Investigating gradable predicates, comparison, and degree constructions in underrepresented languages*

M. Ryan Bochnak\textsuperscript{1} and Elizabeth Bogal-Allbritten\textsuperscript{2}

\textsuperscript{1}University of California, Berkeley; \textsuperscript{2}University of Massachusetts, Amherst

bochnak@berkeley.edu; eba@linguist.umass.edu

November 19, 2013

1 Introduction

The meaning of gradable predicates, comparison, and related constructions has been a topic of interest for linguists for decades. Many key insights into this domain were noticed as early as Sapir (1944), and researchers in formal semantic frameworks have since made great progress in our understanding of the relevant facts in this area (e.g., Kamp 1975, Cresswell 1976, Klein 1980, von Stechow 1984, Heim 1985, Bierwisch 1989, Kennedy 1999, Kennedy & McNally 2005, among many others). Additionally, typological studies including Ultan (1972), Stassen (1985), and Bobaljik (2012) have explored the morphosyntactic landscape of comparison and gradability.

\*To appear in Methodologies in Semantic Fieldwork, M. Ryan Bochnak and Lisa Matthewson (eds.), New York: Oxford University Press. We thank Washo consultants Ramona Dick and Steven James, and Navajo consultants Ellavina Perkins and Louise Kerley. We would also like to thank Robert Henderson and Lisa Matthewson for comments on a previous version that helped us improve this chapter. MRB’s work was supported by grants from the Jacobs Fund of the Whatcom Museum, the American Philosophical Society, and the National Science Foundation under Grant #1155196. EBA's work was supported by a National Science Foundation Graduate Research Fellowship.
in the world’s languages.

Only recently, however, have the theoretical proposals, based largely on data from English and German, been tested by formal semanticists researching understudied and typologically disparate languages in detail (e.g., Beck et al 2009). First steps towards integrating data from underrepresented languages into theories of comparison and gradability can be found in studies such as Galant (1998), Bogal-Allbritten (2008, 2013), Beck et al (2009), Eckardt (2009), Pearson (2010), Bochnak (2012, 2013a,b), Hohaus (2012), and Francez & Koontz-Garboden (2013). This semantic domain is thus relatively under-represented within the small but growing field of formal semantic fieldwork on understudied languages.

Probably the biggest challenge facing the fieldworker investigating comparison and gradability is the fact that the meaning of gradable predicates is highly context-dependent. A single individual may count as tall in some contexts, but not in others, all the while maintaining the same absolute height. A further complication is the fact that gradable predicates are not a uniform class. Variables such as scale structure, ordering polarity, and whether a predicate is associated with a measurable dimension have important consequences for entailment patterns and the distribution of modifiers, among other things. These factors therefore need to be taken into consideration when investigating constructions involving gradable predicates.

In this chapter, we discuss these issues and offer methods for eliciting certain key data in the field. The discussion is based on our own experiences investigating the semantics of gradability and comparison in Washo and Navajo.¹ Using case studies from these two languages, we propose methodologies for obtaining data on semantic distinctions that have been highlighted in the theo-

---

¹Washo is highly endangered, spoken by about 10 elderly native speakers in the Lake Tahoe area of California and Nevada. It is typically considered a language isolate, though it has been linked to the proposed Hokan group of languages (Mithun, 1999). English was used as the contact language for this work. Navajo is a member of the Southern (or Apachean) subfamily of the Athabaskan language family. Navajo is spoken by between 120,000 and 170,000 individuals based in the southwestern United States. All Navajo consultants are fully bilingual in English, and English was used as the contact language. For discussion of contact language choice, see AnderBois and Henderson (this volume) and Matthewson (2004).
retical literature and briefly outline how the data obtained using those methodologies fit into the formal semantic frameworks developed for English and other languages.

The chapter is organized as follows. In section 2, we review Kennedy & McNally’s (2005) theory of scale structure and scalar modifiers, and discuss data from Washo and Navajo, showing that scalar modifiers are not necessarily sensitive to all or only those aspects of scale structure identified as linguistically relevant by Kennedy & McNally. In section 3, we discuss the notions of norm-relatedness and crisp judgements, and propose that certain visual stimuli can be effective in eliciting crucial contrasts in the target language. We suggest that use of visual stimuli can help obviate some problems posed by the inherent context sensitivity of gradable predicates. Section 4 also addresses norm-relatedness and its interaction with lexical competition, where we argue for certain non-visual elicitation techniques for eliciting key data. We summarize and conclude in section 5.

Before continuing, we would like to flag from the outset that this chapter does not contain a questionnaire of sentence types that must be collected by a fieldworker investigating gradable predicates and related constructions in their research language. Rather, the chapter is largely focused on methods for digging deeper into the semantics of these constructions, once a range of sentence types has been collected. For a questionnaire that outlines the basic sentence types to collect, see Beck et al (2009), especially p. 13 and the appendices.

2 Adjective classes and modifier licensing

In this section, we review adjective classes that have been argued in the theoretical literature to be linguistically relevant, for example in their ability to license scalar modifiers. We then present case studies from modifiers in Washo and Navajo that are translated into English as very, but which have different distributions from their English counterpart.
2.1 Methods of classifying adjectives

Much of the literature on the semantics of gradable adjectives has focused on the fact that a subset of them, and indeed the most prototypical of them, are vague in the bare form.\textsuperscript{2} Sentences with vague predicates are subject to contextual variability in their truth conditions. That is, it is not a matter of fact whether Joe at a height of 5’8” counts as tall or not. If Joe is a fifth-grader and we are comparing him with other boys in his class, then the sentence \textit{Joe is tall} is intuitively true. However, if Joe is an adult male and we are comparing him with other members of his professional basketball team, then the same sentence is intuitively false. Not all context-sensitive predicates are vague, but all vague predicates seem to be context-sensitive. Gradable adjectives such as \textit{tall}, \textit{short}, \textit{deep}, and \textit{expensive} show properties of vagueness.

While vague gradable predicates have historically received the most attention in the analysis of gradable adjectives and comparison, there is a subset of gradable adjectives that do not have vague interpretations. An example is \textit{wet}: an object counts as wet if it has a non-zero amount of moisture. In this case, there is a fact of the matter for when an object counts as wet. In this sense, such predicates are not subject to the same kind of contextual variability as predicates like \textit{tall}.

Thus, the first major distinction that we observe between different types of adjectives is the type of \textbf{standard} they are associated with, namely whether the standard is \textbf{relative} or \textbf{absolute}. Vague gradable adjectives like \textit{tall} have a relative standard that is context-dependent. Non-vague gradable predicates like \textit{wet} have endpoint-oriented standards, which do not vary across contexts in the same way. Specifically, we can say that \textit{wet} has a minimum (lower endpoint) standard. Conversely \textit{closed} has a maximum (upper endpoint) standard; a door counts as closed only if it is completely closed. The term endpoint-oriented already makes reference to the types of scales these predicates encode. In the theoretical literature, Kennedy & McNally (2005) and Kennedy (2007b) argue that a gradable predicate whose scale contains a maximum or minimum element will have its

\textsuperscript{2}The ‘bare form’ of scalar predicates is much more frequently referred to as the ‘positive’ form. We use ‘bare’ here to avoid confusion with positive-polar adjectives, such as \textit{tall} and \textit{heavy}, in contrast to negative-polar adjectives, such as \textit{short} and \textit{lightweight}.
standard fixed to one of those values, and will have an absolute interpretation. Meanwhile, a scale that does not contain endpoints, i.e., an ‘open’ scale, cannot possibly have an endpoint-oriented standard, and therefore must have a relative interpretation.³

The scale structure associated with a predicate is correlated with certain entailment patterns observed with the comparative form. As shown in (1), the comparative forms of relative-standard predicates like tall and pretty do not entail that the bare form holds of the subject. By contrast, the comparative form of predicates with lower-bounded scales like wet entails that the bare form of that predicate holds of the subject. In order for the bare form to hold, the predicate must hold of the subject to a degree exceeding some standard.

(1) a. Mary is taller than Ellen. ⇒ Mary is tall.
   b. Mary is prettier than Ellen. ⇒ Mary is pretty.
   c. The table is more wet than the counter. ⇒ The table is wet.

Kennedy & McNally (2005) argue that the acceptability of certain combinations of degree modifiers with scalar predicates can diagnose whether the corresponding scale contains a maximum or minimum element. That is, the distribution of certain modifiers is tied to the scale structure of the predicate. Intuitively, the maximality modifier completely targets a maximal element on a scale. It is therefore only compatible with predicates that contain a maximal element. Meanwhile, the intensifier very is only compatible with relative-standard predicates, and infelicitous with maximum-standard predicates. Thus we observe the contrasts in acceptability in (2):

(2) a. The bottle is very tall/pretty/#closed.
   b. The bottle is completely closed/#tall/#pretty.

The cross-linguistic generality of these scale-structural distinctions and the corresponding contrasts in (2) have yet to receive much cross-linguistic scrutiny. While the distribution of modifiers

³See Kennedy (2007b) and Bochnak (2013a) for more details on the distinction between relative and absolute gradable adjectives.
can be an informative locus of study, the fieldworker should not take the availability (or, unavailability) of a particular modifier-predicate pair as necessarily indicative of a particular scale structure. In the next subsection, we discuss modifiers in Washo and Navajo that are often translated into English as very, but which have different distributions from their putative English counterpart. In both languages, the modifiers appear to track semantic distinctions other than scale structure.

A second major division in gradable adjectives is Bierwisch’s (1989) classification of adjectives into dimensional and evaluative classes. Dimensional properties are quantitative and are usually associated with physically real systems of measurement (e.g., tall, short, heavy). Evaluative properties are qualitative and include properties describing personality (e.g., hardworking, lazy), color, and texture. Such properties are not typically associated with systems of measurement. Bierwisch’s dimensional properties generally pattern as open scale adjectives with relative standards. Descriptive properties are more heterogeneous: while some pattern as open scale adjectives (e.g., pretty) others pattern more like closed scale adjectives. While modifier selection in English does not seem to track this distinction, there are other areas of the grammar of gradability where this distinction is linguistically relevant, for example in licensing norm-related interpretations in how questions and other degree constructions (see Bierwisch 1989; Rett 2007). It remains to be explored how this distinction may be exploited in other languages.

Another class of adjectives that has recently received some attention in the theoretical literature is the class of so-called ‘extreme’ adjectives like gigantic and gorgeous (Paradis, 2001; Morzycki, 2012). While we will not discuss this class of adjectives in this paper, we highlight it here for two reasons. First, this represents another class of adjectives that shows special restrictions on modifier co-occurrence in English. Specifically, as shown in (3), this class licenses a distinct class of intensifiers, but rejects modification by very, even when they arguably share a scale with a non-extreme counterpart.

(3)  a. Your shoes are downright/positively gigantic/gorgeous/#big/#pretty.

We refer the interested reader to Bogal-Allbritten (to appear) for more discussion of how evaluative properties fit into Kennedy & McNally’s (2005) typology of gradable adjectives.
Second, extreme adjectives represent a class that has not yet received any attention in the cross-linguistic literature, as far as we know. This therefore represents an area within degree semantics that deserves further investigation from semantic fieldworkers.

Finally, we point out the distinction between positive- and negative-polar adjectives. Most gradable adjectives come in antonymic pairs such as tall/short, wet/dry, open/closed. Under many theories, it is claimed that both adjectives share the same scale, but have the opposite perspective. To put it another way, tall makes reference to the positive degrees on the scale of height, while short makes reference to negative degrees on the same scale. This distinction is linguistically relevant in English for licensing a norm-related interpretation in certain degree constructions, as shown in (4). While the positive-negative distinction does not appear to have a reflex in modifier selection in English, we will see in the next section a modifier in Navajo that does track this distinction.

(4) a. Charlie is as tall as David. ⇒ Charlie and David are tall.
    b. Mary is as short as Ellen. ⇒ Mary and Ellen are short.

Before moving on to our examples from Washo and Navajo, we would also like to point out the typological work of Dixon (1982) in distinguishing several lexical classes of adjectives. His work focuses on establishing a hierarchy of the types of scalar attributes that tend to be lexicalized as a distinct syntactic category (i.e., adjectives) in the world’s languages.

The distinctions between adjective classes discussed in this section have so far remained relatively unexplored in semantic fieldwork on understudied languages. This is definitely an area of research on the semantics of gradable predicates and degree constructions that deserves more cross-linguistic scrutiny. We submit that these distinctions are important for the fieldworker to be aware of, even if not all of them turn out to be relevant for modifier selection or norm-related entailments in the language being studied.

5We discuss norm-relatedness in more detail in sections 3-4; also see Rett 2007 for overview and discussion.
2.2 Scalar modifiers cross-linguistically

As shown above, scale structure distinctions are linguistically relevant for the distribution of degree modifiers. However, as seen in both case studies discussed below, English translations of modifiers may make incorrect predictions about the distribution of a modifier in another language. We therefore want to test modifiers with predicates with different scale structures, and indeed across the other distinctions described above, to arrive at the correct modifier and predicate meanings in the language under investigation.

Our first example from the field comes from Washo. Before showing the data on modification, we first introduce the basic morphosyntax of gradable predicates in this language. Washo lacks a distinct morphosyntactic category of adjectives; scalar concepts are typically lexicalized as verbs, as in (5).6 These verbal predicates take the regular morphology found on verbs in this language, for example the ‘imperfective’ suffix -i (Jacobsen 1964). Alternatively, a verbal gradable predicate may appear with the nominalizing prefix de-, in which case a copula hosts verbal morphology, as shown in (6).7

(5) verbal predicate:

\[
\text{mé:hu ?ílkayki?i-í}
\]

boy tall-IPFV

‘The boy is tall.’

(6) nominalization + copula:

\[
\text{mé:hu de-?ílkayki?i k’ë?-í}
\]

boy NMLZ-tall 3.COP-IPFV

---

6Baker (2003) argues that all languages make the distinction between the syntactic categories V, N, and A. In making the above claim that scalar concepts are encoded as verbs in Washo, we are simply referring to the fact that they are verb-y in that that occur with verbal morphology. Future research may reveal a category of adjectives that has a distinct syntactic distribution from verbs in this language, in line with Baker’s cross-linguistic claim. For now, we simply note that the question of whether or not a language has degree-denoting expressions is likely separate from whether it has expressions of category A. Thanks to Robert Henderson for discussion on this point.

7We use the following short forms for glosses: 3 = 3rd person; AOR = aorist; CA = comparative aspect; COP = copula; IPFV = imperfective; NCA = non-comparative aspect; NEG = negation; NMLZ = nominalizer.
‘The boy is tall.’

The modifier šemu is often translated into English by speakers as very when it modifies relative-standard predicates. However, as we see below, šemu has a much wider distribution than very in English. In fact, šemu is compatible with predicates that have either relative or absolute standards. Correspondingly, the translation ‘really’ is more appropriate when šemu modifies absolute standard predicates.

(7) a. ?ilkáykayi? šémuyi
   tall    ŠEMU-IPFV
   ‘He is very tall’
   (relative standard predicate) WASHO

b. ?ilšúšíbi? šémuyi
   straight ŠEMU-IPFV
   ‘It’s really straight.’
   (maximum standard predicate) WASHO

c. ?ilk’únk’núi? šémuyi
   bent    ŠEMU-IPFV
   ‘It’s really bent.’
   (minimum standard predicate) WASHO

Based on such data, it appears then that šemu is not a modifier that tracks scale structure or the relative/absolute distinction. The fact that šemu can apply to predicates of all scale types may indicate that the semantics of šemu is different from English very (despite translations offered by speakers), or that scalar predicates in Washo don’t lexicalize scale structure in the same way that scalar predicates in English do, or both. In any case, a translation task that only focused on predicates that license very in English would have missed the distributional differences between very and šemu, despite initial translational equivalence.

Our second example comes from Navajo. Once again, we consider a modifier (yee’) that is translated by speakers and Young and Morgan (1987) as very. Like in Washo, the modifier’s dis-

See Bochnak (2012) for an analysis.
tribution does not track the distinction between adjectives with relative and absolute standards. However, in contrast with Washo, the Navajo modifier has a narrower distribution than English *very*. The modifier *yee*’ is felicitous with open scale adjectival predicates denoting negative dimensions (8a,b) but not with predicates denoting positive dimensions (8c).9

(8) a. shideezhí 'áłts'óózí yee’
   my.little.sister 3-slender YEE’
   ‘My little sister is very slender.’

b. díí dibé yázhi 'áltśíí yee’
   this sheep small 3-little,small YEE’
   ‘This lamb is very small.’

c. # shideezhí nineez yee’
   my.little.sister 3-tall YEE’
   (Intended: ‘My little sister is very tall.’)

The distribution of *yee*’ suggests that it is sensitive to the distinction between positive- and negative-polar adjectives. Given an antonymic pair of dimensional adjectives, *yee*’ is only felicitous with the negative member. Although this distinction does not appear to be relevant for licensing modifiers in English, we noted above that it is relevant for licensing norm-related interpretations in some degree constructions (section 2.1) (Rett, 2007).

In this section, we considered modifiers in Washo and Navajo that are both translated into English as *very* but which do not have the same distribution as *very*. The availability of English *very* has been argued to indicate that an adjective has a relative standard and a scale lacking endpoints. However, it would be a mistake for a fieldworker to conclude based on the grammaticality of Washo >?ilšíšíí? šemuyi (‘straight ōemuyi) that >?ilšíšíí? ‘straight’ has an open scale and a relative standard in Washo. Other diagnostics point away from this conclusion. For instance, absolute-standard predicates are pragmatically marked in certain comparison constructions in Washo (Bochnak, 2013a,b). Similarly, it would be a mistake for a fieldworker to conclude from

---

9To the second author’s knowledge, the only positive adjective that appears to be felicitous by *yee*’ is *nizhóní* ‘pretty.’ This also seems to be the only evaluative predicate with which *yee*’ is felicitous.
the infelicity of Navajo #nineez yee’ (‘tall yee’) that nineez ‘tall’ does not have an open scale and relative standard in Navajo. We will see in section 4.2 that nineez has a relative standard.

It is crucial to first establish the distribution of the modifier in the language of study before taking the modifier as diagnostic of a particular feature of scale structure. This is particularly important for work on underrepresented languages, where not only might English translations be misleading but where existing linguistic materials may not contain complete descriptions of the distribution of degree modifiers.

3 Norm-relatedness and crisp judgments

In the remainder of the paper, we turn our attention to comparative constructions. We address the study of comparative constructions themselves, as well as their usefulness in diagnosing the scale structures of gradable predicates. In this section, we examine the distribution of norm-relatedness and crisp judgments in Washo comparative constructions, and consider how they are related to each other. We show that using the right kind of visual stimuli in the field can be exploited to bring out the relevant semantic distinctions, and illustrate the success of such techniques with data from Washo.

A construction is NORM-RELATED if it entails that the bare form of the adjective holds (Bierwisch, 1989). For instance, the comparative form of the dimensional, relative-standard predicate tall in (9a) is not norm-related, since it does not entail that the bare form in (9b) holds.

(9) a. Alice is taller than Brian.
    b. Alice is tall.

In English, the comparative, superlative, and excessive constructions of dimensional predicates lack the property of norm-relatedness (Bierwisch 1989). One environment in English where

10 Other terms for norm-relatedness found in the literature includes ‘orientedness’ (Seuren, 1978) or ‘evaluativity’ (Neeleman et al., 2004; Rett, 2007).
11 As already mentioned in section 2, certain other constructions may vary depending on the polar variant of the scalar predicate. See Rett (2007) for an analysis of such facts in English.
the bare form is not norm-related occurs in the *compared to* construction, as in (10). In fact, as argued by Sawada (2009), there is actually an implicature that the bare form does not hold of either individual being compared.

(10) Alice is tall compared to Brian.

What counts as tall varies both with the sorts of objects under consideration (e.g., adult humans versus skyscrapers) and also with the particular context of utterance. Correspondingly, during the investigation of degree constructions in the field, it is crucial that the fieldworker and consultant share a conversational background that allows the fieldworker to evaluate whether a certain construction is norm-related. We use examples from the first author’s investigations of Washo to illustrate. In this language, the primary comparative construction is a CONJOINED COMPARISON (Bochnak, 2013a,b), whereby two clauses containing antonymous predicates are juxtaposed, and a comparison between two objects is inferred (Stassen, 1985). An example is given in (11):

(11) *télìlìwù delkáyki? kë?ì* šáwłamhu delkáyki?ë:s  k’a?-aš
    man   NMLZ-tall    3.COP-IPFV  girl   NMLZ-tall-NEG 3.COP-AOR
    ‘The man is taller than the girl.’
    (literally: ‘The man is tall, the girl is not tall.’)  WASHO

Since the comparison in (11) makes use of the bare form of scalar predicates, we may wonder whether the scalar predicates receive a norm-related interpretation, or whether this is similar to the English comparative construction in being non-norm-related.

In some cases, presenting a consultant with a context verbally may suffice if we are confident that both the fieldworker and the consultant share a common ground to make the relevant judgments with respect to norm-relatedness. For instance, the following context was presented verbally in the contact language (English), whereby the fieldworker and consultant share the common ground that heights of five feet or four and a half feet do not count as tall for adult humans:

(12) a. Context: comparing a man who is five feet tall and a woman who is four and a half feet tall (i.e., both are clearly short for adult humans)
We observe that the conjoined comparison using the predicate delkáykayi? ‘tall’ is unacceptable when neither individual being compared counts as tall. Note that the English comparative in (12c) is perfectly felicitous in the same context. We thus have some evidence that conjoined comparisons in Washo are norm-related, in contrast with English comparative constructions, which are not.

However, many contexts that one would want to use to test norm-relatedness are not as straightforward to explain verbally to a field consultant. The context in (12a) works well because the contact language (English) provides a means by which we can precisely talk about heights (e.g., measure phrases like five feet). But there are many scalar predicates in English for which there is no standardized measurement system. Take big for example. Most native speakers probably have good intuitions about when an object counts as big, though there is no conventional means for precisely measuring bigness. Additionally, big is a scalar predicate that is associated with many dimensions. Indeed, an adult human can be considered big if s/he is quite tall, or quite wide, or quite heavy. Furthermore, our intuitions about what precise measurements count as big seem to break down here: is 50 cm wide considered big for an adult human? We therefore run into difficulty when trying to verbally establish a context like (12a) to test norm-relatedness in constructions involving the scalar predicate big.\(^\text{12}\)

We argue that the use of targeted visual stimuli can help overcome this difficulty. The task we have in mind involves the fieldworker collecting a set of objects that differ along some dimension (e.g., height, width). This set should have two properties. First, it should consist of objects of the same type. For instance, comparing a pencil and a piece of rope with respect to length is undesirable, since what counts as long for a pencil may not count as long for a piece of rope. Visually presenting objects of the same type obviates this difficulty. Second, this set should contain

\(^{12}\)Another potential problem is that you may be working in a contact language that has very few or even no measure phrases, making it very difficult to talk about precise measurements in the first place.
objects that differ along the relevant dimension within a wide enough range such that it is clear that a certain subset holds the relevant property, while some other subset does not. In the investigation of degree constructions in Washo, the first author has used sets of pinecones for this task, since pinecones are plentiful in the part of the world where the fieldwork is conducted. They additionally come in a variety of shapes and sizes, which allowed us to compare them along several different dimensions.

To begin, the fieldworker and consultant come to an agreement over which members of the set count as big, which ones count as small, and which ones are in-between. This can be done in either the contact or the research language. However, we recommend that the consultant confirm set membership using the targeted predicates in the research language, in case the extensions of gradable predicates do not quite overlap in the two languages.\(^\text{13}\) After coming to an agreement with the consultant about which pinecones count as big, which ones count as small, and which ones are medium-sized, the consultant is presented with two pinecones from the ‘small’ set, but which still differ in size, and is asked to compare them. In this context, the consultant utters (13b):

\begin{enumerate}
\item Context: comparing two small pinecones; both are clearly not big (for pinecones)
\textit{this small-AOR one this big slightly-IPFV}
\textit{‘This one is small, that one is a little bit big.’} \textit{WASHO}
\item That one is bigger than this one. \textit{ENGLISH}
\end{enumerate}

We observe that the consultant resists using the unmodified predicate \textit{t’i:yeli?} ‘big’ to describe a pinecone that does not fall under its extension, even though it is bigger than another pinecone. Indeed, a follow-up elicitation shows that the comparison in (13b) is unacceptable without the modifier \textit{wéwši} ‘slightly’. Note once again that the English comparative in (13c) is perfectly acceptable in the same context. We thus more evidence for our hypothesis from above that conjoined comparisons in Washo are norm-related.

\(^{13}\)For example, in some languages, one might not use the same predicate for ‘tall’ to describe humans and non-human objects.
To summarize, using this elicitation technique can help elucidate the properties of degree constructions with respect to norm-relatedness through the creation of contexts where the fieldworker and consultant share a common ground of what objects count as big, for example, in a particular context. The use of real-life visual stimuli facilitates the investigation since it controls for the fact that such predicates are highly context-sensitive, and ensures that the fieldworker and consultant have the same context in mind. We now go on to show that a similar task can also be used to elicit judgments on whether comparative constructions support crisp judgments.

Crisp judgment contexts are those in which the two objects being compared differ minimally along the relevant dimension (Kennedy, 2007a,b). Different comparison constructions differ with respect to whether they support crisp judgments or not. For instance, the English comparative in (14b) using the comparative morpheme -er is felicitous in the crisp judgment context given in (14a), while the compared to construction in (14c) is infelicitous.

(14)  

| a. Context: Charlie is 5’6” tall and John is 5’5” tall. |
| b. Charlie is taller than John. |
| c. # Charlie is tall compared to John. |

In fact, it is a general property of the bare form of vague predicates that they do not support crisp judgments. That is, if two objects $x$ and $y$ differ only slightly with respect to a vague predicate $P$, then if $x$ is $P$ is true, then we are unable or unwilling to accept $y$ is $P$ as false (the ‘Similarity Constraint’: Fara, 2000). Investigating whether a comparative construction supports crisp judgments thus requires setting up contexts such as in (14a) where objects differ minimally with respect to some property.

The visual stimulus task described above can be modified to create the relevant contexts for testing crisp judgments. The fieldworker simply needs to juxtapose two objects that differ minimally with respect to the relevant dimension. In the case of Washo, the following sentence was offered by a speaker comparing two pinecones that differ in a perceptible but minimal amount with respect to size:
Here, we have a first clue that conjoined comparisons in Washo do not support crisp judgments. The speaker avoids using the canonical \( x \text{ is } P, y \text{ is not } P \) conjoined comparison in this context, but rather opts to use two distinct predicates to compare the pinecones (\( t'\text{í:yeli}' \text{ 'big'} \) and \( d\text{eltétebi}' \text{ 'fat'} \)). If this construction in Washo does not support crisp judgments, then it would be infelicitous for a speaker to use the canonical \( x \text{ is } P, y \text{ is not } P \) construction, due to the Similarity Constraint on the use of vague predicates.

A follow-up elicitation session confirms this hypothesis. In this case, images of ladders were presented to the speaker on a computer screen. While the use of real-life visual stimuli allow the fieldworker to create more natural contexts (rather than having a speaker imagine some hypothetical scenario), it may not always be possible to find a set of objects that meets the ideal specifications described above. This is where technology can be useful: electronic images can be manipulated and scaled by the fieldworker to create the appropriate set of visual stimuli for testing the relevant distinctions. For instance, the image in Figure 1 shows a set of ladders that differ along the dimension of height.\(^{14,15}\)

In the follow-up elicitation, the fieldworker presented the following conjoined comparison to a Washo speaker in the context of comparing the second and third ladders from the left in Figure 1 (i.e., a crisp judgment context). The sentence is rejected by the speaker in this context:

\[ (16) \quad \begin{align*}
\text{a. Context: comparing two ladders, one only slightly taller than the other} \\
\text{b. } & w'\text{í:di}? \ t'\text{é:we}? \ d\text{ewgí?iš} \ k'\text{é}?-i \quad w'\text{í:di}? \ t'\text{é:weja} \ d\text{ewgí?išēs} \ k'\text{é}?-a\text{s} \\
\text{this} & \text{ much} \text{ height} \quad 3.\text{COP-IPFV this} \text{ much-NC height-NEG} \quad 3.\text{COP-AOR} \\
\end{align*} \]

\(^{14}\text{Of course, we still need to make sure the fieldworker and consultant agree about which objects count as tall and short in this scenario, since even the tallest ladders in Figure 1 could in principle be miniscule.}\)

\(^{15}\text{Thanks to Yaron McNabb for sharing images used in experimental studies on the use of gradable predicates in English, upon which Figure 1 is based.}\)
As discussed in Bochnak (2013a,b), the findings for norm-relatedness and crisp judgment effects suggest that Washo gradable predicates should be analyzed as vague predicates rather than scalar expressions. In other words, Washo gradable predicates represent an alternative type of semantic object to adjectives as usually conceived (cf. Cresswell 1976, von Stechow 1984, Kennedy 1999, Kennedy & McNally 2005). We refer the interested reader to Bochnak (2013a,b) for more discussion.

Crucially, the investigation of norm-relatedness and crisp judgment effects was facilitated by the use of visual stimuli. These visuals can be used both for initial explorations where the speaker is invited to spontaneously offer a sentence to describe a certain situation, and also for targeted follow-up elicitations where the fieldworker tests a sentence against the pictured context. The advantage of this methodology is that the fieldworker and consultant can explicitly negotiate together over whether objects hold a certain property, obviating some of the difficulties related to the context-sensitivity of the predicates being investigated.

We close this section with a couple of caveats about using this type of visual stimuli in elicitation. First, since we have shown that the same type of stimulus can be used to test both norm-relatedness and crisp judgments, the fieldworker must make sure to test for only one semantic feature at a time. For instance, in testing norm-relatedness, it is crucial that the context presented is not also a crisp judgment context, since the fieldworker needs to be able to identify the source of...
infelicity if a speaker judges a sentence unacceptable. Second, while we have shown the applicability of these visual stimuli for testing dimensional predicates, it can be difficult to find or create appropriate stimuli to test norm-relatedness with non-dimensional (i.e., evaluative) predicates. We discuss this issue further in the next section.

4 Norm-relatedness and lexical competition

To this point, we have concentrated on elicitation of judgments for sentences containing dimensional predicates, namely tall and big. Even when the fieldworker focuses on this relatively small set of gradable predicates, we have seen that the format of the elicitation materials is key. In this section, we consider the methodological considerations that arise when the fieldworker expands the study to include evaluative adjectival properties. The data in this section come from the second author’s investigation of Navajo. Elicitation techniques for dimensional and evaluative predicates are discussed in the context of a fieldwork study that probes for norm-relatedness within the adjectival domain. We argue that competition between different morphological forms of a single property leads to norm-relatedness in comparative constructions. We continue to advocate for the value of visual and tactile stimuli, but also make a case for the use of verbal stimuli in certain cases. We argue that verbal stimuli are especially useful for evaluative adjectival predicates, such as pretty, since verbal stimuli allow the context to be more carefully defined and agreed upon by the fieldworker and the consultant.

4.1 Background on Navajo adjectival predicates

The core semantic meaning of Navajo adjectival predicates — like event-denoting predicates — is carried by the stem, which invariably occurs at the right edge of the predicate. Stems are necessarily marked with prefixes encoding nominal arguments and may bear additional prefixes. For event-denoting predicates, the additional prefixes include markers of situational and viewpoint aspect, iterativity, and distributivity. Adjectival stems also bear additional prefixes. We will refer collectively to the prefixes borne by adjectival stems as COMPARATIVE ASPECT (CA) and NON-
COMPARATIVE ASPECT (NCA). Choice of CA vs. NCA morphology also may affect pronunciation of the stem (e.g., tones). For detailed discussion of CA and NCA morphology, please see Young & Morgan (1987) and Bogal-Allbritten (2010, 2013).16

Not all adjectival stems can bear CA morphology: CA morphology is largely restricted to Bierwisch’s (1989) dimensional properties (Young & Morgan, 1987). The existence of a CA-marked form of a given stem entails the existence of an NCA-marked form of the same stem. CA- and NCA-marked forms of the stem -nee ‘tall’ are shown in (17).

(17)  a. -nee
    TALL

      b. nineez
         3-tall

      c. ’áníínéez
         3-tall

Many more adjectival stems can only be NCA-marked. Adjectival predicates with only NCA-marked forms are more heterogenous in meaning. This set of predicates includes predicates whose English translations have closed scales (e.g., digiz ‘crooked’) and predicates whose English translations have open scales, according to Kennedy & McNally’s (2005) typology. Among the latter are predicates expressing Bierwisch’s (1989) evaluative properties, e.g., nízhóní ‘pretty’.17

(18)  a. -zhóní
      PRETTY

---

16The use of the term ‘aspect’ with specifically CA or NCA morphology should not be taken as a theoretical claim that these morphemes are, in some sense, true aspectual morphemes. We use the term ‘aspect’ here to maintain continuity with previous literature on Athabaskan languages, including Young & Morgan (1987). The term is used by the authors to capture the linear proximity of CA/NCA prefixes on adjectival predicates to viewpoint and situational markers on eventive predicate.

17A small number of CA-marked forms of evaluative properties are reported by Young & Morgan (1987). However, the second author found that these forms were rarely volunteered by consultants in comparative constructions. Consultants found some of the CA-marked forms to be unfamiliar, or disagreed on their morphological form. This suggests that these forms are becoming obsolete or irregular in a way that dimensional CA-marked predicates are not.
b. *nizhóní*
   3-pretty\textsubscript{NCA}

c. No CA-marked form

To summarize, dimensional properties are expressed in Navajo by predicates that reliably have both CA- and NCA-marked forms. By contrast, evaluative properties are only expressed by predicates with NCA-marked forms. In the next section, we show that with respect to norm-relatedness in comparative constructions, predicates with both CA- and NCA-marked forms pattern differently from predicates with only NCA-marked forms. In order to draw this conclusion, it was necessary to test for norm-relatedness in comparative constructions containing dimensional and evaluative predicates. Different types of elicitation materials were found to be ideal for each class of property. While visual and tactile contexts were useful for dimensional predicates, verbal contexts were used with evaluative predicates.

### 4.2 Testing for norm-relatedness

We first examine the semantics of bare adjectival predicates in Navajo (i.e., predicates not further modified by comparative morphology). Recall that in English, bare English adjectives are always norm-related, except when used in the *compared to* construction (sect. 3). We confirm below that when NCA-marked Navajo adjectival predicates appear without further modification, they are felicitous in the same types of norm-related contexts that would license English bare adjectives with relative standards.

(19) a. Context: You are showing me some rugs. One of the rugs stands out for the high quality of the weaving, the color of the yarn, and the intricateness of the pattern. You are telling me about this rug. You say to me:

   b. *díí díyógi nizhóní*
      this rug  3-pretty\textsubscript{NCA}

   Comment: “You can say that, you’re saying this rug looks good.”

(20) a. Context: You are describing your family to me. Your little sister is tall for a young girl. She is only 10 years old but is already 5’ tall.
b. \textit{shidéezhí nineez}\textsuperscript{NCA}
\begin{center}
my.younger.sister 3-tall\textsuperscript{NCA}
\end{center}
Comment: “You can say this.”

If the context was not norm-related, consultants judged sentences with unmodified adjectival predicates to be infelicitous.

(21) a. Context: You are showing me some rugs. One of the rugs has uneven weaving and unattractive colors. You are telling me about this rug. You say to me:

b. \# \textit{díí diyóqí nizhóní}\textsuperscript{NCA}
\begin{center}
this rug 3-pretty\textsuperscript{NCA}
\end{center}
Comment: “If the rug isn’t pretty, why are you saying it is?”

Only NCA-marked adjectival predicates are shown above because CA-marked predicates can never appear without further modification (22). Reasons for the ungrammaticality of (22) are discussed at length in Bogal-Allbritten (2013).

(22) * \textit{shidéezhí} \textit{'ánínléez}\textsuperscript{CA}
\begin{center}
my.younger.sister 3-tall\textsuperscript{CA}
\end{center}
(Extended: ‘My younger sister is tall.’)

We now turn our attention to norm-relatedness of adjectives in comparative constructions. Both CA- and NCA-marked adjectival predicates can appear in the comparative construction.\textsuperscript{18} Kennedy & McNally (2005) use norm-relatedness in comparative constructions as a diagnostic of minimum endpoints on adjectival scales (sect. 2.1). In Navajo, lexical competition appears to be another source of norm-relatedness in comparatives:

(23) a. If an NCA-marked adjective has a CA-marked counterpart, then the comparative construction with the NCA-marked adjective is necessarily norm-related.

\textsuperscript{18}The ‘greater than’ ordering relation is expressed by the standard marker -lááh ‘beyond.’ The predicate does not bear any special morphology (cf. English MORE/-ER). This is a cross-linguistically common method of expressing comparative relations (Stassen, 1985). For more discussion of the syntax of Navajo comparative constructions, see Bogal-Allbritten (2013).
b. If an NCA-marked adjective does not have a CA-marked counterpart, then the comparative construction with the NCA-marked adjective is not necessarily norm-related.

We first examine comparative constructions containing adjectival predicates for which only an NCA-marked form is available. Consultant comments in (25) indicate that \textit{nizhóní} ‘pretty’ is not obligatorily norm-related in the comparative construction. Norm-relatedness was tested for by setting up contexts in which both objects of interest were established to meet, or not meet, the relative standard of comparison for the relevant property. We discuss the contexts used to elicit the judgments in (24) and (25) in section 4.3.

(24) a. Context: Verbal context identifying two rugs as attractive (30).

b. \textit{dii} \textit{diyógí} ‘eii \textit{diyógí} yiláah \textit{'át'éego nizhóní} this rug that rug 3-beyond 3-being 3-pretty\textit{NCA}

\textit{‘This rug is prettier than that rug.’}

c. Comment: “You can say this. This rug is pretty but this one is better looking.”


b. \textit{dii} \textit{diyógí} ‘eii \textit{diyógí} yiláah \textit{'át'éego nizhóní} this rug that rug 3-beyond 3-being 3-pretty\textit{NCA}

\textit{‘This rug is prettier than that rug.’}

c. Comment: “It is relative to what you’re comparing. If you have two ugly rugs, then you’re just saying that this one is better than the other.”

Navajo \textit{nizhóní} behaves like its English translation \textit{pretty}: neither is obligatorily norm-related in comparative constructions. This suggests that both are open scale adjectives according to Kennedy & McNally’s (2005) diagnostics.

We now turn to Navajo adjectival predicates for which both CA- and NCA-marked forms exist. (26) shows that CA-marked adjectives can be used in comparative constructions in non-norm related contexts.

(26) a. Context: You are describing your family to me. Your mother and your younger sister are both short women. Your mother is 5’2” and your younger sister is 4’11”.
b. *shimá* *shidéézhí* *yilááh* *’áníítnééz*
   my.mother my.younger.sister 3-beyond 3-tall_{CA}
   ‘My mother is taller than my younger sister.’

c. Comment: “This sounds good. They aren’t tall, but my mom is taller.”

Both the Navajo predicate *’áníítnééz* and its English translation *tall* behave like open scale adjectives: neither is obligatorily norm-related in comparative constructions.

English and Navajo diverge when we look at Navajo *NCA*-marked predicates that have a *CA*-marked form. The comparative containing *NCA*-marked *nineez* ‘tall’ is judged infelicitous in the context in (27), where the individuals under comparison do not meet the relative norm or standard *tall*. The consultant suggested that the *CA*-marked form of the predicate be used instead.

(27)  
   a. Context: Same as (26)
   b. *# shimá* *shidéézhí* *yilááh* *’át'éego nineez*
      my.mother my.younger.sister 3-beyond 3-being 3-tall_{NCA}
      (Intended: ‘My mother is taller than my younger sister.’)
   c. Comment: “You don’t want to use *nineez*. This implies that they’re tall people. You’d have to use *’áníítnééz*.”

By contrast, in a context like (28) where both individuals under comparison meet the standard for *tall*, both *NCA*- and *CA*-marked adjectives were felicitous:

(28)  
   a. Context: You are describing your friend’s family. Her mother and younger sister are both tall women. Her mother is 6’4” and her sister is 6’2”.
      *bimá* *bidéézhí* *yilááh* *’át’éego nineez*
      her.mother her.younger.sister 3-beyond 3-being 3-tall_{NCA}
   b. Comment: “This sounds good. You can say *’áníítnééz*, too.”

In both English and Navajo, adjectives that occur in the bare form are obligatorily norm-related. In both languages, norm-relatedness is sometimes retained when adjectives appear in comparative constructions. Kennedy & McNally (2005) cite the presence of minimum endpoints on adjectival scales as a source of norm-related comparative constructions. In Navajo, competition between
NCA- and CA-marked predicates is another source of norm-relatedness in comparative constructions. If a CA-marked form of a particular predicate exists, then a comparative construction containing the NCA-marked form of that predicate is obligatorily norm-related.

### 4.3 Constructing the contexts

We saw above that the distribution of CA and NCA morphology is linked to Bierwisch’s (1989) distinction between dimensional and evaluative properties. Dimensional properties are expressed by stems with both CA- and NCA-marked forms. Evaluative properties are expressed by stems for which only NCA morphology is generally available. In the study described above, the second author probed for norm-relatedness in comparative constructions with dimensional and evaluative predicates. In this section, we discuss the elicitation materials used by the second author to obtain the reported judgments. Visual and tactile materials were found to be best-suited for dimensional properties while we suggest that verbal contexts can be valuable for studying evaluative properties.

In initial elicitation sessions examining adjectival predicates with dimensional meanings (e.g., -neez ‘tall’), verbal contexts like (29) were used on their own. Before a judgment about a Navajo target sentence was given, the fieldworker and consultant agreed that the two individuals were short. This was established by giving each individual’s height and referring to them as ‘short women.’

(29) Context: You are describing your family to me. Your mother is 5’2” tall and your younger sister is 4’11”. Your mother and your younger sister are both short women but there is a difference in their heights.

As discussed for Washo in section 3, adjectives like tall are compatible with verbal contexts because the contact language (English) permits us to precisely talk about heights using measure phrases. However, targeted visual materials were used in later elicitation sessions. Although felicity judgments were consistent regardless of whether visual materials were used or not, visual materials helped consultants to keep track of individuals under comparison. After many contexts like (29) have been discussed in succession, it is easy to lose track of which individual is meant to
be the taller one and which the shorter one. Figure 2 was presented to consultants in conjunction with contexts like (29). Names and identifying information were added below the relevant individuals in accordance with the context. Each individual was indicated by the fieldworker when mentioned in the context.

![Figure 2: ‘Height chart’ visual stimulus for norm-relatedness judgments](image)

To this point, discussion of elicitation materials for both Washo and Navajo has focused on dimensional properties. A different set of challenges is presented by evaluative properties like pretty. First, pretty and its Navajo translation nizhóní lack a single dimension of measurement: in contrast with tall, individuals’ levels of beauty cannot be described precisely with measure phrases. One potential solution to this problem would be to use visual and tactile materials like those used to study the predicate translated as big in Washo. Recall that big not only lacks a precise measurement system but is also potentially associated with many dimensions. A similar challenge is presented by pretty: while pretty is not clearly associated with multiple dimensions, prettiness is a subjective property (Lasersohn, 2005, Stephenson, 2007). Different speakers may take different factors into account when determining whether an object counts as pretty or nizhóní. As a result, speakers may disagree about whether the same object is ‘pretty’ or not.

Second, because of the subjectivity of pretty and nizhóní, it proved difficult to locate ideal objects for use in visual or tactile contexts. While pinecones that differed minimally in size could be found, fieldworkers may find it difficult to locate objects that both the consultant and the fieldworker agree to be ‘pretty.’ It is even more difficult to locate pairs of objects which only differ slightly in their beauty.
Given these considerations, we argue that verbal contexts are particularly suitable for the study of evaluative properties. The following verbal contexts were used to elicit the judgments reported in (24) and (25), respectively.\footnote{In context (30), the English adjective pretty was named in the context. In other contexts, the English adjective does not appear. The presence or absence of the English adjective in the context did not seem to affect consultants’ judgments.}

(30) Context: We’re looking at rugs at a trading post. There is a wide range of rugs, from very attractive to very unattractive. You hold up two rugs that are pretty. Their wool is attractively dyed and the weaving was done very straight. One of the rugs is even prettier than the other. I ask you what you think of them.

(31) Context: We’re looking at rugs at a trading post. There is a wide range of rugs, from very attractive to very unattractive. You hold up two rugs that are not good-looking: their wool is dyed in strange colors, the weaving was not done in straight lines, and the design is not complex. One is of slightly better quality than the other, however. I ask what you think of the two rugs.

By presenting the context verbally, the fieldworker is able to invoke a degree of subtlety that might not be possible if real world objects were used. Furthermore, the verbal context in (31) explicitly establishes what qualities the rugs have (crooked weaving, strange coloration, etc.). These qualities were independently confirmed by consultants to be attributes of unattractive rugs. By making background assumptions explicit in this way, the consultant and fieldworker can come to a clear agreement on whether the objects under comparison meet the standard of comparison for the property in question.

A final note can be made about the study of bare adjectives. In order to elicit judgments about sentences like (32), contexts were presented verbally rather than visually. This was a deliberate decision, even though the predicate in (32) is dimensional.

(32) a. Context: You are describing your family to me. Your mother is tall: she is 5’11”. You are telling me about her.
b. \textit{shimá ninee\textsubscript{NCA}z}
\hspace{1cm} my.mother 3-tall\textsubscript{NCA}

Comment: “You can say this if your mother is tall.”

Visual materials depicting more than one object — the target (‘my mother’) and a standard of comparison (e.g., a tree or building) — were initially constructed to test sentences like (32). However, consultants found it most natural to describe the these pictures with comparative constructions rather than unmodified adjectives. If the standard of comparison were eliminated from the visual materials, it was still necessary for the context to be clarified verbally. Verbal presentation allow the fieldworker and consultant to come to agreement about the context without the additional complication introduced by visual materials.

The distinction between dimensional and evaluative properties is relevant to the distribution of adjectival morphology in Navajo. While dimensional adjectival stems can bear either \textit{CA} or \textit{NCA} morphology, evaluative properties are only expressed by \textit{NCA}-marked adjectival predicates. The distribution of \textit{CA} and \textit{NCA} morphology was in turn shown to be relevant to determining whether a comparative construction is obligatorily norm-related or not. In many — perhaps even most — languages, the distinction between evaluative and dimensional properties may not affect the distribution of adjectival morphology. However, it is still important for the fieldworker to be aware of the distinction between dimensional and evaluative meanings. Different types of elicitation materials (verbal, visual, and tactile) are best suited to the two classes of adjectives.

5 Conclusion

The conclusions we have drawn can be categorized into the theoretical and the methodological, although there is, of course, interaction between the two categories. Considering first the theoretical conclusions, the fieldworker should be aware of predictions made by theoretical accounts of scale structure and predicate meaning. The application of Kennedy & McNally’s (2005) theory permitted both authors to make initial predictions about the types of scale structures that might appear in their languages of study. Two of Kennedy & McNally’s scale structural diagnostics —
the distribution of modifiers and norm-relatedness in comparatives — were applied to gradable predicates in Washo and Navajo. Both the modifier and norm-relatedness findings suggest interesting differences between Washo, Navajo, and English. For instances, both Washo and Navajo were shown to have modifiers translated as English *very* but with different distributions from *very*. We argued that the modifiers in both languages were sensitive to aspects of adjectival meaning that were not relevant to the distribution of English *very*. Further cross-linguistic studies of the semantics of gradable predicate and degree constructions will likely find constructions sensitive to other aspects of adjectival meaning, and may also uncover additional complexities and patterns that should be incorporated into the typological picture of adjectives.

On the methodological side, we have drawn two conclusions. First, fieldwork on a relatively restricted subject — such as the semantics of gradable predicates — still benefits if the fieldworker has access to a broader base of knowledge about the language of study. In both Navajo and Washo, norm-relatedness of adjectives arose in comparative constructions containing adjectives for which norm-relatedness is not predicted by Kennedy & McNally (2005). For both case studies, these departures suggested a reanalysis of gradable predicates. Their reanalysis in each language was supported by additional knowledge about the language. In Washo, the findings for norm-relatedness in adjectives were considered alongside the unavailability of crisp judgments in comparative constructions. Both sets of facts follow if gradable predicates are analyzed as vague predicates (Bochnak 2013,a,b). In Navajo, the patterns of norm-relatedness became clear once the distribution of adjectival morphology (*CA* vs. *NCA*) was taken into account.

Second, even if a given way of classifying adjectives does not seem relevant to the language of study, different classes of adjectives may still be best suited to different types of elicitation materials. We found that for dimensional predicates, both visual and tactile contexts were effective. For evaluative adjectives, however, verbal contexts permitted more careful definition of the context and more explicit agreement between fieldworker and consultant about the context’s crucial aspects.

While we have not exhausted all the possible avenues of research on gradable predicates and degree constructions that a fieldworker may undertake, we hope to have provided basic overview
of some of the issues that may arise and the methodologies we have found useful in investigating this area of meaning in the field. As we have already mentioned, though it bears repeating, this is an area of semantics that is ripe for making new discoveries based on cross-linguistic research, which can then inform future typological and formal research.

References

AnderBois, Scott & Robert Henderson. This volume. Linguistically establishing discourse context: two case studies from Mayan languages.


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


