, optimistic, pessimistic}.

[3] Sassoon (in progress) argues that the derived quantificational reading in (15c) is even more dominant than the counting-based reading in (15a). Quantificational readings are naturally restricted; e.g., in the context of *healthy,* only life-threatening dimensions, or no life-threatening dimensions can be quantified over. Such restrictions may be less readily available when counting, rendering comparisons of number of health dimensions inappropriate, because any life-threatening dimension weighs more than any non-life-threatening dimension. Counting-based readings are more appropriate for *healthier {food, attitude}*, where it is easier to regard the dimensions as equally important.

(18) This scholar is a linguist.

(19) This bird is a duck.

(20) a. Ann is healthy in all (contextually relevant) respects.
b. The neighborhood is safe in all respects.
c. Bill is sick in some respects.
d. The dish is dangerous in some respects.
e. They are {clever, optimistic, pessimistic} in as many respects as context requires.

(21) This scholar is a linguist in {some, most, all} respects.

(22) This bird is a duck in {some, most, all} respects.

The first question is whether, unlike classification forms, quantificational forms with adjectives, social nouns and natural-kind nouns, as in (20)–(22), systematically exhibit different levels of naturalness. If an interpretation based on dimension-counting is a precondition for licensing respect arguments and the quantifiers that bind them, then such quantifiers are predicted to combine naturally with adjectives that have interpretations that are readily based on dimension-counting, to combine slightly less naturally with additive social nouns whose interpretations become counting-based only by shifting to a representation of their dimensions as binary and equally important, as explained above, and to combine even less naturally with multiplicative natural-kind nouns, which do not become counting-based even upon such a shift, since multiplicativity should also be overridden in favor of additivity.

In sum, naturalness is expected to mirror the ease with which dimensions can be accessed and quantified over (e.g., (20) > (21) > (22)). Confirmation of this prediction would highlight the role of dimension accessibility for binding by counting-operations in predicting the naturalness of dimensional-quantifier constructions, while disconfirmation would speak against such a distributional restriction.

Importantly, the three naturalness levels predicted for dimensional-quantifier constructions are not expected to occur in basic classification constructions (e.g., (17) to (19)), where licensing of predicates is not dependent on dimension accessibility for binding by counting operations. Thus, classification constructions serve as baselines in the reported study, and the following predictions are tested:

(23) **Predictions for dimensional-quantifier vs. classification constructions**
a. Assuming that dimension accessibility for binding by counting operations is higher in adjectives than nouns, a significant interaction is expected between *predicate type* (counting-based adjectives vs. similarity-based nouns) and *construction* (classification vs. quantification), due to a predicate-type effect (adjectives > nouns) only in quantificational constructions that require dimension-counting.
b. An interaction is expected also between *entity-type* (social additive nouns vs. natural-kind multiplicative nouns) and *construction* (classification vs. quantification) with an entity-type effect (social > natural) only in quantification-forms.
c. No interaction is expected in adjectival predicates, assuming that they are based on counting by default, regardless of the construction in which they occur or the entities to which they apply.

The second question is to what extent the naturalness of classification and quantificational forms correlate. When no particular dimension is made salient, adjectival classification forms such as (17a–e) are predicted to have quantificational readings as in (20a–e), respectively. But no corresponding equivalences are predicted to hold between the nominal classification constructions in (18)–(19) and respective dimensional-quantifier constructions in (21)–(22).

The past research reviewed in section 1.1 indirectly supported this view by considering the acceptability and frequency of use of dimensional exception phrases (indicators of universal quantification). The present research investigates this view by directly comparing classification forms and explicit quantificational forms with *every, most,* and *some*. If adjectival classification forms are based on dimension-counting (and thus are readily equivalent to quantificational forms), whereas nominal classification forms are not (they are based on averaged similarity), we expect to see a correlation between classification and quantification forms in adjectives, but not nouns.

This question is even more basic than the former. Confirmation of its prediction would support the role of quantification over dimensions in adjectival as opposed to nominal classification constructions, while disconfirmation would speak against a conceptual difference and in favor of the more standard one-dimensional representations of adjectives' meanings.

The proposed account is easily refutable. The interpretation of predicates may not involve dimensions as building blocks, or the dimensions of adjectives and nouns may not be bound by counting- and averaging-operations, respectively. And even if they do, these dimension-binding operations may be altogether irrelevant to the semantics of dimensional- and degree-constructions, as the standard theory assumes. As opposed to the hypothesized readings of comparatives (cf., (15a–c)), their dominant readings may actually involve **averaging**; e.g., "x is more P than y" may predominantly convey that x's averaged degree in P's dimensions is higher than y's (cf., Bylinina 2014; McNally & Stojanovic 2015; Solt 2017).

In this view, *healthy*, for example, means healthy on average. More specifically, in the standard account, different health measurements are not encoded by grammar. Rather, entities' degrees in those measurements are combined creating a unique 'averaged health' scale, and it is this scale alone that grammar is assumed to encode.

In this account, adjectives are not different than nouns in their conceptual structure, but they differ from nouns in their semantic type. Adjectives have a degree argument, nouns do not. Thus, adjectives are morphologically gradable, while nouns, social as well as natural, are not. Dimensional quantification (as in *healthy/chair/pine in some/all respects*) in this account must involve metalinguistic reference to multiple aspects of a word even in adjectives, since the semantics of, e.g., *healthy* only represents a single dimension, e.g., health. Entities' degrees on this dimension are built through operations like averaging, similarly to the way they are built in nouns. Thus, neither are any conceptual differences expected to affect naturalness, nor are any higher correlations between classification and dimensional-quantifier constructions expected to occur in adjectives than nouns. In sum, the predictions of the proposed conceptual account contrast with those of the standard account.

1.2.2 Quantification in degree constructions

The current study considers also comparative forms such as *more healthy* (or *healthier*), *more a pine,* and *more a chair*. In most current literature, all or most nouns (Kennedy 1999; Baker 2003; Morzycki 2009; de Vries 2010) are analyzed as denoting non-gradable properties (type $<e,t>$), while most adjectives are analyzed as denoting gradable properties (e.g., relations between entities and degrees, type $<d,<e,t>>$; Cresswell 1976;

von Stechow 1984; 2009; Heim 2000; Beck 2011). Thus, nouns are hypothesized not to license degree modification (e.g., comparative forms), because they do not have a degree argument (Kennedy 1999; Baker 2003).

By contrast, the Accessibility Hypothesis incorporates the cognitive psychological gradable conception of nouns into their formal semantics. This conception is consistent with the assignment of **a uniform type** to adjectives and nouns, the only distinction between them being that they denote **gradable properties of different sorts**. Adjectival degree relations predominantly utilize scales representing dimension-counting, whereas nominal relations predominantly utilize scales based on averaged similarity (for multiple gradable sorts, see also Doetjes 2008: 149–154; Bochnak 2010; Morzycki 2012).

Predicates with dimension-counting interpretations are hypothesized to be licensed in degree constructions. Thus, social nouns, whose additive interpretations shift relatively easily to dimension-counting, are expected to be better in comparison constructions than natural-kind nouns, whose multiplicative interpretations do not shift easily (Wattenmaker 1995).

Introspective judgments reported in the literature indicate that some +human nouns more readily accept degree modification than other nouns (de Vries 2010; 2015; Morzycki 2011; 2012). But the predictions of the present proposal are more radical. Social nouns in general, whether +human or not, by virtue of being additive, are predicted to be judged more natural in comparison constructions than natural-kind nouns, which are generally multiplicative.

Confirmation of this prediction would highlight the role of dimension-binding (relative accessibility for binding by counting-operations) in predicting the naturalness of degree constructions, over and above the role of the noun/adjective and –/+human features. However, disconfirmation would speak against the dimension-accessibility hypothesis and in favor of a more standard non-gradable representation of most nouns.

Confirmation of the prediction would not entirely rule out a standard degree-less theory of nouns (see, e.g., Constantinescu 2011 and references therein). It would merely imply that when nouns combine relatively naturally with degree morphemes they also shift their type to that of adjectives. But such a shift would be necessary in all nouns. Thus, higher naturalness ratings for social than natural-kind nouns in comparatives would nonetheless require an additional explanation in terms a conceptual shift (the conceptual shift needed in social nouns is indeed smaller because of their additivity). Hence, assuming a degree relational type for nouns in the first place is preferable, since it is more parsimonious.

In addition, big differences between the naturalness of social nouns and adjectives in degree constructions may suggest that a costly type shift is indeed involved in addition to a conceptual shift in dimension representation. However, small or insignificant differences between adjectives and social nouns would render the degree-based account for nouns, where only a conceptual shift is needed for licensing in degree constructions, more appealing.

Importantly, the three naturalness levels predicted for comparative constructions (adjectives > social additive nouns > natural multiplicative nouns) are not expected to occur in basic classification constructions, where licensing of predicates is not dependent on dimension accessibility for binding by counting operations.

The study considers three different comparison constructions ('more P than {y, Q, y is Q}'), to test the rival hypotheses that nouns are or are not gradable by asking whether nouns manifest lower or equal naturalness compared with adjectives, and whether any such differences arise in nouns generally (consistent also with a syntactic explanation), or only in the natural-kind noun domain (supporting a conceptual explanation). Hence, the following predictions of the conceptual account (Accessibility Hypothesis) are tested:

(24) **Prediction for comparison vs. classification constructions**

a. Assuming dimension-counting is a predictor of licensing in degree constructions, a significant interaction is expected between *predicate-type* (counting-based adjectives vs. similarity-based nouns) and *construction* (classification vs. comparison), due to a predicate-type effect (adjectives > nouns) only in comparisons.

b. Since additive nouns shift to dimension counting more easily, assuming dimension-counting is a predictor of licensing in degree constructions, a significant interaction is also expected between *entity-type* (social additive nouns vs. natural-kind multiplicative nouns) and *construction* (classification vs. comparison), with an entity-type effect (social > natural-kind nouns) only in comparisons.

c. No interaction is expected in adjectival predicates, assuming they are based on counting as a default, regardless of the construction or entities to which they apply.

d. For each comparison type, we ask whether an adjectival advantage is manifested over additive nouns, to test whether the construction is restricted syntactically (to adjectives) or only conceptually (to quantificational or additive concepts).

Moreover, the reported study enables a more direct testing of the Accessibility Hypothesis than mere introspection over single examples does, because it makes it possible to test whether the naturalness of nominal constructions involving comparison (*more healthy/a pine/a chair*) and quantification over dimensions (*heathy/a pine/a chair in {some, most, all} respects*) correlates. According to the Accessibility Hypothesis, the prominent readings of degree constructions with multidimensional predicates involve counting, accessing, and quantifying over dimensions (cf., (15a–c)). The study tests this hypothesis by asking to what extent the naturalness of comparative and quantificational forms correlates.

A correlation would support the proposal that multiple dimensions, and the ease with which they are accessed and bound by counting-operations, affect the status and truth conditions of both dimensional and degree-constructions. By contrast, absence of a correlation would speak against these hypotheses ((3) and (15)) and in favor of the standard theory, by which comparative interpretations are based on single dimensions. For example, *Ann is healthier than Bill* may after all predominantly convey that Ann is (on average) healthier than Bill (for literature reviews on comparative interpretations see Kennedy 1999; Schwarzschild 2008; Beck 2011; for averaging-based accounts of various adjective types see Bylinina 2014; McNally & Stojanovic 2015; Solt 2017). Let us turn to the study that tested these predictions.

2 A study of naturalness judgments

2.1 Method

2.1.1 Participants

139 adult English speakers were recruited using Amazon mechanical Turk (AMT), an online labor market-place which has been shown to provide a quick and efficient method to acquire high-quality experimental results that do not differ significantly in performance from standard experimental settings (Buhrmester et al. 2011). All participants were American with an approval rate – an index of reliability – exceeding 95%. They were awarded 2 cents for the completion of each experimental item, with an average hourly rate of $6.5.

2.1.2 Stimuli

The **subject positions** of the sentences we examined included 10 natural nouns denoting plants and animals and 10 social nouns denoting human traits and artifacts, listed in (25a–d), respectively.

(25) **Subjects denoting natural (a, b) and social (c, d) entity types**
- a. tree, vegetable, flower, bush, piece of fruit
- b. farm animal, insect, bird, reptile, predator
- c. journalist, artist, football player, person, colleague
- d. place, piece of clothing, car, booklet, container

Each of these 20 subject-nouns was paired with a pair of **nominal predicates** denoting more specific members of the same conceptual category (e.g., *bird* was matched with *crow* and *pigeon*, and *container* was matched with *carafe* and *vase*). Further, the 20 subject-nouns were paired with 10 pairs of multidimensional **adjectival predicates**, appropriate for the description of both natural and social subjects (e.g., *exciting/boring, efficient/inefficient*). The complete list of subject-nouns alongside their nominal and adjectival predicates is provided in Appendix A.[4]

The subject nouns, nominal predicates and adjectival predicates were used to construct 20 sets of 15 sentences. Each set featured one of the 20 subject-nouns, and its respective nominal and adjectival predicates in 15 different target conditions, including the three comparison structures in (26), the three quantification structures in (27), the basic classification structure in (28a), and the structure in (28b) with *small* in attributive position of nominal predicates. This condition was included to control for the reliability of the judgments, as it was expected to exhibit a reversed entity-type effect compared with the target conditions (intuitively, *small* is more naturally attributed to nouns denoting natural-kinds and artifacts than to +human nouns; cf., *small {duck, chair, #linguist}*).

(26) **Comparisons**
- a. Between-predicate comparison structure (*bet1s*)
  1. Nominal: This tree is more a pine than an oak.
  2. Adjectival: This tree is more safe than dangerous.
- b. Between-subject-and-predicate comparison structure (*bet2s*)
  3. Nominal: This tree is more a pine than that one is an oak.
  4. Adjectival: This tree is more safe than that one is dangerous.[5]
- c. Within-predicate comparison structure (*within*)
  5. Nominal: This tree is more a pine that that one.
  6. Adjectival: This tree is safer than that one.

[4] The nominal pairs were mostly withdrawn from the stimuli reported in Hampton et al. (2009). The pairs of human traits and professions were chosen from a list created by searching the corpus of contemporary American English (COCA; Davies 2010) for the string "(s)he is a" followed by a noun. The adjectives were chosen from a list of multidimensional gradable adjectives, created by searching COCA for "more adjective and adjective" and removing the one-dimensional adjectives (which were compared with nouns on a separate study). The classification of an adjective as multidimensional was based on the results of a pretest which involved 95 AMT participants that were presented with 61 randomly chosen adjectives from the above list. They were asked to respond to the question "Can individuals be classified as ADJECTIVE in some respect and NOT ADJECTIVE in some other respect?" by providing a rating between 1 (certainly not) and 7 (certainly yes). The multidimensional adjectives for this study were chosen from the 20 highest ranked adjectives. 20 participants rated each adjective. On average, each participant rated 20.5 adjectives (SD = 18.9).

[5] Three adjectives (*healthy*, *safe* and *clean*) could combine both with a free and a bound comparison morpheme, as in, e.g., *{more safe, safer) than that one is dangerous*. Thus, we also included the bound forms in a set of additional fillers (see below). However, as their ratings (detailed in Appendix B) did not differ systematically from those of the free forms, there was no need to include the bound forms in the statistical analyses.

(27) **Quantifications over dimensions**
- a. *Every respect* structure (*every*)
  7. Nominal: This tree is a pine in every respect.
  8. Adjectival: This tree is safe in every respect.
- b. *Most respects* structure (*most*)
  9. Nominal: This tree is a pine in most respects.
  10. Adjectival: This tree is safe in most respects.
- c. *Some respect* structure (*some*)
  11. Nominal: This tree is a pine in some respect.
  12. Adjectival: this tree is safe in some respect.

(28) **Control structures**
- a. Basic structure
  13. Nominal: This tree is a pine.
  14. Adjectival: This tree is safe.
- b. *Small* structure (only nominal)
  15. Nominal: This tree is a small pine.

In sum, 300 (20×15) experimental items were formed by manipulating 3 factors: *Entity type* (natural/social subject), *Predicate type* (nominal/adjectival predicate), and *Structure* (comparisons/quantifications/control structures). Given the diverse variety of structures and concepts, this design did not demand inclusion of fillers. However, to assure that participants provide genuine naturalness judgments, 24 fillers were added, including clearly natural and clearly unnatural items (see Appendix B), along with 60 sentences of a separate experiment.[6]

2.1.3 Design and procedure

In the online AMT platform, participants are usually asked to complete small tasks named HITs (Human Intelligence Tasks). In our study, each HIT consisted of one item. The 384 sentences were presented as a single randomly ordered list, out of which different participants completed different subsets. Each participant chose how many hits to fill out. Our factors are therefore completely manipulated within-item but only partially manipulated within-subject (i.e., there are missing observations). The following instructions preceded each HIT.

(29) *[This hit is for English Native Speakers Only]*
RATE THE FOLLOWING SENTENCES BY HOW NATURAL THEY SOUND TO YOU AS AN ENGLISH NATIVE SPEAKER.
FOR EXAMPLE: (i) To me, the sentence "This child is rarely sick" is a perfectly natural sentence, so I give it '7'; (ii) However, the sentence "This child is rarely tall" is perfectly unnatural and makes no sense (me or my friends would never use such a combination of words), so I give it '1'".

Following the instructions, the participants saw a sentence and a 7-point scale with the labels PERFECTLY NATURAL and PERFECTLY UNNATURAL adjacent to 7 and 1, respectively. Participants then provided their rating for the hit by clicking on the radio button adjacent to the relevant number. Once 25 participants filled out a hit, it was no longer visible. In total, 139 participants answered an average of 69.06 different hits each (SD = 81).

[6] Their structures were as in *I consider this tree {a pine, safe}* and *This tree is a real pine*.

2.2 Results

2.2.1 Mean ratings and factor analyses

Mean ratings by structure (per entity type and predicate type) are presented in Figures 1 and 2 (see tables of mean ratings and standard deviations in Appendices C-D; for detailed tables by items see Appendices E-F). In Figure 1, which presents the mean ratings of structures with adjectival predicates, we see that social items were rated higher than natural items across structures. By contrast, in Figure 2, which presents the mean ratings of structures with nominal predicates, we observe an interaction between structure and entity type. Entity type did not affect ratings of basic structures, while quantificational and comparison structures exhibited lower ratings with natural entities, and the *small* control condition exhibited lower ratings with social ones. We elaborate on these distinctions below.

To analyze our results, we used the *lmerTest* package in R (Kuznetsova et al. 2014) to fit mixed effects models to our data, with naturalness ratings as the dependent variable and participants and items as random effects.[7]

Our first mixed model defined the three manipulated factors (Structure, Predicate type and Entity type) as fixed factors and was applied on the entire set of observations. Following Barr et al. (2013), we started out by running a maximal model including subject and item random intercepts and random slopes for all fixed effects and their interactions. Due to convergence failure, all slopes were removed for both subjects and items. Let us refer to this model as the Overall Model.

As Table 1 indicates, this model yielded significant main effects of Structure, Entity type and Predicate type. Further, this analysis yielded significant interactions between

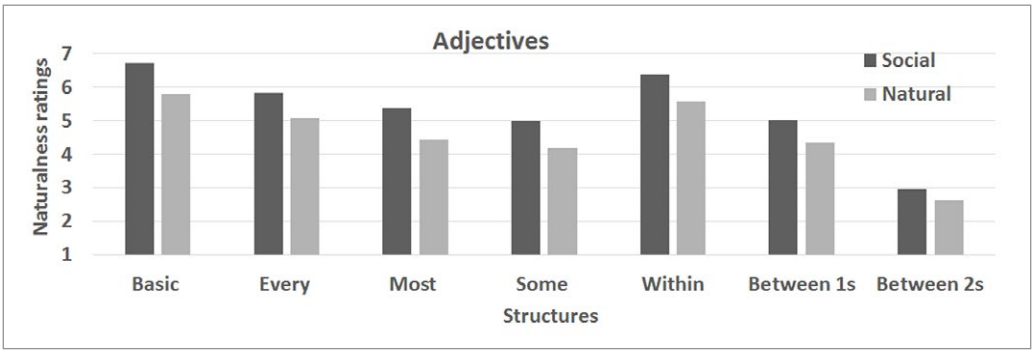


Figure 1: Mean ratings by Structure and Entity type for adjectives.

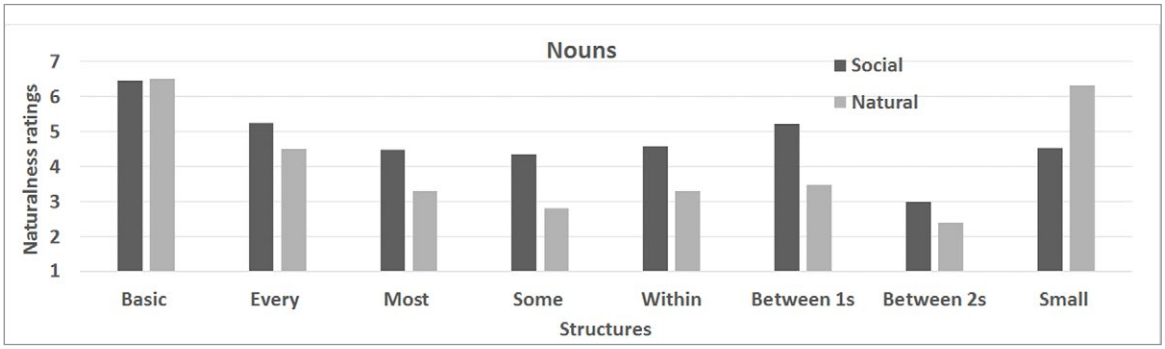


Figure 2: Mean ratings by Structure and Entity type for nouns.

[7] The test-statistics were obtained by the application of the functions *ANOVA* (for *F* and p-values evaluating the role of the fixed factors as predictors), *summary* (for estimates, labeled as β, standard errors and *t* and p-values evaluating the difference between β and 0), and *difflsmenas* (for post-hoc pairwise comparisons where p-values in the omnibus F test indicate a significant interaction).

Structure and Predicate type, Structure and Entity type, and Predicate type and Entity type, as well as a three-way interaction between Structure, Entity type and Predicate type.

To reveal the sources of these patterns, more detailed analyses were required. Focusing on the predictions in (23)–(24), we conducted post-hoc pairwise comparisons and fitted additional mixed models for relevant subsets of our data when needed. Their results are presented in the following sections, followed by correlation-test results. These sections are organized around the research questions. For readability, research questions and predictions are presented together with the models testing them, their results, and a short indication of whether the predictions are borne out. A general discussion is found in section 3.

2.2.2 Quantification over dimensions in classification structures: Predictions and results

	F	***p*-value**
Structure	$F(8,8338.8) = 452.51$	< .001
Entity	$F(1, 18) = 15.54$	< .001
Predicate	$F(1,8834.4) = 328.19$	< .001
Str×Pred	$F(7, 8339.1) = 52.07$	< .001
Str×Entity	$F(8, 8339.9) = 50.38$	< .001
Entity×Pred	$F(1, 8343.3) = 5.93$	= .015
Str×Pred×Ent	$F(7, 8337.8) = 10.33$	< .001

Table 1: Results of the Overall Model analysis applied on the entire data set.

The first research question is whether, unlike classification structures (e.g., 'x is P'), quantificational structures ('x is P in {some, most, every} respect') with adjectives, social nouns and natural-kind nouns systematically exhibit different levels of naturalness. If interpretations based on dimension-counting are a precondition for licensing respect arguments and quantifiers that bind them, then such quantifiers are predicted to combine naturally with adjectives whose interpretations are often based on counting, slightly less naturally with additive social nouns whose interpretations become counting-based only by shifting to a representation of their dimensions as binary and equally important, and even less naturally with multiplicative natural-kind nouns, which do not become counting-based even upon such a shift, since multiplicativity should also be overridden in favor of additivity.

In sum, naturalness in quantification structures is expected to correspond with the ease with which dimensions can be accessed and counted. Importantly, the three naturalness levels predicted for dimensional-quantifier constructions are not expected to occur in basic classification constructions, where licensing of predicates is not dependent on dimension accessibility for binding by counting operations.

Thus, the first prediction (23a) is an interaction between predicate-type and structure. As reported above, such an interaction was indeed observed. More specifically, a predicate-type effect – higher naturalness of counting-based adjectives (e.g., *safe*) than similarity-based nouns (e.g., *a pine* or *church*) – has been expected in quantificational structures (e.g., (30a, b) < (30c)), but not in basic classification structures (e.g., (31a, b) vs. (31c)).

(30) Quantification structures
- a. Natural nominal: This tree is a pine in {some, most, every} respect.
- b. Social nominal: This place is a church in {some, most, every} respect.
- c. Natural/Social adjectival: This tree/place is safe in {some, most, every} respect.

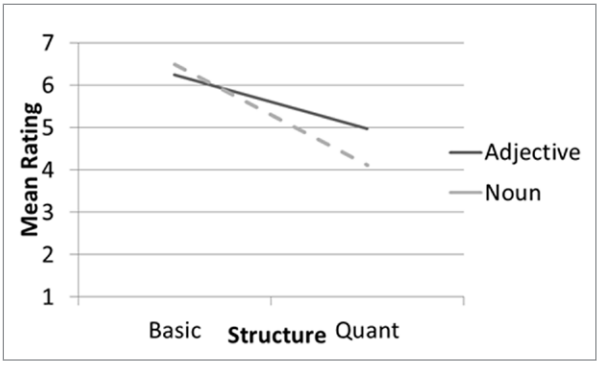


Figure 3: Mean ratings for quantifications and basic structures by Predicate type.

(31) Basic classification structures
  a. Natural nominal: This tree is a pine.
  b. Social nominal: This place is a church.
  c. Natural/Social adjectival: This tree/place is safe.

Planned pairwise comparisons with an application of a Bonferroni correction for multiple comparisons (X11) revealed this prediction to be correct.[8] As demonstrated in figure 3, nominal quantifications (M = 4.11, SD = 2.08) were rated significantly lower than adjectival quantifications (M = 4.90, SD = 1.92, *Some respects* $t(8340.8) = 9.78$, $p < .001$, *Most respects* $t(8337.3) = 10.58$, $p < .001$, *Every respect* $t(8337.3) = 5.48$, $p < .001$), while basic nominal items (M = 6.48, SD = 1.16) were not significantly different than basic adjectival items (M = 6.23, SD = 1.52, $t(2329.2) = -2.61$, $p = .1$).

Moreover, if dimension accessibility is higher in additive than in multiplicative nouns, then the second prediction (23b) is an interaction between Entity-type and Structure. An Entity-type effect – higher naturalness of social additive nouns (e.g., *church*) than natural-kind multiplicative nouns (e.g., *pine*) – is expected in quantificational structures (e.g., (30a) < (30b)), but not in classification structures (e.g., (31a, b)).

The third prediction (23c) is absence of interaction in adjectival predicates, which are counting-based as a default, regardless of the structure or whether they are applied to social or natural entities (as in, e.g., (30c) and (31c)).

To test these predictions, a model with the fixed effects Structure (basic vs. quantificational) and Entity type (natural vs. social) was applied once to the set of nominal quantificational and basic items and once to the set of adjectival quantificational and basic items. We initially ran a maximal model. Due to convergence failure of these models with both data sets, all slopes were removed for both subjects and items.

As Table 2 indicates, the predictions were confirmed. While the analysis of the nominal data-set yielded a significant main effect of both Structure ($p < .001$) and Entity type ($p = .005$) and a significant interaction ($p = .003$), the analysis of the adjectival data-set only yielded main effects of Structure ($p < .001$) and Entity type ($p < .001$), while the interaction was not significant ($p = .87$). Figures 4 and 5 demonstrate this distinction.

Planned pairwise comparisons with an application of a Bonferroni correction for multiple comparisons (X2) revealed the cause for the significant interaction in the nominal

[8] We applied a Bonferroni correction for multiple comparisons once for each model, according to the number of pairwise comparisons performed under it. For example, as we report 11 planned pairwise comparisons ran under the Overall Model, the *p*-values we report are the ones that were obtained by the function *difflsmeans* multiplied by 11. Where uncorrected *p*-values were already quite high, we report the original values followed by the notation "uncorrected".

Nouns	***F***	***p-value***	***β***	**SE**	***t***	***p-value***
Structure	$F(1, 75.9) = 168.3$	**< .001**	$\beta_{quantification} = -2.9$	.26	$t(75.9) = -11.4$	< .001
Entity	$F(1, 75.8) = 8.43$	**= .005**	$\beta_{social} = -.03$	.32	$t(75.9) = -0.12$	= .9
Str×Entity	$F(1, 76.03) = 9.65$	**= .003**	$\beta_{quantific:social} = 1.14$	.36	$t(76.03)=3.106$	= .003
Adjectives						
Structure	$F(1, 76.05) = 44.9$	**< .001**	$\beta_{quantification} = -1.26$	.27	$t(75.08)= -4.85$	< .001
Entity	$F(1, 76.2) = 21.64$	**< .001**	$\beta_{social} = .83$	.32	$t(76.26) = 1.55$	= .011
Str× Entity	*F(1,76.3) = .026*	**= .87**	$\beta_{quantific:social} = .058$	.36	*t(76.35) = .160*	= .87

Table 2: Results of separate analyses for nominal and adjectival items.

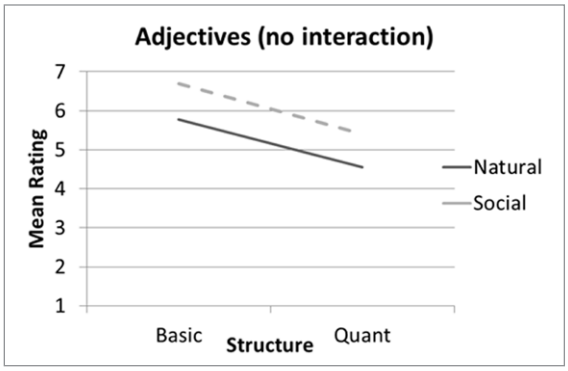


Figure 4: Mean ratings for basic and quantification structures by Structure and Entity type for adjectives.

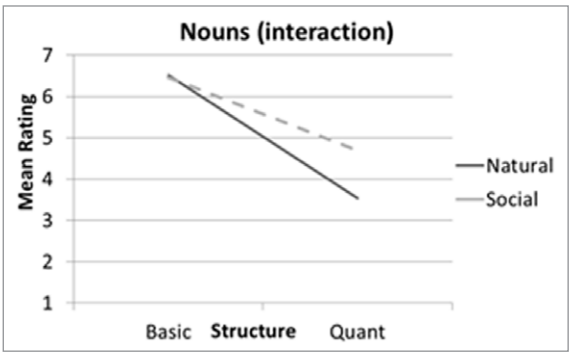


Figure 5: Mean ratings for basic and quantification structures by Structure and Entity type for nouns.

data. While social nominal quantification structures (M = 4.7, SD = 1.95) were rated significantly higher than natural nominal quantification structures (M = 3.6, SD = 2.03, $t(76) = -6$, $p < .001$), there was no significant difference between the social nominal basic structures (M = 6.46, SD = 1.09) and natural nominal basic structures (M = 6.51, SD = 1.2, $t(75.12) = .12$, $p = .9$, uncorrected). This differentiated the nouns from the adjectives, where the entity-type difference observed in the adjectival quantificational forms was no bigger than the difference already manifested in the basic forms. Hence, predictions (23b, c) were confirmed.[9]

[9] For additional statistical analyses suggesting a quantifier type effect (*every* > *most, some*) see online appendix I in https://sites.google.com/site/weidmansassoon/docs/general-documents/Online%20appendices.pdf.

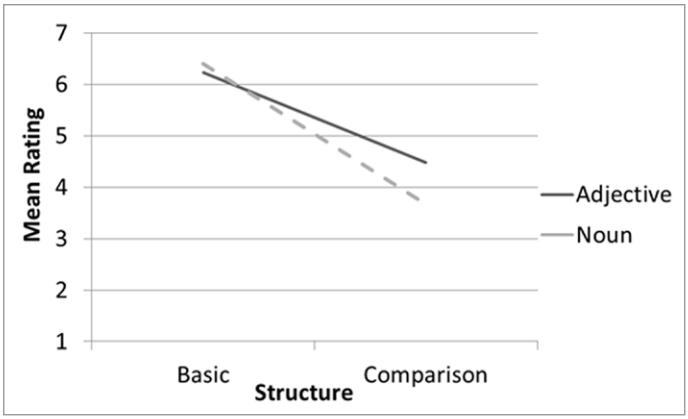


Figure 6: Mean ratings for comparison and basic structures by Predicate type.

2.2.3 Quantification over dimensions in comparison structures: Predictions and results

The second research question is whether dimension accessibility for binding by counting operations raises the naturalness of predicates in degree constructions, i.e. whether, being additive, social nouns are more natural in comparison structures such as those in (32) than natural-kind nouns, which are multiplicative. Confirmation of this prediction would support the role of dimension-binding (dimension accessibility) in predicting the naturalness of degree constructions, over and above the role of the noun/adjective and –/+human features. However, disconfirmation of the prediction would speak against the Dimension-Accessibility Hypothesis and in favor of a more standard non-gradable representation of most nouns.

Importantly, the three naturalness levels predicted for comparative structures (adjectives > additive nouns > multiplicative nouns) are not expected in basic classification structures such as those in (33), where licensing of predicates is not dependent on dimension accessibility for binding by counting operations.

Thus, the first prediction (24a) is an interaction between predicate type and structure. A predicate-type effect – higher naturalness of counting-based adjectives like, e.g., *safe*, than similarity-based nouns like, e.g., *pine* or *church* – has been expected in comparison structures (e.g., (32a, b) < (32c)), but not in classification structures (e.g., (33a–c)). As the reported results of the Overall model indicate, such an interaction (demonstrated in figure 6) was indeed observed. To test the source of the interaction, we went back to the Overall Model and inspected further pairwise comparisons.

(32) Comparison structures
- a. Natural nominal: This tree is more a pine than {that one, an oak, that one is an oak}.
- b. Social nominal: This place is more a church than {that one, an art gallery, that one is an art gallery}.
- c. Natural/Social adjectival: This tree/place is more safe than {that one, dangerous, that one is dangerous}.

(33) Basic classification structures
- a. Natural nominal: This tree is a pine.
- b. Social nominal: This place is a church.
- c. Natural/Social adjectival: This tree/place is safe.

After an application of a Bonferroni correction for multiple comparisons (X11) this analysis revealed that our prediction was confirmed for Within comparisons (“more P than that one”) and Bet1s comparisons (“more P than Q”), in which nominal comparisons

were rated lower than adjectival comparisons (adjectival *within*: M = 5.97 , SD = 1.62; nominal *within*: M = 3.94, SD = 2.17; $t(8344.7) = 20.62$, $p < .001$; adjectival *bet1s*: M = 4.68, SD = 1.92; nominal *bet1s*: M = 4.35, SD = 2.21; $t(8336.7) = 3.4$, $p = .007$).

The prediction was not confirmed for *Bet2s* comparisons ("more P than that one is Q"), in which ratings for nouns and adjectives did not differ significantly (adjectives: M = 2.79, SD = 1.67; nouns: M = 2.7, SD = 1.73; $t(8338.1) = .9$, $p = .3$, uncorrected).

Furthermore, as predicted and reported above, basic nominal items (M = 6.4 , SD = 1.6) did not differ significantly from basic adjectival items (M = 6.23, SD = 1.5 , $t(2329.2) = –2.61$, $p = .1$). Hence, the predicted difference between syntactic categories (nouns and adjectives) was only confirmed in two out of three comparison types.

Moreover, if dimension accessibility for binding by counting operations, which is higher in additive than in multiplicative nouns, licenses morphological gradability, then the second prediction (24b) is an interaction between entity-type and structure. An entity-type effect – higher naturalness for social additive nouns like *church* over natural-kind multiplicative nouns like *pine* – is expected in comparisons ((32a) < (32b)), but not in classification structures (33a, b). Further, the third prediction (24c) is absence of interaction in adjectival predicates, which are based on counting as a default, regardless of the entities they apply to or the structure (as, e.g., (32c) and (33c) illustrate).

To test these predictions, a model with the fixed effects Structure (basic vs. comparison) and Entity type (natural vs. social) was applied, once to the set of the nominal comparisons and basic items and once to the set of the adjectival comparisons and basic items. We initially ran a maximal model. Due to convergence failure of these models with both data sets, all slopes were removed for both subjects and items.

As Table 3 indicates and Figures 7 and 8 illustrate, the predictions were borne out. While the analysis of the nominal data-set yielded a significant main effect of both Structure ($p < .001$) and Entity type ($p = .007$), and a significant interaction ($p = .002$), the analysis of the adjectival data-set only yielded main effects of Structure ($p < .001$) and Entity type ($p = .03$). The interaction was not significant ($p = .68$).

Planned pairwise comparisons with an application of a Bonferroni correction for multiple comparisons (X2) revealed the cause for the interaction in the nominal data. While the social nominal comparisons (M = 4.3, SD = 2.15) were rated significantly higher than the natural nominal comparisons (M = 3.06, SD = 2, $t(76.1) = –5.89$, $p < .001$), there was no significant difference between the social nominal basic structures (M = 6.46, SD = 1.09) and natural nominal basic structures (M = 6.51, SD = 1.2, $t(76.1) = .23$, $p = .82$, uncorrected). This differentiated the nouns from the adjectives, where the entity type

Nouns	***F***	***p-value***	***β***	**SE**	***t***	***p-value***
Structure	$F(1,76.05)=182.23$	**< .001**	$\beta_{comparison}$ = -3.5	.3	$t(76.05)=-11.76$	< .001
Entity	$F(1,76.08) = 7.56$	**=.007**	β_{social} = -.08	0.36	$t(76.07)=-0.226$	= .8
Str×Entity	$F(1,76.08) = 7.86$	**=.002**	$\beta_{comp:social}$ = 1.3181	.42	$t(76.08) = 3.14$	=.0024
Adjectives						
Structure	$F(1,76.975) = 27.9$	**<.001**	$\beta_{comparison}$ = -1.64	.46	$t(75.9)= -3.45$	< .001
Entity	$F(1, 76.05) = 4.87$	**= .03**	β_{social} = .88	.6	$t(76.06) = 1.55$	= .13
Str× Entity	*$F(1,76.07) = .17$*	**= .68**	$\beta_{comp:social}$ = -.28	.67	*$t(76.07) = -.41$*	=.68

Table 3: Results of separate analyses nominal adjectival items.

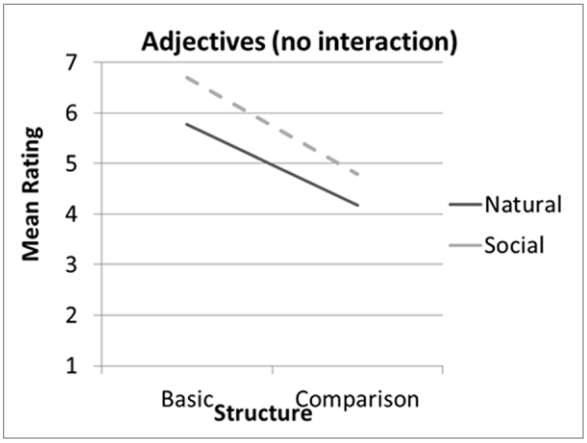


Figure 7: Mean ratings for comparisons and basic structures by Structure and Entity type for adjectives.

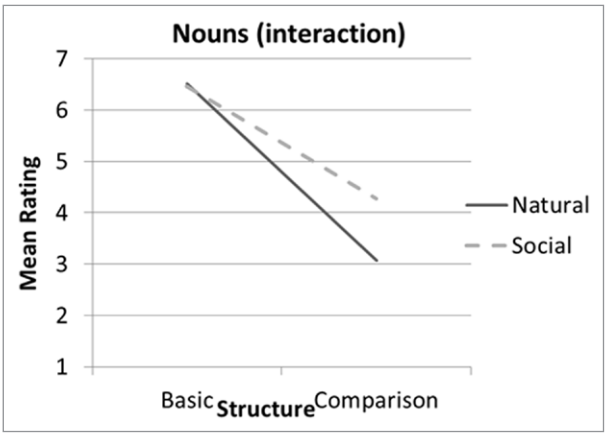


Figure 8: Mean ratings for comparisons and basic structures by Structure and Entity type for nouns.

difference observed in the comparison forms was no bigger than the difference already manifested in the basic forms. Hence, predictions (24b, c) were confirmed.

Finally, for each comparison type, we ask in (24d) whether an adjectival advantage is manifested over additive nouns, i.e., whether the construction is restricted to adjectives, or to predicates which can be interpreted additively (adjectives and additive nouns). That is, we ask whether mainly the natural nouns contribute to the differences observed between nouns and adjectives ((32a) < (32c)), while the social nouns do not (e.g., (32b) vs. (32c)).

Thus, to find out whether adjectives do or do not differ from additive social nouns, a model with the fixed effects Comparison type (Within vs. *Bet1s* vs, *Bet2s*) and Predicate type (nominal vs. adjectival) was applied, once to the set of the comparisons with a Natural subject and once to the set of the comparisons with a Social subject. We initially ran a maximal model. Due to convergence failure of these models with both data sets, all slopes were removed for both subjects and items.

Planned pairwise comparisons with an application of a Bonferroni correction for multiple comparisons (X3, as three comparisons were inspected for each data set) revealed that Within-predicate comparisons ("more P than that one") were ranked as significantly more natural with adjectival than nominal predicates both in the case of subjects denoting natural entity types (Adjectives: M = 5.57, SD = 1.9; Nouns: M = 3.3 ,SD = 2.1; t(53.8)

= 8.43, $p < .001$) and subjects denoting social entity types (adjectives: M = 6.4 , SD = 1.12; nouns: M = 4.58 ,SD = 2.04; $t(54.1) = 8.26$, $p < .001$), whereas *bet2s* comparisons ("more P than that one is Q") were ranked as equally natural with adjectival and nominal predicates both with subjects denoting natural entity types (adjectives: M = 2.63, SD = 1.6; nouns: M = 2.4, SD = 1; $t(53.8) = 1$, $p = .9$, uncorrected) and subjects denoting social entity types (adjectives: M = 2.96, SD = 1.75; nouns: M = 3, SD = 1.84; $t(53.9) = –.66$, $p = .51$, uncorrected).

However, bet1s comparisons ("more O than Q") were significantly more natural with adjectival than nominal predicates when subjects denoted natural entity types (adjectives: M = 4.35, SD = 2.00; nouns: M = 3.48, SD = 2.14; $t(53.5) = 3.35$, $p = .002$), but not social entity types (adjectives: M = 5.01, SD = 1.77; nouns: M = 5.22 ,SD = 1.93; $t(54.1) = .57$, $p = .57$, uncorrected).

In sum, within-predicate comparisons ("more P than that one") are natural for adjectives more than nouns, in line with a strict licensing restriction to dimension-counting scales, or a mere syntactic account that bans nouns. By contrast, between-noun comparisons bet1s ("more P than Q") are natural for adjectives and additive nouns more than multiplicative nouns, in line with a weaker licensing restriction to additive scales, not necessarily based on counting. It is less obvious to see how a syntactic account might explain the advantage of adjectives and nouns denoting human traits and artifacts over nouns denoting plants and animals. Finally, the *bet2s* (subdeletion) comparisons tested were neither natural with adjectives, nor nouns.[10]

2.2.4 Correlations between structures: Predictions and results

Last but not least, two separate analyses by subject-noun (n = 20) were conducted to test correlations between the ratings provided for the different structures, once for the nominal items and once for the adjectival items. The results are presented in Tables 4 and 5, respectively (for full tables see Appendices G-H).[11]

Two predictions were tested. First, if by default adjectival classification forms are quantificational, but nominal classification forms are similarity-based, the prediction is a

Spearman's rho		***Every***	***Most***	***Some***
Basic	r^3	**0.182**	**0.047**	**-0.135**
	p	0.443	0.843	0.572
Within	r^3	**.698****	**.855****	**.769****
	p	0.001	0.000	0.000
Bet1s	r^3	**.567****	**.801****	**.751****
	p	0.009	0.000	0.000
Bet2s	r^3	**.557***	**.711****	**.713****
	p	0.012	0.000	0.000

Table 4: Correlation coefficients and *p* values for the nominal items.
* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

[10] For comparisons between the naturalness of the different comparison structures see online Appendix J in https://sites.google.com/site/weidmansassoon/docs/general-documents/Online%20appendices.pdf.
[11] *p*-values with three zeros after the dot are indicated as 0.000. The contribution of any additional digits is negligent.

Spearman's rho		*Every*	*Most*	*Some*
Basic	r^3	**.634****	**.797****	**.721****
	p	0.003	0.000	0.000
Within	r^3	**.594****	**.661****	**.683****
	p	0.006	0.002	0.001
Bet1s	r^3	**.587****	**.676****	**.595****
	p	0.007	0.001	0.006
Bet2s	r^3	**.597****	**.584****	**.468***
	r^3	0.005	0.007	0.037

Table 5: Correlation coefficients and *p* values for the adjectival items.
* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

correlation between the naturalness of classification- and quantification-structures with **adjectives, but not nouns**.

As the first line of Tables 4 and 5 indicates, this prediction was borne out. Moderate to high correlations (marked in boldface) are exhibited between the quantification and basic structures with adjectival predicates, where coefficients were .63 ($p = .003$), .79 and .72 ($p < .001$), but not with nominal ones, where coefficients were very low ($p > .4$).

Second, if dimension accessibility is a predictor of licensing in both dimensional- and degree-structures, the prediction is a correlation between the naturalness of comparison- and dimensional-quantifier structures in **both adjectives and nouns**. In other words, if licensing in comparisons involves a shift to dimension counting and quantification (cf. the readings in (15)), correlations with dimensional-quantifier structures are expected.

As the second to fourth lines of Tables 4 and 5 indicate, this prediction was confirmed as well in all the 18 structure-pairs tested. Moderate to high correlations are exhibited between the quantification and comparison conditions both with nominal predicates (where Spearman coefficients range between .557 and .855, with $p < .000$ in 6 out of 9 cases), and adjectival ones (where coefficients ranged between .468 and .683, with $p \leq .007$ in 8 out of 9 cases).

Note also that null or even negative correlations were observed between the comparison or quantification structures and the control *small* structure (as the rightmost column of the table in Appendix G indicates, Spearman coefficients ranged between –.430 and –.628 with $p < .05$ in 4 out of 6 cases). This suggests that the results are **not** due to a putative strategy, which the participants might have developed, of automatically rating social items higher than natural ones. Rather, the participants attentively processed the sentences of the various structures and provided reliable ratings.[12]

3 Discussion

Degree constructions, and especially different types of comparisons, are extensively investigated within linguistics, and the literature on their syntax and semantics is wide ranging. The study presented in this paper aimed toward a more solid descriptive basis of data against which to test theories and improve them. In that sense, it is a preliminary

[12] Indeed, analyses indicating a reversed entity type effect in *small* structures are found in online Appendix K in https://sites.google.com/site/weidmansassoon/docs/general-documents/Online%20appendices.pdf. Additional indication for the reliability of the results comes from the fillers (see Appendix B); e.g., the average rating of the clearly bad sentence *#The inspector saw anything* was 1.88, as opposed to the clearly good sentence *The inspector didn't see anything*, which averaged rating was 6.32.

attempt to assess systematically the naturalness of different predicates in different degree constructions.

Moreover, the reported study assessed the naturalness of different predicates in different dimensional-quantifier constructions, in order to test predictions of a theory built of two main components. First, the lexical representations of adjectives and two types of nouns were assumed to reflect conceptual distinctions between quantificational and similarity-based (additive and multiplicative) dimension-binding. Quantification is by default based on counting, while similarity is not.

The link between quantification and counting goes back to generalized quantifier theory, where the interpretation of determiners is recast in terms of the cardinalities of the intersection and/or difference between the denotations of the arguments; e.g., *some boys walk* is true iff the cardinality of the intersection (set of walking boys) is at least 1, and *every boy walks* is true iff the cardinality of the difference (set of boys that do not walk) is 0 (Barwise & Cooper 1981). This also directly applies to quantification over respects.

Furthermore, Wattenmaker (1995) has shown that upon a shift of similarity-based concepts to a flattened representation of dimensions as binary and of equal weights, additive binding, unlike multiplicative binding, reduces to counting.

Second, according to the Accessibility Hypothesis, when the dimensions of predicates are accessible for binding by counting operations, these predicates are licensed in degree- and dimensional-quantifier constructions. In fact, counting and quantification have been hypothesized to dominate the interpretation of dimensional and degree constructions.

A link between comparison and counting has been created in the literature on quantity words such as *many* and *much* and their derived comparative *more* (Hackl 2001; Schwarzschild 2002; 2006; Solt 2009; Wellwood et al. 2012; Wellwood 2015); e.g., *more boys than girls arrived* directly relates to the cardinality of the arguments' sets (arriving boys and girls), but so is also *John smokes more,* which relates to cardinalities of events, and arguably *more optimistic*, which relates to dimension-set cardinalities.

Hence, nouns were hypothesized to be licensed in comparisons to the extent that they allow interpretations based on counting. A potential example of such an interpretation of a comparison between social nouns is (34) (Sassoon in press).

(34) COCA (Davies 2010-)
A sagging bunk bed, a straight chair with turned legs, a shelf of books, a sink, and a gas fired hotplate – *more a cell than a room.*

Fluctuations in naturalness seem to reflect the ease of shift to an interpretation based on counting. In additive nouns such a shift only requires a flattened representation of the dimensions. In multiplicative nouns it requires also a shift to additivity.

This highly refutable theory emerged from past research of categorization and dimensional exception phrases. The present study aimed to further test it by looking directly into the naturalness of dimensional-quantifier constructions, and degree constructions, considering the additive-multiplicative distinction in the nominal domain. Nine predictions were derived, tested, and confirmed, supporting the indicated theoretical components and elucidating some comparison-specific details. Section 3.1 and 3.2 summarize the results, and elaborate on a formal account which incorporates them into the standard framework for the analysis of gradability (Sassoon in progress).

Despite the confirmation of all the predictions of the theory, a cautious interpretation is recommended, because alternative accounts have not been directly ruled out, and the processing aspect, among other aspects of the account, has not been addressed. Various

directions for future research seem essential for these goals to be accomplished. Alternative accounts of morphological gradability are therefore considered in section 3.3 and important questions for future research are indicated.

3.1 Quantification on dimensions in classification constructions

The results suggest a positive response to our first research question. It was hypothesized that classification based on dimension-counting is a precondition for licensing respect arguments and quantifiers that bind them, as in *P in some/most/all respects*. Thus, these structures were predicted to be rated natural with adjectives (e.g., *safe*), slightly less natural with additive social nouns (e.g., *a church*), and even less natural with multiplicative natural-kind nouns (e.g., *a pine*) which do not become counting-based even upon a shift to a flattened representation of dimensions. Multiplicativity should also be overridden in favor of additivity. Classification constructions (e.g., "x is P") served as baselines, where licensing of predicates was not assumed to be dependent on dimension accessibility for binding by counting operations.

Indeed, **predicate-type effects** were observed in the structures involving explicit quantification over dimensions by *every, most* and *some*. multidimensional adjectives were judged more natural than nouns in these constructions, although they were judged equally natural in classification constructions. Moreover, **entity-type effects** were observed, i.e., additive nouns were judged more natural than multiplicative nouns in the quantification constructions, suggesting the relevance of the additive-multiplicative distinction to the semantics of dimensional quantifier constructions.[13]

These results undermine the competing standard account whereby neither adjectives nor nouns are associated with sets of multiple dimensions (adjectives are associated with a single dimensions and nouns with no dimensions at all). In this competing view, dimensional quantifier constructions should be costly with adjectives and nouns of all types alike, because in all cases a costly device for metalinguistic reference is need in order to access conceptual dimensions.

Rather the results support an account in which grammar associates multidimensional adjectives and social and natural-kind nouns with lexically encoded dimension sets and default dimension binding operations based on counting and additive vs. multiplicative similarity, respectively. Dimensional-quantifier constructions select counting-based concepts thus they are best with adjectives, and better with additive nouns than multiplicative nouns, where the shift toward dimension counting is bigger and thus more costly.

Finally, there were absolutely no correlations between the ratings of the nominal basic and quantificational structures (e.g., *a pine* vs. *a pine in some/ most/ every respect*), as opposed to moderate to high correlations between the ratings of the adjectival basic and quantificational structures (e.g., *safe* vs. *safe in some/ most/ every respect*). These results support an analysis of classification forms with multidimensional adjectives in terms of

[13] The multidimensional adjectives chosen for this study were systematically ranked higher when applied to social subjects than natural ones (see Figure 1, 4 and 7). Importantly, this was the case in all the seven adjectival conditions. Since this effect is not related to structure type and is even observed in the basic condition, it seems related to properties of the multidimensional adjectives chosen, which are orthogonal to the questions we ask.
By contrast, the difference between the social and natural nominal items was directly related to structure. While the basic constructions manifested no naturalness difference, the quantificational constructions manifested a social additive noun advantage. This interaction between structure and entity type in the nominal data (figures 5 and 8) and its absence in the adjectival data (figures 4 and 7) highlight the role of the social-additive/natural-multiplicative distinction for **nominal** naturalness in **quantificational** constructions.

quantification over dimensions (e.g., where *safe* means safe in all or most contextually relevant respects).

In such an account, adjectives can be associated with dimension-counting scales. An entity's degree reflects the number of dimensions in which it is within the norm. The norm determines how many dimensions are required for classification (e.g., all, most or some). Such an account conforms to conventional assumptions of the standard framework for the analysis of gradability (Cresswell 1976; von Stechow 1984; 2009; Kennedy 1999; 2007; Heim 2000; 2006; Hackl 2001; Kennedy & McNally 2005; Fox & Hackl 2006; Schwarzschild 2008; Rett 2008; 2015; Solt 2009; Beck 2011; Wellwood et al. 2012). In this framework, a gradable adjective like *tall* usually denotes the relation in (35a) between degrees d and individuals x who are at least d tall. A null constituent, *pos*, introduces a membership norm (standard) into classification forms, as in (35b), which is true if Jane's maximal height is at least as big as *tall*'s norm. A comparison like (35c) is analyzed as true if Jane's maximal height exceeds Bill's.

(35) a. [tall] = λdλx: **tall**(x) ≥ d
b. [Jane is pos tall] = 1 iff Max(λd.**tall**(j,d)) ≥ norm(**tall**).
c. [Jane is taller than Bill is] = 1 iff Max(λd.**tall**(j,d)) > Max(λd. **tall**(b,d)).

To model multidimensionality, assume that the relation *dimension of* associates each predicate (e.g., *bird*, *healthy*, *similar*), in each context, with a set of one **or more** predicates – its contextual dimensions (e.g., *has a beak*, *healthy with respect to the flu*, *similar in shape*). As dimensions are themselves predicates, they too denote degree relations. Let $Dim_{optimistic}$ symbolize the set of degree relations R denoted by dimensions of *optimistic*, whose classification norms are norm(R).

Optimistic denotes the **dimension-counting relation** in (36a), $R_{optimistic}$, that holds between a degree n and any entity x that is classified under at least n many dimensions of *optimistic*. A membership norm introduced via *pos*, norm($R_{optimistic}$), renders (36b) true if the number of dimensions under which Ann is classified is at least as high as the norm; e.g., Ann is optimistic in many, most or all dimensions. Since counting and quantification are introduced into these truth conditions by means of the traditional adjectival apparatus of scales and norms, gradability directly follows. Within-adjective comparisons such as (36c) are true if Ann is classified under more optimistic dimensions than Bill is. Between-adjective comparisons such as (36d) are true if Ann is classified under more *optimistic*- than *pessimistic*-dimensions.

(36) a. $R_{optimistic}$ = λnλx.|{R∈$Dim_{optimistic}$: Max(λd.R(x,d)) ≥ norm(R)}| ≥ n.
b. [Ann is pos optimistic] = 1 iff Max(λn.$R_{optimistic}$(a,n)) ≥ norm($R_{optimistic}$)
= 1 iff |{R∈$Dim_{optimistic}$: Max(λd.R(a,d)) ≥ norm(R)}| ≥ norm($R_{optimistic}$).
c. [Ann is more optimistic than Bill is] = 1 iff
Max(λn. |{R∈$Dim_{optimistic}$: Max(λd.R(a,d)) ≥ norm(R)}| ≥ n) >
Max(λn. |{R∈$Dim_{optimistic}$: Max(λd.R(b,d)) ≥ norm(R)}| ≥ n).
d. [Ann is more optimistic than pessimistic] = 1 iff
Max(λn. |{R∈$Dim_{optimistic}$: Max(λd.R(a,d)) ≥ norm(R)}| ≥ n) >
Max(λn. |{R∈$Dim_{pessimistic}$: Max(λd.R(a,d)) ≥ norm(R)}| ≥ n).

However, the results suggest that nominal classification forms merit a different account. The predicate type effect observed in the study is consistent with a syntactic account in terms of differing semantic types for adjectives and nouns that render adjectives readily

gradable, while requiring a costly type shift for nouns. However, an additional conceptual account would still be necessary to capture the entity-type effect observed, where additive nominal labels (of human traits and artifacts) scored higher than non-additive labels (of animals and plants). By parsimony (and considering that in two out of three comparison constructions, additive nouns were as natural as adjectives, cf., section 3.2), a conceptual account should be preferred where adjectives and nouns only differ in sort (but see discussion in section 3.2).

Hence, although a gradable account of nouns is less conventional, the experimental results appear to be in line with the view that nouns denote **additive and multiplicative similarity-based relations**, as in (37a, b). These relations are based on weighted sums or products of entities' dimensional degrees (Rosch 1973; Hampton 1998; Murphy 2002). Hence, nominal classification forms like (37c) involve the morpheme *pos* and are true if, e.g., Tweety's maximal degree of averaged similarity to a duck is at least as high as the norm. Additive nouns like *philosopher* are more similar to adjectives like *optimistic* than multiplicative nouns like *duck* are, because upon a shift c to a flattened representation of dimensions (with binary scales consisting of 0 and 1 and equal weights), they denote dimension-counting relations as in (37d).

(37) a. $[\text{philosopher}] = R_{\text{philosopher}} = \lambda d \lambda x.\ (\mathbf{w_{F1}f_{F1}(x)} + \ldots + \mathbf{w_{Fn}f_{Fn}(x)}) \geq d.$

b. $[\text{duck}] = R_{\text{duck}} = \lambda d \lambda x.\ (\mathbf{f_{F1}(x)^{WF1}} \times \ldots \times \mathbf{f_{Fn}(x)^{WFn}}) \geq d$

c. $[\text{Tweety is pos a duck}] = 1 \text{ iff } \text{Max}(\lambda d.\ R_{\text{duck}}(t,d)) \geq \text{norm}(R_{\text{duck}})$
$= 1 \text{ iff } \text{Max}(\lambda d.\ w_1 f_1(t) + \ldots + w_n f_n(t)) \geq d) \geq \text{norm}(R_{\text{duck}}).$

d. $[\text{philosopher}]_c = R_{\text{philosopher,c}} = \lambda x \lambda n.\ |\{R \in \text{Dim}_{\text{philosopher}}: \text{Max}(\lambda d.R(x,d) \geq 1\}| \geq n.$

Turning to dimensional constructions such as *optimistic {with respect to finding love, in every respect}*, since regular adjectives don't lexicalize a respect argument, a special operation WRT (the denotation of propositions like *with respect to*) must be used to introduce a dimension argument into the derivation. The results suggest that WRT only applies to dimension-counting relations, namely either adjectival denotations, or denotations of additive-nouns when shifted as in (37d). Thus, in (38a), WRT takes a dimensional degree relation, R (such as R_{love}) and a dimension-counting relation R_A (e.g., $R_{\text{Optimistic}}$), and returns R, providing that R is one of the dimensions generating R_A through counting (i.e., if there is a set of relations X of which R is a member and R_A is the relation between degrees n and entities x that classify under at least n many relations in X). Hence, the modified adjective in (38b) denotes the dimensional relation, R_{love}, providing that it is one of the dimensions generating the adjectival interpretation, $R_{\text{optimistic}}$. The (potentially implicitly) modified positive form in (38c) is true if Jane's optimism with respect to love is sufficiently high. The comparison in (38d) is true if Jane is more optimistic with respect to love than Bill is, and the between-adjective comparison in (38e) is true if Jane is more optimistic than pessimistic with respect to love.[14]

[14] Other degree constructions, such as *Bill is {somewhat, very, completely} optimistic*, seem similarly ambiguous. They can either mean that Bill is optimistic in {few, many, all} respects (if *optimistic* denotes a dimension-counting relation), or that Bill is {somewhat, very, completely} optimistic with respect to a contextually given dimension (e.g., finding love). Thus, experimentation is needed to test whether such constructions also exhibit dimension-binding effects.

(38) a. [with respect to] = WRT =
$\lambda R \lambda R_A$: $\exists X$, $R \in X$ & $R_A = (\lambda n \lambda x. |\{R \in X: \forall d \in I_R, R(x,d)\}| \geq n)$. R

b. [optimistic with respect to finding love] = $WRT(R_{love})(R_{optimistic}) = R_{love}$
(Providing that $R_{optimistic}$ is a dimension-counting relation and R_{love} is one of the relations generating it: $\exists X$, $R \in X$ & $R_{optimistic} = \lambda n \lambda x. |\{R \in X: Max(\lambda d.R(x,d)\}| \geq n)$.

c. [Jane is pos optimistic (about love)] = 1 iff $Max(\lambda d.R_{love}(d)(j)) \geq norm(R_{love})$.

d. [Jane is more optimistic (about love) than Bill is] = 1 iff $Max(\lambda d.R_{love}(j,d)) \geq Max(\lambda d.R_{love}(b,d))$.

e. [Jane is more optimistic than pessimistic (about love) = 1 iff $Max(\lambda d.R_{op\text{-}love}(j,d)) \geq Max(\lambda d.R_{pes\text{-}love}(j,d))$.

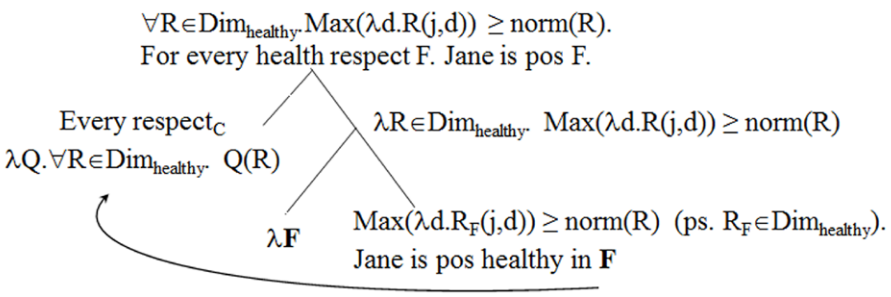
The PP complement may also include a dimensional quantifier, as in (39a), which is true if for all optimistic respects F, Jane is classified under F. As (39a) suggests, the first argument of *every* denotes a set of *respects*, namely, generators of some dimension-counting relation. However, quantification is contextually restricted to health respects (through the variable C; von Fintel 1994), as being a health respect is a presupposition of *every*'s scope introduced by *in*. The quantifier *every respect* is raised leaving a respect trace, $F_{<d,<e,t>>}$, in the PP adjoined to the adjective. Abstraction over the trace yields the second argument of *every*, which denotes the function in (39b) from degree relations R to truth if setting the value of F to R renders *Jane is pos healthy in F* true. The relation $[\textit{healthy in F}]_{F/R}$ is the relation $WRT(R)(R_{healthy})$, which by definition (39a) is the relation R (providing that $R_{healthy}$ is a dimension-counting relation generated by a set of relations that includes R; i.e., that $R \in Dim_{healthy}$.) Hence, the second argument of *every* denotes the set of health dimensions under which Jane is classified. The final truth conditions in (39c), then, require that Jane would be classified in every contextual health respect, as desired. The tree in (39d) illustrates the quantifier's raising.

(39) a. [Jane is pos healthy in every respect] =
$[Every]([respect_C]_{C/Dim_{healthy}})(\lambda R.[\text{Jane is pos healthy in F}]_{F/R})$.

b. $[\lambda F.$ Jane is pos healthy in F$]$ =
$= \lambda R \in Dim_{healthy}.Max(\lambda d.WRT(R)(R_{healthy})(j,d)) \geq norm(WRT(R)(R_{healthy}))$
$= \lambda R \in Dim_{healthy}.Max(\lambda d.R(j,d)) \geq norm(R)$.
(ps.: $R_{healthy}$ is a dimension-counting relation and $R \in Dim_{healthy}$).

c. [(39a)] = 1 iff $\forall R \in Dim_{healthy}.Max(\lambda d.R(j,d)) \geq norm(R)$.

d.



An (implicitly) modified comparison like (40) is true if in {some, many, most, all} optimistic respects F, Ann is more optimistic than Bill is about F.[15]

(40) [Ann is (generally, in most ways) more optimistic than Bill is] =
[Generally]([respect$_c$])(λR.[Ann is more optimistic (wrt F) than Bill is]$_{F/R}$
= 1 iff Generally for R∈Dim$_{optimistic}$,
Max(λd.WRT(R$_{optimistic}$,R)(d)(a)) > Max(λd.WRT(R$_{optimistic}$,R)(d)(b)).
= 1 iff Generally for R∈Dim$_{optimistic}$, Max(λd.R(a,d)) > Max(λd.R(b,d)).

3.2 Quantification over dimensions in comparative constructions

An additional research question was whether dimension accessibility for binding by counting operations predicts naturalness of predicates in degree constructions, namely morphological gradability, as the Accessibility Hypothesis predicts, or not. An intuitive competing alternative is the counter-hypothesis whereby comparatives with multidimensional predicates predominantly compare averaged degrees of, e.g., health or optimism, of entities, not dimension-counts. This hypothesis conforms to the more standard view that adjectives are lexically associated with a unique dimension, rather than a whole set, and nouns are not associated with scalar dimensions at all. Thus, we tested whether, being additive, social nouns are more natural in comparison structures ("more P than {that one, Q, that one is Q}") than natural-kind nouns, which are generally multiplicative.

The results suggest a positive answer to this question as well. **Entity-type effects** were observed in all the comparison constructions, as predicted assuming that degree constructions as well involve counting of, accessing and quantification over dimensions (as in, e.g., (36c, d), (38c, d), and (40), respectively). Only in the nominal data, the social-noun advantage was related to structure, occurring in comparisons but not in classification forms, thus indicating the role of the social-additive vs. natural-multiplicative distinction in predicting naturalness of **nominal comparisons**. Confirmation of this prediction highlights the role of dimension-binding (dimension accessibility) in predicting the naturalness of degree constructions, over and above the role of the noun/adjective and –/+human features.

Furthermore, **predicate type effects (adjectives > nouns)** were only observed in two out of the three comparison structures tested (no category differences were observed in subdeletion comparisons ("x is more P than y is Q"), and in particular, differences between social adjectives (adjectives applied to human or artifact entities) and social additive nouns (labels of human and artifact concepts) were only observed in one comparison structure, the within-predicate comparison ("x is more P than y"). These results undermine the importance of lexical category (noun-adjective) in determining acceptability. They stress the role of conceptual structure. Moreover, they suggest that comparisons with multidimensional predicates actually tend to have readings based on dimension-counting, and therefore comparisons are more natural with nouns whose interpretation more readily involves dimension-counting.

Further support for the Accessibility Hypothesis and detailed theoretical interpretation presented in section 3.1 comes from the moderate to high correlations between the naturalness of all the comparison types tested and all the dimensional-quantifier structures tested, especially in nouns. This finding directly supports the view that nouns are natural in comparisons to the extent that quantification over their dimensions is natural supports

[15] The complex trace type (<d,<e,t>>) is supported by the fact that *optimistic wrt F*, which denotes the semantic value of F, is degree-modified.

the view the dominant readings of comparisons involve counting and quantification over dimensions. Alternative explanations for the higher naturalness of comparisons with social than natural nouns, which ignore the role of dimensions, do not seem capable of explaining the tight correlations observed.[16]

Nonetheless, although the proposed account explains the data, alternative account can in principle be possible. While it is hard to see how models that reject the roles of dimensions and their binding would predict the correlations (but see next section for discussion), various dimension-models co-exist in psychology (Photos & Wills 2011). On one alternative account, both social and natural-kind nouns are multiplicative (Medin & Schafer 1978), but they differ in the strictness of their norms. Thus, one alternative explanation is that natural-kind nouns are not licensed in comparisons because their instances can only manifest small deviations from the prototype. While this account has to be examined in the future, notice that strict standards normally do not block licensing in comparisons; e.g., while *full* and *empty* tend to have maximal standards, *fuller* and *emptier* are perfectly natural (Kennedy 1999; 2007; Kennedy & McNally 2005).

Furthermore, the present study revealed correlations between licensing in dimensional and comparison constructions, but future experimentation is needed to test any claims for causality. That is, in order to test whether some nouns are licensed in degree constructions *because* they are additive, other potential reasons should be neutralized, including, e.g., vagueness, subjectivity, or any other factor that may distinguish between social and natural-kind nouns (for a discussion of such differences, see Keil 1986; Kalish 1995; Gelman 2003; Estes 2004; Sassoon 2017; in press). To eliminate alternative influencing factors, pseudo-nouns with additive and non-additive conceptual structures can be used.

Notice also that this study included only nouns and positive multidimensional adjectives. Its results confirmed those of a previous study with positive, one-dimensional, context relative adjectives (e.g., *big*; Sassoon 2017). Future studies should also consider negative adjectives (like *sick*), one-dimensional absolute adjectives (like *empty* and *closed*), and predicates of additional semantic domains, for instance, abstract nouns (like *problem* or *love*). Of special interest are additive natural-kind nouns or multiplicative social nouns, if such exist, and nouns like *flu, depression* and *schizophrenia*, which may have an intermediate status between the natural and social ends. Future research should address them theoretically and experimentally.

Additional degree constructions should also be examined; e.g., the role of mediating particles like *of* in *more of a linguist* should be addressed. Although *more of* comparisons are more natural than comparisons without *of*, they are predicted to exhibit an additive-noun advantage and correlations with quantified constructions, because they as well seem intuitively to involve access to dimensions of nouns and their counting.

But above all, research should address the online processing of multidimensional gradability. According to the Accessibility Hypothesis, the same problem causes the reduction of naturalness of nominal predicates in dimensional and degree constructions, namely difficulty to access and count or quantify over the dimensions of predicates that do not denote dimension-counting relations. Thus, future research should test a predicted correlation between the online processing time of nominal predicates in dimensional and degree constructions (Fadlon et al. in progress), and a predicted delay in the processing

[16] The slightly higher correlations in nouns than adjectives may suggest that dimension accessibility for binding by counting operations is a main predictor of naturalness in nouns, while in adjectives, where all items are readily counting-based, the effects of other irrelevant factors on naturalness of comparisons as opposed to quantification structures reduce the correlations. At any rate, the correlations are still rather strong.

of additive nouns due to the shift necessary for their licensing, in comparison with adjectives and multiplicative nouns, where no shift seems to be needed or successfully used, respectively.

The next section considers alternative accounts related to specific comparison types. It includes provisional observations about open questions and directions for their future research.

3.3 A closer look at specific constructions: Conceptual vs. syntactic accounts

Considering ordinary degree constructions first, within-noun comparisons such as those in (41) favor adjectives over nouns.

(41) *Within-predicate comparisons* (comparisons of two entities in a single predicate)
a. This woman is more optimistic than Bill.
b. ?This scholar is more a linguist than that one.
c. #This bird is more a duck than that bird.

The contrast between (41a) and (41b, c) may stem from a syntactic account in which nouns do not have degree arguments, or from a conceptual account with gradable similarity-based nominal semantics. Unlike the former, the latter view explains the fact that examples with multiplicative nouns like (41c) are judged worse than examples with additive nouns like (41b). But a syntactic constraint is not ruled out here, given the pronounced differences between social nouns and adjectives in this construction.

Syntactic restrictions admit high levels of cross-linguistic variation, while semantic aspects are more stable across languages. Thus, the need for experimental research of dimension accessibility in different languages emerges (see, e.g., Fadlon et al. in progress). Neeleman, van de Koot & Doetjes (2004) and Doetjes (2008) consider the hypothesis that, e.g., intensifiers like *very*, which are restricted to the adjectival domain, are syntactic heads that select for an adjective, while similar intensifiers in other languages, which apply across lexical categories, are adjuncts. Within-predicate comparisons do license different lexical categories (Wellwood et al. 2012; Wellwood 2014).

Within-predicate comparisons contrast with comparisons between-predicates like (42a–c), which more easily license predicates that are typically considered non-gradable like nouns (Constantinescu 2011; Giannakidou & Yoon 2011; Morzycki 2011; 2012). Their accounts often resort to *metalinguistic scales*; e.g., (42b) seems to imply that it is more appropriate or preferable to call the relevant scholar *a linguist* than *a philosopher*, possibly by virtue of completely subjective contextual interests and desires of the discourse participants (Giannakidou & Yoon 2011). (42b) further implies that (43a) is closer to the truth than (43b), in the sense of Lasersohn's (1999) account of imprecision (Morzycki 2011).

(42) *Between-predicate comparisons* (comparisons of one entity in two predicates)
a. This woman is more optimistic than pessimistic.
b. This scholar is more a linguist than a philosopher.
c. ?This bird is more a duck than a goose.

(43) a. This scholar is a linguist.
b. This scholar is a philosopher.

A metalinguistic account can take dimensions of nouns to have a cognitive role, but no grammatical or truth-conditional role (cf., Lewis 1970; Kripke 1972; Putnam 1975; Kamp & Partee 1995; Kennedy 1999; Baker 2003). However, metalinguistic comparison, by definition, can access non-linguistic features of predicates, i.e., bring the linguistically inert role of the nominal dimensions and their binding operations into the semantics of a sentence. Thus, metalinguistic comparatives behave as if nouns have linguistically accessible dimensions. Moreover, the lack of correlation between responses to the nominal basic and the nominal quantificational condition is expected.

An alternative metalinguistic account may reject any role for conceptual distinctions assuming that comparisons like (42a–c) involve solely ad-hoc contextual scales of, e.g., speakers' subjective attitudes to prepositions. Natural-kind nouns could then be regarded as less natural than social nouns and adjectives in this construction (where the observed advantage of adjectives over nouns stems solely from them) because they have relatively precise conventional interpretations that render ad-hoc subjective attitude scales less appropriate.

However, any account in terms of metalinguistic reference fails to explain the generality of the observations. Entity type effects occurred in all comparison types, including the most ordinary within-predicate comparison (e.g., (41)), where correlations with dimensional quantifier constructions were also quite high.

A different viable option is that the between-predicate Bet1s comparison in (42) does not involve metalinguistic interpretations any more than any other ordinary construction. Rather, it has a more relaxed licensing constraint, which only requires additivity (Sassoon in press). Thus, additive nouns and adjectives are equally natural, and multiplicative nouns are worse

This account captures the fact that between-noun comparisons have readings involving degrees of similarity to two nominal prototypes, where similarity is not reduced to dimension-counting (cf., Lakoff 1987; de Vries 2010; Bochnak 2010; Morzycki 2011); e.g., a person, who is closer to the psychologist's prototype than the linguist's only due to few high-weight features, like PhD in psychology, may count as *more a linguist than a psychologist* by virtue of resembling a linguist more on many low weight features, thus having a higher number of properties of linguists than of psychologists (a dimension-counting reading).[17] But a person not strictly having any properties of linguists (say a philosopher or a child) may still count as *more a linguist than a psychologist* due to higher similarity to the linguist's prototype (a similarity-based reading). Being closer to that prototype, *linguist* is a preferred label for this person, and a more precise one than *psychologist*.

Indeed, comparisons like (44a) seem less marked than clearly metalinguistic comparisons like (44b), where it is necessary to resort to an ad-hoc scale based on speaker attitudes toward propositions because *tall* and *intelligent* are one-dimensional open-scale adjectives that do not seem to have prototypes (Klein 1991). If nouns denote similarity relations (as in (37)), such a resort is possible but not necessary in (44a).[18]

(44) a. This car is more a taxi than an ambulance.
b. ?Dan is more tall than Ram is intelligent.

[17] Readings involving stereotypical features may actually emerge from shifts to dimension-counting. Such a shift involves treating the dimensions as equally important, thus raising the weight of minor features on account of the weights of important features. More research is needed to better understand these cases.

[18] Decompositional accounts of *more* as *much-er* constrain the type of scales with which *much* and thus *more* can be contextually associated, potentially ruling out similarity scales (Hackl 2001; Schwarzschild 2006; Solt 2009; Wellwood et al. 2012). But similarity predicates are odd also with degree morphemes that do not derive from *much* (e.g., *very, slightly* and *too*), and constraints on *much* do not straightforwardly distinguish different types of comparison (but see in-depth discussion in Wellwood 2014).

Furthermore, similarity scales usually include a maximum and dimension-counting scales always include a minimum, in line with Doetjes's (2008) suggestion that degree semantics outside the adjectival domain always involves **closed** scales. Degree words which are restricted to adjectives like *very* are actually restricted to vague predicates whose scales are open.

In sum, syntactic and semantic differences between ordinary within-predicate comparisons and between-predicate comparisons exist (Morzycki 2011), but they cannot explain the main results of this study, because differences between natural and social nouns and correlations with quantifier constructions occurred in all the tested comparison constructions. All in all, the results seem to support the view that nouns can denote at the type of degrees. Otherwise, the fact that additive classification criteria predict an increase in the naturalness of a noun with degree morphology in **both** within- and between-predicate comparisons (including subdeletion comparisons) would remain unexplained. At the same time, future experimentation is much needed to characterize the set of readings of each construction and their competing accounts more directly and systematically.[19]

Finally, in subdeletion *bet2s* comparisons, adjectives and nouns were rated similarly and lower than in other comparisons. These results, however, should be viewed with caution. First, the *bet2s* sentences were the longest. Future studies of comparison types should aim to balance sentence length in different comparisons. Second, additional factors may have lowered the rankings of adjectives. In particular, Büring (2007) suggests that one-dimensional structures like (45a) are less natural than reversed structures like (45b). The *bet2s* sentences of the present study, e.g., (45c), resembled (45a). Their average ratings was 2.8 (1.67), as opposed to 2.7 (1.73) in nouns. The minimal pair in (45a,b) was actually included among the fillers. Its average ratings were 2.48 and 3.33, respectively, suggesting that either way naturalness was low, but in order to fully understand the status of adjectives as opposed to nouns, future study should also assess the naturalness of reversed structures, possibly also with negative nouns, like *fool*.

(45) a. #This house is higher than that ladder is short.
b. This ladder is shorter than that house is high.
c. This car is more interesting than that one is dull.
d. This car is more interesting than dull.

4 Conclusions

Notwithstanding the questions left open, the current study examined the naturalness of three comparison conditions, and compared it with the naturalness of three conditions involving quantification over dimensions. In addition, it included adjectives and two types of nouns as items and considered the differences between them. Furthermore, it aimed toward a unified view that takes into account both the cognitive psychological perspective, which is based on experimental research of concepts denoted by adjectives and nouns, and the semantic perspective, which is based on theoretical and empirical work on degree constructions, quantification, nouns and adjectives. The unified view has been developed into a testable theoretical proposal in Sassoon (2017; in progress; in

[19] Another possibility that merits investigation is that between-predicate comparisons tend to involve comparisons of number of dimensions (e.g., x is more optimistic than pessimistic if x is optimistic in more dimensions), whereas ordinary within-predicate comparisons tend to involve quantification (x is more optimistic than y is if in {all, most, some} dimensions F, x is more F than y.) Such readings require gradable dimensions. This requirement is met by adjectives, but not by additive nouns, where a shift to a binary representation takes place.

press), whose predictions were confirmed. Significant correlations and interactions were obtained, and only where they were predicted.

The main result was that the naturalness of a given noun or adjective in comparisons and dimensional-quantifier constructions tightly correlates, suggesting that the accessibility of dimensions for binding by counting and quantifying operations is an important predictor of morphological gradability. Hence, evidence was provided suggesting the relevance of dimension-binding operations to the grammatical status of constructions that do not involve explicit counting or quantification over dimensions, but seem to manifest them in their truth conditions. Much experimental and theoretical work is needed to address the implications of the findings, the generality of the phenomena, and the viability of their various alternative accounts. Hopefully, raising these issues will help to improve our understanding of morphological gradability and its underlying cognition.

Additional File

The additional file for this article can be found as follows:

- **Appendices.** DOI: https://doi.org/10.5334/gjgl.155.s1

Acknowledgements

This article has profited from invaluable advice given by the editor Chung-Hye Han and the anonymous referees. Special thanks to Petra Schumacher and to the audience of the McGill / MIT Workshop on Gradability and Quantity in Language and the Brain, including especially Irene Heim and Kai von Fintel for helpful critical remarks. Many thanks for financial support from the GIF, the German-Israeli Foundation for Scientific Research and Development.

Competing Interests

The authors have no competing interests to declare.

References

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511615047

Barr, Dale J., Roger Levy, Christopher Scheepers, & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Bartsch, Renate. 1986. Context-dependent interpretations of lexical items. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Foundations of pragmatics and lexical semantics* (GRASS 7), 1–26. Dordrecht: Foris.

Barwise, Jon & Robyn Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4. 159–219. DOI: https://doi.org/10.1007/BF00350139

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

Bochnak, Ryan. 2010. Quantity and gradability across categories. In *Proceedings of SALT 20*, 251–268. Ithaca, NY: CLC Publications. DOI: https://doi.org/10.3765/salt.v20i0.2570

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86. 186–213. DOI: https://doi.org/10.1353/lan.0.0189

Buhrmester, Michael, Tracy Kwang, & Samuel D. Gosling. 2011. Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science* 6. 3–5. DOI: https://doi.org/10.1177/1745691610393980

Büring, Daniel. 2007. Cross-polar nomalies. In *Proceedings of Semantics and Linguistic Theory (SALT) 17*. Ithaca, NY: CLC Publications. DOI: https://doi.org/10.3765/salt.v17i0.2957

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

Constantinescu, Camelia. 2011. *Gradability in the nominal domain*. Leiden: Rijksuniversiteit Leiden dissertation.

Cresswell, Max J. 1976. The semantics of degree. In Barbara Partee (ed.), *Montague Grammar,* 261–292. New York, NY: Academic Press. DOI: https://doi.org/10.1016/b978-0-12-545850-4.50015-7

Davies, Mark. 2010. Corpus of contemporary American English (http://corpus.byu.edu/coca). Utah: Brigham Young University.

de Vries, Hanna. 2010. *Evaluative degree modification of adjectives and nouns*. Utrecht: Utrecht Institute of Linguistics MA thesis.

de Vries, Hanna. 2015. *Shifting sets, hidden atoms – the semantics of distributivity, plurality and animacy*. Utrecht: Utrecht University.

Doetjes, Jenny. 2008. Adjectives and degree modification. In Louise McNally & Chris Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics and discourse,* 123–155. Oxford: Oxford University Press.

Estes, Zachary. 2004. Confidence and gradedness in semantic categorization: Definitely somewhat artifactual, maybe absolutely natural. *Psychonomic Bulletin & Review* 11. 1041–7. DOI: https://doi.org/10.3758/BF03196734

Fadlon, Julie, Galit W. Sassoon, & Petra Schumacher. In progress. Discrete dimension accessibility in multidimensional concepts. Ms. University of California San Diego, Bar Ilan University.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29. 537–86. DOI: https://doi.org/10.1007/s10988-006-9004-4

Gelman, Susan A. 2003. *The essential child: Origins of essentialism in everyday thought.* Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195154061.001.0001

Giannakidou, Anastasia & Suwon Yoon. 2011. The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language and Linguistic Theory* 29. 621–655. DOI: https://doi.org/10.1007/s11049-011-9133-5

Grinsell, Timothy Wood. 2013. Avoiding predicate whiplash: Social choice theory and linguistic vagueness. In Anca Chereches (ed.), *Semantics and linguistic theory* 22. 424–440.

Hackl, Martin. 2001. *Comparative quantifiers and plural predication*(WCCFL XX). Somerville, MA: Cascadilla Press.

Hampton, James A. 1995. Testing the prototype theory of concepts. *Journal of Memory and Language* 34. 686–708. DOI: https://doi.org/10.1006/jmla.1995.1031

Hampton, James A. 1998. Similarity based categorization and fuzziness of natural categories. *Cognition* 65. 137–65. DOI: https://doi.org/10.1016/S0010-0277(97)00042-5

Hampton, James A, Gert Storms, Claire L. Simmons, & Daniel Heussen. 2009. Feature integration in natural language concepts. *Memory and cognition* 37(8). 1150–1163. DOI: https://doi.org/10.3758/MC.37.8.1150

Hampton, James A & Yoad Winter. In progress. *Compositionality and concepts in linguistics and psychology*. Forthcoming in Springer.

Heim, Irene. 2000. Degree operators and scope. In *Proceedings of SALT* 10. Ithaca, NY: CLC Publications. DOI: https://doi.org/10.3765/salt.v10i0.3102

Heim, Irene. 2006. Remarks on comparatives as generalized quantifiers. Cambridge, MA: MIT Press.

Hoeksema, Jack. 1995. The semantics of exception phrases. In Jan van Eijck & Jaap van der Does (eds.), *Quantifiers, logic, and language*, 145–77. Stanford, CA: Stanford University Press, CSLI.

Kalish, Charles W. 1995. Essentialism and graded membership in animal and artifact categories. *Memory & Cognition* 23. 335–353. DOI: https://doi.org/10.3758/BF03197235

Kamp, Hans. 1975. Two theories about adjectives. In Edward L. Keenan (ed.), *Formal Semantics of natural language*, 123–155. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/cbo9780511897696.011

Kamp, Hans & Barbara Partee. 1995. Prototype theory and compositionality. *Cognition* 57. 129–191. DOI: https://doi.org/10.1016/0010-0277(94)00659-9

Keil, Frank C. 1986. The acquisition of natural kind and artifact terms. In William Demopoulos & Ausonio Marras (eds.), *Language learning and concept acquisition: Foundational issues*, 133–153. Norwood, NJ: Ablex.

Kennedy, Chris. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York, NY: Garland.

Kennedy, Chris. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45. DOI: https://doi.org/10.1007/s10988-006-9008-0

Kennedy, Chris & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381. DOI: https://doi.org/10.1353/lan.2005.0071

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

Kuznetsova, Alexander, Per Bruun Brockhoff, & Rune Haubo Bojesen Christensen. 2014. lmertest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). http://CRAN.R-project.org/package = lmerTest.

Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the Mind*. Chicago, IL: Chicago University Press. DOI: https://doi.org/10.7208/chicago/9780226471013.001.0001

Landman, Fred. 1989. Groups II. *Linguistics and Philosophy* 12. 723–44. DOI: https://doi.org/10.1007/BF00632603

Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75. 522–551. DOI: https://doi.org/10.2307/417059

Mahowald, Kyle, Peter Graff, Jeremy Hartman, & Edward Gibson. 2016. SNAP judgments: A small N acceptability paradigm (SNAP) for linguistic acceptability judgments. *Language* 92(3). 619–635. DOI: https://doi.org/10.1353/lan.2016.0052

McCloskey, Michael E & Sam Glucksberg. 1978. Natural categories: Well defined or fuzzy sets?. *Memory & Cognition* 6. 462–72. DOI: https://doi.org/10.3758/BF03197480

McNally, Louise & Isidora Stojanovic. 2015. Aesthetic adjectives. In James Young (ed.), *The semantics of aesthetic judgmen*t. Oxford: Oxford University Press.

Medin, Douglas L & Marguerite M. Schaffer. 1978. Context theory of classification learning. *Psychological Review* 85 (3). 207–238. DOI: https://doi.org/10.1037/0033-295X.85.3.207

Moltmann, Frederique. 1995. Exception phrases and polyadic quantification. *Linguistics and Philosophy* 18. 223–80. DOI: https://doi.org/10.1007/BF00985445
Morzycki, Marcin. 2009. Degree modification of gradable nouns: Size adjectives and adnominal degree morphemes. *Natural Language Semantics* 17(2). 175–203. DOI: https://doi.org/10.1007/s11050-009-9045-7
Morzycki, Marcin. 2011. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics* 19(1). 39–86. DOI: https://doi.org/10.1007/s11050-010-9063-5
Morzycki, Marcin. 2012. The several faces of adnominal degree modification. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL)*. Somerville, MA: Cascadilla Proceedings Project.
Murphy, Gregory. 2002. *The big book of concepts*. Cambridge, MA: MIT Press.
Neeleman, Ad, Hans van de Koot, & Jenny Doetjes. 2004. Degree expressions. *The Linguistic Review* 21. 1–66. DOI: https://doi.org/10.1515/tlir.2004.001
Nouwen, Rick. 2011. Degree modifiers and monotonicity. In Paul Egre & Nathan Klinedinst (eds.), *Vagueness and Language Use*. Basingstoke: Palgrave MacMillan. DOI: https://doi.org/10.1057/9780230299313_7
Rett, Jessica. 2008. *Degree modification in natural language*. New York, NY: The State University of New York dissertation.
Rett, Jessica. 2015. *The semantics of evaluativity*. Oxford: Oxford University Press.
Rosch, Elinor. 1973. Natural categories. *Cognitive Psychology* 4 (3). 328–50. DOI: https://doi.org/10.1016/0010-0285(73)90017-0
Rosch, Elinor & Caroline B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7. 573–605. DOI: https://doi.org/10.1016/0010-0285(75)90024-9
Roth, Emilie & Edward J. Shoben. 1983. The effect of context on the structure of categories. *Cognitive Psychology* 15. 346–378. DOI: https://doi.org/10.1016/0010-0285(83)90012-9
Sassoon, Galit W. 2012. The double nature of negative antonymy. *The Proceedings of Sinn und Bedeutung* 16. Utrecht: Utrecht University Press.
Sassoon, Galit W. 2013a. *Vagueness, gradability, and typicality, the interpretation of adjectives and nouns*. Leiden: Brill. DOI: https://doi.org/10.1163/9789004248588
Sassoon, Galit W. 2013b. A typology of multidimensional adjectives. *Journal of Semantics* 30. 335–380. Sassoon, Galit W. 2017. Comparisons of nominal degrees. *Language* 93 (1). 1–36.
Sassoon, Galit W. 2017. Dimension accessibility as a predictor of morphological gradability. In James A. Hampton & Yoad Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology*. Springer.
Sassoon, Galit W. In progress. *Multiudimensionality in the grammar of gradability*. Bar Ilan university.
Schwarzschild, Roger. 2002. The grammar of measurement. In Brendan Jackson (ed.), *Proceedings of SALT XII*, 225–245. Ithaca, NY: CLC Publications. DOI: https://doi.org/10.3765/salt.v12i0.2870
Schwarzschild, Roger. 2006. The role of dimensions in the syntax of noun phrases. *Syntax* 9(1). 67–110. DOI: https://doi.org/10.1111/j.1467-9612.2006.00083.x
Schwarzschild, Roger. 2008. The semantics of comparatives and other degree constructions. *Language and Linguistics Compass* 2.2. 308–331.
Shamir, Adi. 2013. Disjunctivity. Ms.,The Hebrew University of Jerusalem.

Solt, Stephanie. 2009. The semantics of adjectives of quantity. New York NY: City University of New York.

Solt, Stephanie. 2018. Multidimensionality, subjectivity and scales: Experimental evidence. In Elena Castroviejo, Galit Weidman Sassoon, and Louise McNally (eds.), *The Semantics of gradability, vagueness, and scale structure, experimental perspectives*. (Language, Cognition, and Mind series). Switzerland: Springer.

Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115. 1497–1524. DOI: https://doi.org/10.1016/j.lingua.2004.07.002

Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1. 118–129.

Sprouse, Jon. 2013. Acceptability judgments. In Mark Aronoff (ed.), *Oxford Bibliographies Online: Linguistics*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/obo/9780199772810-0097

Tversky, Amos. 1977. Dimensions of similarity. *Psychological Review* 84. 327–52. DOI: https://doi.org/10.1037/0033-295X.84.4.327

von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Cambridge, MA: University of Massachusetts dissertation.

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

von Stechow, Arnim. 2009. The temporal adjectives früh(er)/spät (er) and the semantics of the positve. In Anastasia Giannakidou & Monika Rathert (eds.), *Quantification, definiteness, and nominalization*, 214–233. Oxford: Oxford University Press.

Wattenmaker, David W. 1995. Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology* 28. 274–328. DOI: https://doi.org/10.1006/cogp.1995.1007

Wellwood, Alexis. 2015. On the semantics of comparison across categories. *Linguistics and Philosophy* 3. DOI: https://doi.org/10.1007/s10988-015-9165-0

Wellwood, Alexis, Valentine Hacquard, & Roumyana Pancheva, 2012. Measuring and comparing individuals and events. *Journal of Semantics* 29(2). DOI: https://doi.org/10.1093/jos/ffr006

How to cite this article: Sassoon, Galit Weidman and Julie Fadlon. 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, DOI: https://doi.org/10.5334/gjgl.155

Published: 09 May 2017

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS