On the scalar properties and telicity of degree achievements

Louise McNally

Universitat Pompeu Fabra

Kennedy & Levin (2008) argue that the aspectual properties of so-called degree achievement (DA) verbs (e.g. *darken*) can largely be predicted from the scale structure of the adjectives to which they are derivationally related (e.g. *dark*). Specifically, when the adjective is evaluated on a scale that is upper closed and the standard for the adjective to truthfully apply is the upper endpoint on that scale (i.e., when the adjective is absolute; see e.g. Kennedy & McNally 2005), the corresponding DA can be either telic or atelic. In contrast, when the adjective's scale is open and the standard is context-dependent (i.e., when the adjective is relative), the corresponding DA is atelic. In this paper, I defend, following Kearns (2007), the position that telic interpretations of DAs are not directly a function of the standards for the adjectives from which the verbs are derived. Rather, the telic interpretation simply depends on it being possible to characterize the amount of change undergone in terms of the part structure of the event described, without reference to a specific comparison class. This conclusion will emerge from reflection on how the notions of relative and absolute standards can be recast in terms of similarity- vs. rule-based classification (as proposed in McNally 2011), extended from the adjectival to the verbal domain.

Keywords: lexical semantics; degree achievement; telicity; scale structure

1. Introduction

Dowty (1979) introduced the term *degree achievement* (hereafter, DA) to describe change of state verbs like *to darken*, *to cool*, or *to widen*, which manifest variable telicity: some, like *to darken*, can describe telic changes (e.g. darkening completely) or atelic changes (darkening by some unspecified degree), as illustrated in (1a), where the *in* adverbial signals a telic reading, and the *for* adverbial, an atelic one. Others, like *to widen* in (1b) seem to describe only atelic processes.

(1)   a.   The room darkened {for a few minutes/in a few minutes}
      b.   The road widened {for a few minutes/??in a few minutes}

Kennedy & Levin (2008), building on Hay, Kennedy & Levin (1999), argue that the aspectual properties of DAs can largely be predicted from the scale structure of the adjectives to which they are derivationally related (e.g. *dark*, *cool*, *wide*).[1] Specifically, when the adjective is evaluated on a scale that is upper closed and the standard for the adjective to truthfully apply is the upper endpoint on that scale (i.e., when the adjective is *absolute*; see Kennedy & McNally 2005 and below for further details), the corresponding DA can be either telic or atelic, as for *dark*/*to darken* in (1a). In contrast, when the adjective's scale is open and the standard is context-dependent (i.e., when the adjective is *relative*), as for *wide*/*to widen* in (1b), the corresponding DA is atelic.[2]

However, this generalization has been challenged by Kearns (2007), who argues that DAs derived from relative adjectives can also have telic interpretations, contrary to what (1b) suggests; Fleischhauer (2013) makes a similar claim. The goal of this paper is to supplement Kearns' and Fleischhauer's argumentation. While I will maintain, along with Kennedy and Levin and others (e.g. Krifka 1998, Beavers 2012), that in general it is illuminating to relate scalar semantics and telicity, I offer a new argument, based on a set of deadjectival verbs such as *awaken*, which are morphologically identical to DAs but are not generally grouped with them due to their differing telicity properties, that the scale structure of the adjectival property from which a DA is derived is not directly accessible to the derived verb. Rather, what matters for the telicity of a deadjectival verb is simply that the standard for the underlying adjective to hold be clearly identifiable. The relative ease or difficulty of identifying this standard value is arguably what Kennedy and Levin's generalization is getting at. Absolute adjectives have standards that are, as a rule, easier to identify than those for relative adjectives. The former therefore lend themselves better to telic interpretations than the latter. This discussion of DAs will naturally lead to reflection on a bigger question that has received comparatively

---

1.   Dowty's original definition of this class was not strictly based on morphology, but rather on whether the verb entailed a change in a vague gradable property. Thus, his classification of DAs included verbs like *to sink* and *to improve*. Though Kennedy and Levin also do not specifically discuss the internal morphology of DAs, their semantics (2008: example (25)) clearly takes them to be derived compositionally from gradable adjective denotations. In this paper, I will take as the domain of interest the class of verbs derived morphologically from adjectives via the suffix –*en* (e.g. *to darken*) plus those that are surface-identical to those adjectives (e.g. *to cool*) and generally assumed to be derived from them (see e.g. Marchand 1964). I will not consider those verbs with no obvious morphological relation to adjectives, such as *to improve*, as well those derived from adjectives via other suffixes, such as –*ify* (e.g. *to purify*), which are not generally included in the class of DAs and whose semantic properties might differ from those of verbs derived in –*en* due to differences in the semantics of the affixes.

2.   Kennedy and Levin do not discuss DAs related to adjectives with upper closed scales whose standards are the lowest positive value on the scale, such as *to awaken*; see below for more on these verbs.

little attention in the literature, namely: If verbal predicates are associated with scalar meanings, what are the properties of these scales and how are they established?

The structure of the paper is as follows. In Section 2, I begin by briefly reviewing Kennedy and Levin's analysis of DAs and the challenges it faces, adding some new examples to those discussed by Kearns and Fleischhauer. I then discuss the expectations that their analysis raises for deadjectival change of state verbs like *awaken*. In Section 3, I provide the background for a different perspective on the variable telicity of DAs by briefly presenting the typology of adjective gradability in McNally (2011) and sketching how it could extend to gradability in the verbal domain. An alternative account of verbal telicity follows in Section 4.

2.   Adjective scale structure and the telicity of DAs

2.1   The measure of change analysis of DAs

Kennedy and Levin's (2008) analysis of DAs builds on the intuition that these verbs can be associated with a scale whose properties reflect the degree of change the theme undergoes in the property described by the underlying adjective over the course of the event described by the verb. On analogy with the treatment of adjectives as measure functions (Bartsch & Vennemann 1972, 1973; Bierwisch 1989; Kennedy 1997), DAs are taken to denote *measure of change* functions. Specifically, the DA takes as its input an individual (the theme participant) and an event (the event of change) and returns a degree corresponding to the difference between the value of the theme on the measure function denoted by the underlying adjective at the beginning of the event, and the value of the theme on the same measure function at the end of the event. An example of the representation for the DA *to darken* appears in (2).

(2)   **darken**$_\Delta$: $\lambda x \lambda e.$**darken**$_\Delta(x)(e)$

As in the case of adjectives, abstract degree morphology (represented as **pos**v in (3)) is needed to turn the verb into a truthfully applicable predicate. This morphology introduces the standard value **stnd(mΔ)** for the measure of change function **mΔ**.

(3)   **pos**v(**darken**$_\Delta$): $\lambda x \lambda e.$**darken**$_\Delta(x)(e) \geq$ **stnd**(**darken**$_\Delta$)

In general, the authors observe, "a DA based on a measure of change function **mΔ** is true of an object x and an event e [iff] the degree to which x changes as a result of participating in e exceeds the standard of comparison for **mΔ**" (i.e., **stnd(mΔ)**, Kennedy & Levin 2008: 174). They posit that **stnd(mΔ)** can always be the lowest positive value on the measure of change scale, corresponding to the smallest change possible; this minimum degree is a natural candidate for a standard value because once even the smallest change has taken place, it is true that there is some event of which the DA can

be truthfully asserted to hold. The choice of this lower endpoint as the standard corresponds to the so-called *comparative* reading of the DA (e.g., of *to darken* as *to become darker*). The authors associate this reading with atelicity insofar as there is no specific non-minimal degree of required change that would mark the end of the event (though see the Section 2.3 for more on this point).

Kennedy and Levin argue that **stnd(mΔ)** can also be the upper endpoint on the measure of change scale if there is an upper limit to the amount of change possible in the property contributed by the adjective underlying **mΔ**; this upper limit corresponds to the maximum degree of change possible. If this standard is chosen, the result is the so-called *positive* reading (e.g. of *to empty* as *to become empty*). If the underlying adjective's scale is open, that is, if there is no maximal degree to which something can have the adjectival property, **mΔ** 's scale is (upper) open, as there is going to be no limit to the change in the adjectival property that the theme can undergo. Nonetheless, in any given context, the adjective can be associated with a standard of comparison, and so the question arises as to why **stnd(mΔ(x)(e))** cannot be equal to that degree of change corresponding reaching the (contextually established) standard value for the adjective. The answer Kennedy and Levin provide relies on the principle of Interpretive Economy (Kennedy 2007):

(4)   Interpretive Economy: Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.

According to this principle, if a scalar expression is associated with a closed scale, one or both of the endpoints on the scale should be chosen as a possible standard, unless the adjective meaning is conventionalized in some other way. The endpoints form part of the conventional meaning of the scalar expression, and thus should be favored as standard values over any degree in the middle of the scale, whose choice as a standard would have to be determined by context.

If Interpretive Economy holds, DAs based on open scale adjectives should have lower endpoint standards. Recall that the scale for such DAs has a minimum value (corresponding to the smallest degree of change), but no maximum value (since no maximal degree of change is possible). Crucially, since the standard for the underlying adjective in such cases depends on context, it should not be possible to choose that degree (or change up to that degree) as an alternative standard, and DAs derived from open scale adjectives should have only atelic, comparative interpretations.

However, Kennedy and Levin are careful to point out that what is crucial is conventionalization. If reaching a non-minimal/non-maximal degree of the adjectival property becomes conventionalized in the meaning of the verb, as they argue happens with *to cool* (see the discussion of (8), below), this standard can be used without

violating Interpretive Economy. In such cases, a telic interpretation should also be possible alongside the atelic reading.

Summarizing, on this analysis, when **stnd(mΔ)** is the upper endpoint on the scale or a conventionalized value between the minimum and the maximum value, the verb is telically interpreted. In contrast, when **stnd(mΔ)** is the minimum value on the scale, the verb is atelically interpreted. The latter standard for the DA should always be available; the former is available only for some lexical items.

2.2   Non-maximum standards and telicity

As mentioned in the introduction, the claim that the telicity properties of DAs follow from the telicity properties of the underlying adjectives has been challenged. Kearns (2007) argues on the basis of examples like (5) that DAs like *to darken* can be given a telic interpretation without entailing that a maximal degree of darkness is reached.

(5)   The sky darkened in an hour, but it wasn't completely dark

Fleischhauer (2013) makes the same argument using German data such as (6), which shows that for something to stabilize entails that that thing reaches the state of being stable, but not necessarily the state of being completely stable.

(6)   a.   #Der Zustand   hat   sich   stabilisiert,   er   ist   aber   nicht   stabil
               the   condition   has   itself   stabilized   he   is   but   not   stable
               'The condition has stabilized, but it is not stable'

      b.   Der Zustand   hat   sich   stabilisiert,   er   ist   aber   nicht
             the   condition   has   itself   stabilized   he   is   but   not

             vollkomme   stabil
             completely   stable

             'The condition has stabilized, but it is not completely stable'

Kennedy and Levin argue in the case of Kearns's example that the definite description is interpreted imprecisely, providing in support of this claim the data in (7), which are intended to show that when the definite description is clearly interpreted maximally, the sentence is no longer felicitous.

(7)   a.   #All of the sky darkened in an hour, but it wasn't completely dark
      b.   #The sky darkened in an hour, but no part of it was completely dark

However, if complete darkness is measured according to the depth of the darkness, as opposed to the extension of darkness across the sky, neither of these examples seems infelicitous.

In (8) we see another sort of data that has been used to challenge the generalization that only DAs related to adjectives with upper endpoint standards can be telic.

(8)   The soup cooled in 10 minutes

Clearly, the truth of (8) does not require the soup to cool to the lower endpoint on the temperature scale; rather, it must only reach some contextually established value, such as room temperature. In this case, Kennedy and Levin argue, the temperature in question has become conventionalized in the meaning of the verb and thus can serve as the endpoint for the event.

However, when we consider further examples, it becomes increasingly difficult to maintain that a specific degree is conventionalized in all cases where a telic reading is available. For example, alongside the examples above with *to darken*, and *stabilisieren*, we find examples such as those in (9):[3]

(9)   a.   The gelatin will finish hardening around the fruit and candy, creating a solid cell model   (<www.ehow.com/how_6306508_make-human-cell-science-project.html>)
         b.   Put the fruit out the next day to finish drying. Fruit will be leathery when dry   (<www.aces.edu/pubs/docs/H/HE-0360/HE-0360.pdf>)

It is clear that for these sentences to be true, the gelatin need not be 100% hard, nor need the dried fruit be completely devoid of moisture. However, it is fairly evident in the context what degree of hardness/dryness counts as the standard. Thus, despite the fact that they involve DAs derived from adjectives with closed scales, these example raise the same sort of problem that the soup cooling example in (8) raises.

Other examples like (8) appear in (10). These DAs are derived from adjectives whose scales lack an upper endpoint, and yet the DAs are interpreted telically, the event in question being finished when some relevant non-minimal degree of the adjectival property is reached.

(10)   a.   She finished sharpening her pencil before she answered him
                                        (from *A Wee Christmas Homicide* by Kaitlyn Dunnett)
          b.   Taste the wine and decide if it needs sweetening… Add sugar or better still grape juice concentrate. Keep doing this until the wine is to your taste… Taste the wine now you have finished sweetening it and decide if it is sharp enough for you
                        (<www.wineworks.co.uk/content/making-wine-from-grapes>)

---

3.   Here I use acceptability as a complement to the verb to *finish* as a diagnostic of telicity. Some of the following examples involve causative/transitive variants of the DAs, but as Kennedy and Levin note, whether the DA is used transitively or intransitively is not does not affect the issue under discussion.v

c. The extra water in the chicken corn mixture will finish moistening the potatoes

(from *Freezer Bag Cooking: Trail Food Made Simple*
by Sarah Svien Kirkconnell)

If one wants to maintain that the data introduced in this section are not counterexamples to Kennedy and Levin's analysis, it becomes necessary to have a theory of when and how it is possible to conventionalize as the endpoint of an event of reaching a non-maximal degree of the relevant adjectival property. However, the more examples one finds of this sort, the more difficult it is to maintain the claim that the relevant standard (of e.g. hardness, dryness) is conventionalized. At most it looks like what could be conventionalized is the fact that the standard must be appropriate to the situation being described. Such data thus suggest that the telicity of a DA is not directly predictable from the scale structure of the adjective from which it is derived. The verbs to which we now turn raise similar sorts of questions.

2.3   Verbs derived from adjectives with lower-endpoint standards

Verbs derived from gradable adjectives whose standard is (or can be) the minimum value on a scale, such as *to awaken*, are not discussed by Kennedy and Levin, probably because they do not behave like DAs.[4] However, since they are derived using the same morphological process as are DAs, we might expect a single account to predict their telicity properties as well. Specifically, we would expect them to have both a positive reading and a comparative reading: the positive reading corresponding to reaching the minimum degree of the adjectival property – this reading would be telic, as the standard for the adjective is that minimum degree – and the comparative reading corresponding to reaching a higher degree of the property in question, beyond the standard.

However, this prediction is not borne out. These verbs lack a comparative reading: (11a) cannot be paraphrased as (11b); more specifically, (11a) entails that the baby had no degree of awakedness immediately before the event.

(11) a. The baby awoke
      b. The baby became more awake

Relatedly, these verbs are always telic, as their behavior on various telicity diagnostics shows. For example, the progressive form of *to awaken* does not entail the perfect (see

---

4.   The scalar properties of the adjective *awake* are discussed in Kennedy & McNally (2005). Other verbs that behave like *to awaken* are *to sicken* and *to dampen* in the sense of to become damp. *To open* and *to wet*, like *to cool* and *to empty*, do not have deadjectival morphology but, as verbs related to adjectives with lower-endpoint standards, also show similar behavior.

(12a)), and these verbs are acceptable as the complement of *to finish* if it is possible to imagine a prior process culminating in the event (see (12b)).[5]

(12)   a.   The baby is awakening. -/- > The baby has awoken
       b.   The baby finished awakening

Nonetheless, the adjective *awake* is gradable (and indeed, is associated with a fully closed scale):

(13)   a.   The baby is more awake than his older sister
       b.   The baby is fully awake

These data suggest that the scale structure of the adjective is not directly accessible to the derived verb: If it were, we would expect the comparative reading to be possible.

3.   Scale structure and standards for verbal predicates

In order to move towards an alternative analysis, it will be useful to leave the issue of telicity aside for a moment and approach from a different angle the question of the role of the underlying adjective's scale structure in determining the semantics of DAs and other, morphologically-related deadjectival verbs. Specifically, I will compare the way in which we decide whether a verbal predication holds to the way in which we decide whether an adjectival predication holds. In other words, I will explore the question of what kinds of standards gradably-used verbal predicates can be associated with. To do this, I first briefly review the analysis of the distinction between relative and absolute standards for adjectival predicates developed in McNally (2011); I then carry this analysis over to the verbal domain. Doing this will elucidate the contribution of the underlying adjective. I will then return to the question of telicity.

3.1   A different view of relative and absolute standards for adjectives

As mentioned above, gradable adjectives have been associated with two kinds of standards: relative and absolute. Kennedy (2007) provides a detailed discussion of the

---

5.   Intransitive *to dampen* and *to sicken* do not sound very good in the progressive at all, as is consistent with their describing an event of reaching the minimum degree of the adjectival property. Though *to awaken* and transitive *to sicken* are acceptable in the progressive, the progressive describes a process prior to the awakening or sickening, and not an ongoing situation that could be described by the verb *to awaken*. See e.g. Piñón (1997) and Kearns (2003) for discussion of this interpretation of the progressive as a diagnostic for punctuality. Transitive *to dampen* is different: its scalar properties can be determined in terms of the moisture applied to an object or the extent of the object dampened, both of which license a process use of the progressive and a comparative reading for the verb.

differences between the two; I briefly mention two of these differences here and refer the reader to Kennedy's paper for further discussion. First, relative standards are context dependent, while absolute standards are not. Thus, an adjective with a relative standard like *old* can be judged to truthfully apply to an individual in some contexts and not apply in others, whereas the applicability of an adjective with an absolute standard like *closed* does not vary in this way, as illustrated by the contrast in (14):

(14) a.   Compared to Mary, John is old; compared to Sue, he is not old
      b.   ??Compared to the small box, the big box is closed

Second, adjectives with relative standards allow for borderline cases, i.e. it may be difficult to judge whether the property in question clearly holds or does not hold of a particular individual. For example, we might feel confident that a 70-year-old person counts as old, and that a 10- year-old person does not (when considering average human lifespan), but it is more difficult to decide how to classify a 45-year-old. Such uncertainty does not arise in the case of adjectives with absolute standards.

Kennedy & McNally (2005) and Kennedy (2007) identify absolute standards with minimum (non-zero) and maximum values on a scale, and relative standards with non-minimum, non-maximum values. However, McNally (2011) and Foppolo & Panzeri (2013) present counterexamples to these generalizations. These include *full* as used in (14a), and color terms on the reading where gradability is measured according to the extension of the color on an object (as opposed to some other dimension such as intensity or prototypicality of the color; see (15b)).

(15) a.   The wine glass is full
      b.   The scarf is blue

A full wine glass is typically only about half full (rather than full to the top), and for a sentence like (15b) to be judged true, it is sufficient that the color predominate (as opposed to being manifest on all parts of the object in question). Nonetheless, the standards for these uses of the adjectives are absolute according to the diagnostics mentioned above. Given these facts, McNally (2011) proposes that what differentiates relative and absolute standards is not the nature of the degree that marks the standard, but rather the type of classification involved in deciding the applicability of the predicate. Specifically, relative standards are claimed to reflect *similarity-based classification*, whereas absolute standards reflect *rule-based classification*, a distinction discussed in detail in Hahn & Chater (1998). Rule-based classification requires strict matching between the classification criterion/a – equivalent to what semanticists typically call the satisfaction conditions for a predicate – and the relevant properties of the object being classified, and involves comparing a representation associated with a specific individual against a more abstract representation. In contrast, similarity-based classification requires only a partial match and involves comparing a representation of

a specific individual or property of that individual against another representation of an equally specific individual or one or more of its properties. Similarity-based classification is inexorably dependent on the set of individuals being classified (in other words, a *comparison class*) and presupposes that these individuals will be partitioned into two or more groups: those labeled by an adjective and other candidate adjectives that are sensitive to the same dimension (e.g. *old* and *young* for age).[6] The intuition is that when we want to decide if some individual is, for example, old, one way we can do this is by considering a particular comparison class and sorting the individual according to whether he more closely resembles in age a prototype we will have of old individuals from that group, or rather a prototype we have of young individuals from that group.[7] Interestingly, the standard for truthful application of an adjective like *old* is not established *a priori* in such cases but rather is a consequence of the way the members of the comparison class are partitioned. It is, therefore, inevitably context dependent. It also allows for borderline cases when an individual is equally similar to the prototype for each of the candidate classes into which it might be sorted. We have seen that these are all characteristics of adjectives with relative standards.

Rule-based classification works very differently. As it involves classification by comparison to an abstract representation, there is no role for a specific comparison class of individuals in deciding whether an individual falls into a given category (labeled by an adjective) or not; there is thus no context dependence related to such a class.[8] As it requires exact matching, the effect of crisp judgments is reproduced and the absence of borderline cases is expected. These are all characteristics of adjectives with absolute standards.

With this characterization of the relative/absolute distinction in hand, we can now turn to the verbal domain.

---

6.   In these respects, similarity-based classification is deeply reminiscent of the semantics for gradable adjectives defended in Klein (1980), which relied crucially on the partition of a comparison class into positive and negative extensions for the adjective.

7.   There may be a number of ways of choosing these prototypes. For example, if we think of deciding which members of a comparison class are old as the task of clustering individuals into two groups – the old and the young –, the prototype for the old individuals could be identified with the centroid for the cluster of old individuals. However, how exactly this prototype is chosen is not crucial for the present discussion. See e.g. Gärdenfors (2000), Barner & Snedeker (2008), and Solt & Gotzner (2012) for relevant discussion.

8.   This is not to say that the standard might not vary according to the general type of object being described. As McNally (2011) discusses, the standard for being full for a wine glass as a type of object is different from that for other sorts of vessels. However, what is crucial is that these standards can be determined for a kind of object in the abstract and do not depend on a choice of particular instances of that kind of object. See also the experimental evidence in Liao, et al. (to appear).

3.2 Relative and absolute standards for verbal predicates

As a first observation, note that the use of any verbal predicate presupposes on the part of the speaker the individuation of an eventuality that he or she wants to describe – otherwise it is difficult to see how we would ever manage to use verbal predicates in the first place. This eventuality may or may not be part of a larger eventuality in progress. The logically prior nature of eventuality individuation will be crucial in the discussion that follows.

It is possible for the standard for verbal property ascription to be derived from classification by similarity, even if the predicate is not what we would normally consider to be gradable. For example, imagine that three music students are supposed to be practicing their instruments: Alex runs her fingers over the keyboard for 5 minutes; Berta diligently plays scales for 5 minutes, but then stops; and Carol spends 45 minutes carefully doing everything she was asked to by her teacher. In such a circumstance, one might utter (16).

(16)   Compared to Alex, Berta practiced; compared to Carol, she didn't

Note that the construction in (16) (in contrast to a normal comparative such as *Berta practiced more than Alex*) entails that whatever Alex did should be called something else – in other words, it is false that she practiced, and what she did should be classified differently (for example, as running one's fingers over the keyboard). In other words, with verbs the issue has to do with appropriateness.

If we order these eventualities to reflect the degree to which each one qualifies as practicing, the result is a function of the different aspects of what we might consider a prototypical practicing event, and not just of one of those aspects. This is because similarity-based classification works on the basis of partial matching, and in principle two objects or eventualities can be considered equally similar to a third even if the properties that match between them, and thus the support the similarity judgment, are not identical. For example, in the case of (16), the amount of time spent at the instrument, the purpose with which the instrument is played, and whether what is played is repeated in a particular way are just a few of the different properties that could be matched.

The sort of comparison illustrated in (16) is similar to the between-noun comparison discussed in Sassoon (2013). It does not correspond to the gradability properties of verbs as captured in the measure function analysis of DAs. What is relevant for the analysis of the relation between verbal telicity and scalarity, I submit, is rather classification by rule. Specifically, we match the eventuality being described against an abstract representation, made available by the satisfaction conditions of the verbal predicate, by checking the relevant parts of the eventuality against the individual satisfaction conditions. For example, we might decide whether it is true that Alex, Berta or Carol practiced by considering a set of necessary and sufficient conditions for the

purposes at hand: if these are met, it will be true that the individual in question practiced; if not, it will not.[9]

Logically prior to checking whether an object or eventuality matches an abstract description is a decision about how to interpret any context-dependent conditions. Recall in the case of adjectival predicates that when we decide whether a wine glass is full using rule-based classification, we must first decide how much of the volume of the glass must be occupied in order to count as full. Being able to do this presupposes having an abstract representation of wine glasses as they are filled in particular sorts of contexts. Once the relevant representation has been identified, an exact match must obtain the occupied volume of the glass under consideration and the occupied volume in the abstract representation in order for the glass to count as full. Crucially, the volume in question need not be either minimal or maximal.[10] Similarly, when we grade color predications according to the extension of the color on an object, we must previously fix (and hold constant) a value for the color along the (relative) gradable dimension related to hue. In the case of DAs, their use will presuppose a decision about what degree of the adjectival property has to be reached (whether minimal, maximal, or somewhere in between) in order for it to be possible to evaluate whether a given eventuality matches the abstract representation embodied in the satisfaction conditions for the DA. On what basis this decision can be made, and how it relates to telicity, are the questions to which we now turn.

4.   Verb scalarity and telicity

For the purposes of discussion, I will assume that telic predicates are those for which *only a non-arbitrary (and typically not proper) subpart* of the eventuality being described satisfies the predicate in question, and that atelic predicates are those for which *any arbitrary subpart* of the eventuality being described can satisfy the predicate

---

9.   It is of course a matter of significant debate to what extent the semantics of lexical items can be characterized in terms of necessary and sufficient conditions (e.g. Lakoff 1987), particularly as the complexity of the content associated with the items increases. However, as our focus is not on the nature of the conditions associated with verbal predicates but rather how these are related to telicity, and because the satisfaction conditions for DAs are relatively simple (see below), we can largely abstract away from this problem.

10.   However, it must be clearly identifiable. For this reason, it seems that non-endpoint absolute standards tend to be related to proportions, which can often be calculated relatively easily without the use of measuring devices. See McNally (2011) for discussion.

in question.[11] To apply this characterization to DAs, let us suppose that their semantics is as proposed by Kennedy and Levin, and thus that the representation for e.g. *to darken* will be as in (17) (repeated from (3) above):

(17)   **pos***v*(**darken**$_\Delta$): $\lambda x \lambda e.$**darken**$_\Delta(x)(e) \geq$ **stnd**(**darken**$_\Delta$)

The standard for the measure of change function (i.e. **darken**Δ) must be established in advance of any attempt to evaluate the truth of a predication of the DA. Let us assume that there are two options for this standard: a specific, possibly non-minimal degree of change; or else the smallest degree of change possible. If the first standard is chosen, it will be the case that only a non-arbitrary part of a given event will satisfy the description in (17), namely, the part of the event that is large enough to include the point at which the change reaches the standard. In this case, the predicate will have a telic interpretation. In contrast, if we accept that a possible standard for **m**Δ can be the smallest degree of positive change, any arbitrary subpart of the event described will satisfy the verbal predication. This is, as noted above, the hallmark of an atelic reading.

Kennedy and Levin's hypothesis is that, for adjectives whose standard is context-dependent, the specific, non-minimal degree of change that supports the telic reading should be strongly dispreferred over the minimal degree of change that supports the atelic reading. They grounded this hypothesis in the notion of Interpretive Economy, introduced in (4), above. Moreover, they identify the atelic reading with the comparative interpretation (*to become* Adj-*er*), and the telic reading with the positive interpretation (*to become* Adj).

However, things are actually a bit more complicated than this. First, there are telic uses of DAs that intuitively have a comparative interpretation to the extent that they do not entail that, at the end of the event, the adjectival property holds of the theme. These cases arise when the amount of change that is undergone is specific and non-minimal, but not necessarily identical to the standard for the underlying adjectival property. The examples in (18) are typical:

---

11.   See e.g. Krifka (1989); Parsons (1990); Pustejovsky (1995); Hay et al. (1999); Borik (2006); Piñón (2008); Rothstein (2008); Landman & Rothstein (2010) and Marín & McNally (2011) for various alternative definitions and comparisons between them. The qualifications to the characterization in the text are in order. First, in the case of (atelic) dynamic predicates such as *sing*, the claim that the predicate will hold for any arbitrary subpart of a singing event holds only to a certain level of granularity (e.g. it won't hold for a subpart that consists only of a breath between notes). Second, it is not obvious how this characterization applies to predicates denoting truly instantaneous eventualities – that is, achievement predicates as characterized by Vendler (1957); Mittwoch (1991) and Piñón (1997) – insofar as for these it is impossible to talk meaningfully about arbitrary subparts. Interestingly, some of these seem to be telic, while others seem to be atelic; see Marín & McNally (2011) for discussion.

(18) a. The construction company took one year to widen the road
      b. They haven't finished deepening the port yet

These sentences do not under any circumstances entail that the road is wide or that the port will be deep, but rather only that the amount of change reaches an implicit goal. In this latter respect they contrast with examples such as (8), repeated in (19a), which does license an entailment to (19b) with *cool* being understood as "the desired temperature", even if that temperature might not count as cool in all contexts.[12]

(19) a. The soup cooled in 10 minutes
      b. The soup is cool

A second complication is that, as discussed in Section 2.3, a minimal degree of positive change is sufficient with some verbs, such as *to awaken*, to entail that the adjectival property associated with a DA holds. These verbs do not license comparative readings, nor are they interpreted atelically, as would be expected if the scale structure of the adjective formed the basis for a measure of change function denoted by the verb.

We can account for the contrast between (18) and (19), and disconnect telicity from the positive reading, by appealing to the hypothesis that all non-stative verbal predication can be modeled as rule-based classification.[13] Such classification requires being able to match an event against an abstract representation, which, as discussed above, presupposes that there is a set of clearly identifiable properties (or conditions) that characterize the abstract representation. In the case of DAs, these properties include the amount of change the theme participant must undergo. The harder it is to establish a standard for the adjective underlying the DA independently of a specific comparison class (i.e., the harder it is to identify an absolute standard as characterized in McNally 2011), the harder it will be to construct an abstract representation for the positive reading of the DA which can support rule-based classification. Since it is more difficult to establish such standards with open scale adjectives (see Kennedy 2007 and McNally 2011 for discussion), it is unsurprising that DAs derived from such adjectives are difficult to associate with a positive reading. Kennedy and Levin note that Kearns (2007) appeals to a similar sort of explanation for the difficulty of obtaining telic readings for verbs like *widen*.

There are different ways of establishing an absolute standard that is neither a minimum nor a maximum value on a scale. For example, it seems relatively easy to identify

---

12. An anonymous reviewer rightly observes that in this respect these examples are similar to telic uses of motion verbs where context supplies a fixed distance that is not overtly expressed in the sentence.

13. As the gradability properties of stative predicates appear to be different from those for non-stative predicates, I set the former aside here. However, this decision does not affect the present discussion.

standards based on the proportion of area or volume of an object that manifest the property in question. However, this of course requires that the scale for the property be one that is defined relative to area or volume. This is not the case for the properties denoted by adjectives like *wide*, *deep*, or *steep*, and thus this way of defining the relevant standard for the positive reading of *to widen*, *to deepen*, and *to steepen* is unavailable. Another way in which a standard can be identified is with respect to some specific metric. Such metrics can be numerical (as when we measure temperature or distance), but they can also be defined in other terms. For example, the standard for *cool* for a physical object might be established with respect to an average person's body temperature, or the ambient temperature; the standard for *dark* for the sky might be established with respect to the predominance of a particular shade of grey that speakers can easily identify, or for a room in functional terms, such as the difficulty with which one can see. The availability of such metrics is arguably what is behind the facility with which a positive reading is available for *to cool* and *to darken*. In contrast, it is more difficult to find such a natural metric for width, depth or steepness, and this may explain why the positive reading for the corresponding DAs is so hard to get.

An additional factor that may impede the perception of the positive reading is the possibility that change in width, depth or steepness may rarely involve reaching whatever the standard is for being wide, deep, or steep, and since the positive reading entails the comparative reading, we may simply have no clear evidence for interpreting these verbs positively rather than comparatively, without it being the case that the positive reading is, strictly speaking, unavailable. Preliminary, if still only suggestive, evidence for this possibility comes from the fact that the 450 million word Corpus of Contemporary American English (Davies 2008) contains only 3 instances of the lemma for *become* followed by *wide*, all of which are predicated of eyes, vs. approximately 4900 tokens of the verb lemma *widen*. Similarly, there were 7 instances of *become deep*, predicated of breathing, pain, snow, and a story, vs. nearly 3400 instances of *deepen*; and 3 of *become steep*, predicated of an incline, terrain, and cost, vs. 89 instances of *steepen*. These numbers suggest that the changes that are associated with the positive reading of these verbs are not a common part of our everyday experience.

If what facilitates the telic reading for DAs derived from open-scale adjectives is the easy availability of an abstract standard for the corresponding adjectival property that would support rule-based classification, we might expect that precisely those sorts of arguments that lend themselves to predications of the *become Adj* sort would allow a positive interpretation for the corresponding DAs. Such a possibility exists at least for the case of *to widen*: (20) easily admits a positive reading.

(20)   His eyes widened

The fact that it may be difficult to identify in absolute terms a standard for the adjectives underlying these DAs need not preclude it being easy to identify a specific, non-minimal amount of change that is able to support a telic interpretation of the

predicate on the comparative reading. For sentences like (18), for example, the relevant amount of change is that which the agent of the action intends to produce. Though the speaker might not know what amount of change this is, s/he can infer that the agent does, and that an event of widening or deepening of this sort will only be complete if the intended degree of width or depth has been reached. This is the source of the telic but comparative reading: no arbitrary subpart of an event described using *to widen* or *to deepen* in this sense will satisfy the verbal description, but it isn't required that the specified degree of change be sufficient for the theme to count as wide or deep. In the case of (20), the amount of change can be identified with the amount that indicates the surprise of the individual whose eyes have widened.[14]

Summarizing, there is, after all, evidence to suggest that pragmatic considerations play an important role in explaining the strong preference for a comparative reading over the positive reading with these verbs. More generally, I conclude that, though being able to characterize a standard in absolute terms eliminates one sort of context dependence in interpretation (insofar as it removes appeal to a specific comparison class), there is no reason to think that it precludes all sorts of context dependence, and the fact that the data are sensitive in part to the nature of the theme argument makes it hard to maintain the position that the standard must be conventionalized in the verb meaning.

Now, let us turn to the question of why verbs like *to awaken* lack comparative readings, that is, why (11a), repeated below in (21), cannot be felicitously used to describe an increase in an already awake baby's degree of awakedness.

(21)   The baby awoke

The lack of comparative reading indicates that it is not systematically possible to choose as the standard for the DA the smallest degree of positive change in the property contributed by the underlying adjective, independently of the initial degree to which the property is held. This is unexpected on the Interpretive Economy account, since the lower endpoint of the measure of change scale is conventionalized.

This fact is reminiscent of the observation made by Kearns (2007) and others, that DAs based on adjectives with upper closed scales such as *to darken* have default telic readings, i.e. an utterance of (22) strongly implies, even if it does not entail, that the room became dark:

(22)   The room darkened

---

14.   Indeed, this correlation would account for the possibility of using (20) strictly metaphorically as a way of expressing surprise.

Kennedy and Levin suggest that this default could be accounted for pragmatically insofar as the positive reading is stronger, and thus more informative, than the comparative reading. This is also true for cases like (21): all situations in which the positive reading is true (i.e. the baby awakens), the comparative one is as well (the baby becomes more awake), but there are cases where the comparative reading is true, but the positive reading will not be true, since the individual in question will already be awake. However, the positive reading for (21) is not merely a default: it is the only interpretation for the sentence. The telicity of verbs like *to awaken* is reminiscent of the telicity of path verbs with overt source arguments, such as *to move* (*from* somewhere). Gehrke (2008) claims, contra e.g. Filip (2003) and Nam (2005), that source arguments can make verbs telic, even though the source is just the beginning of a path that could in principle be open-ended. She argues that whether a predicate is telic or not is orthogonal to whether it contains a source or a goal. This is similar to challenging the claim that the telicity of DAs is necessarily associated with an upper endpoint on a scale, and atelicity, with a lower endpoint. Gehrke further argues that there is morphosyntactic idiosyncrasy in whether or not a given source or goal argument identifies an entailed result state, which on her analysis entails telicity. In other words, it is a fact about the part structure of the event described – namely whether the predicate is true of an arbitrary subpart of the event or not – that determines telicity. But this cannot be predicted from the nature of the path, just as it cannot be predicted from the nature of the scale for the adjective underlying a DA. This is not to deny that the generalization that DAs based on closed scale adjectives are more likely to be telic, and those based on open scale adjectives are more likely to be atelic, but, as Gehrke concludes for the behavior of source and goal arguments with respect to telicity, this generalization has a pragmatic, rather than grammatical, explanation.

5.   Conclusion

In this paper I have defended, following Kearns (2007), the position that telic interpretations of degree achievement predicates are not directly a function of the conventionalization of the standards for the adjectives from which the predicates are derived. Rather, the telic interpretation simply depends on it being possible to characterize the amount of change undergone in terms of the part structure of the event described, without reference to a specific comparison class. This conclusion has emerged from reflection on how the notions of relative and absolute standards can be recast in terms of similarity- vs. rule-based classification, extended from the adjectival to the verbal domain.

Acknowledgements

References

Barner, David & Snedeker, Jesse. 2008. Compositionality and statistics in adjective acquisition. *Child Development* 79(3): 594–608. doi: 10.1111/j.1467-8624.2008.01145.x

Bartsch, Renate & Vennemann, Theo. 1972. The grammar of relative adjectives and comparison. *Linguistische Berichte* 20(1): 19–32.

Bartsch, Renate & Vennemann, Theo. 1973. *Semantic Structures: A Study in the Relation Between Syntax and Semantics*. Frankfurt: Athenäum Verlag.

Beavers, John. 2012. Lexical aspect and multiple incremental themes. In *Telicity, Change, and State: A Cross-Categorial View of Event Structure*, Violeta Demonte & Louise McNally (eds), 23–59. Oxford: OUP. doi: 10.1093/acprof:oso/9780199693498.003.0002

Bierwisch, Manfred. 1989. The semantics of gradation. In *Dimensional Adjectives*, Manfred Bierwisch & Ewald Lang (eds), 71–261. Berlin: Springer. doi: 10.1007/978-3-642-74351-1_3

Borik, Olga. 2006. *Aspect and Reference Time*. Oxford: OUP. doi: 10.1093/acprof:oso/9780199291298.001.0001

Davies, Mark. 2008. The Corpus of Contemporary American English: 450 million words, 1990-present. <http://corpus.byu.edu/coca/>

Demonte, Violeta & McNally, Louise (eds). 2012. *Telicity, Change, and State: A Cross-Categorial View of Event Structure*. Oxford: OUP. doi: 10.1093/acprof:oso/9780199693498.001.0001

Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel. doi: 10.1007/978-94-009-9473-7

Filip, Hana. 2003. Prefixes and the delimitation of events. *Journal of Slavic Linguistics* 11(1): 55–101.

Fleischhauer, Jens. 2013. Interaction of telicity and degree graduation of change of state verbs. In *Studies in Composition and Decomposition of Event Predicates*, Boban Arsenijević, Berit Gehrke & Rafael Marín (eds), 125–152. Dordrecht: Springer. doi: 10.1007/978-94-007-5983-1_6

Foppolo, Francesca & Panzeri, Francesca. 2013. Do children know when their room counts as clean? *NELS 40: Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, Vol. 1, Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds), 205–218. Amherst MA: GLSA.

Gärdenfors, Peter. 2000. *Conceptual Spaces: The Geometry of Thought*. Cambridge MA: The MIT Press.

Gehrke, Berit. 2008. *Ps in Motion: On the Semantics and Syntax of P Elements and Motion Events*. PhD dissertation, Utrecht University. LOT Dissertation Series 184.

Hahn, Ulrike & Chater, Nick. 1998. Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition* 65(2): 197–230. doi: 10.1016/S0010-0277(97)00044-9

Hay, Jen, Kennedy, Christopher & Levin, Beth. 1999. Scale structure underlies telicity in degree achievements. In *Proceedings of SALT IX*, Tanya Matthews & Devon Strolovitch (eds), 127–144. Ithaca NY: CLC Publications.

Kearns, Kate. 2003. Durative achievements and individual-level predicates on events. *Linguistics and Philosophy* 26(5): 595–635. doi: 10.1023/A:1025803912153

Kearns, Kate. 2007. Telic senses of deadjectival verbs. *Lingua* 117(1): 26–66.
doi: 10.1016/j.lingua.2005.09.002

Kennedy, Christopher. 1997. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. PhD dissertation, University of California. Published 1999 by Garland Press, New York.

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

Kennedy, Christopher & Levin, Beth. 2008. Measure of change: The adjectival core of degree achievements. In *Adjectives and Adverbs: Syntax, Semantics, and Discourse* [Oxford Studies in Theoretical Linguistics 19], Louise McNally & Christopher Kennedy (eds), 156–182. Oxford: OUP.

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

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

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

Krifka, Manfred. 1998. The origins of telicity. In *Events and Grammar*, Susan Rothstein (ed.), 197–235. Dordrecht: Kluwer. doi: 10.1007/978-94-011-3969-4_9

Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago IL: University of Chicago Press. doi: 10.7208/chicago/9780226471013.001.0001

Landman, Fred & Rothstein, Susan. 2010. Incremental homogeneity and the semantics of aspectual for phrases. In *Syntax, Lexical Semantics, and Event Structure*, Malka Rappaport Hovav, Ivy Sichel & Edit Doron (eds), 229–251. Oxford: OUP.
doi: 10.1093/acprof:oso/9780199544325.003.0011

Liao, Shen-yi, McNally, Louise & Meskin, Aaron. To appear. *Aesthetic adjectives lack uniform behavior*. Inquiry. doi: 10.1080/0020174x.2016.1208927

Marchand, Hans. 1964. A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes. *Indogermanische Forschungen* 69: 10–19.

Marín, Rafael & McNally, Louise. 2011. Inchoativity, change of state, and telicity: Evidence from Spanish reflexive psychological verbs. *Natural Language and Linguistic Theory* 29: 467–502.
doi: 10.1007/s11049-011-9127-3

McNally, Louise. 2011. The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In *Vagueness in Communication. International Workshop, ViC 2009, held as part of ESSLLI 2009, Bordeaux, France, July 20–24, 2009. Revised Selected Papers*, Rick Nouwen, Robert van Rooij, Uli Sauerland & Hans-Christian Schmitz (eds.), 151–168. Berlin: Springer.

Mittwoch, Anita. 1991. In defence of Vendler's achievements. *Belgian Journal of Linguistics* 6(1): 71–85. doi: 10.1075/bjl.6.05mit

Nam, Seungho. 2005. Directional locatives in event structure: Asymmetry between goal and source. *Eoneohag ('Linguistics')* 43: 85–117.

Parsons, Terence. 1990. *Events in the Semantics of English*. Cambridge MA: The MIT Press.

Piñón, Christopher. 2008. Aspectual composition with degrees. In *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, Louise McNally & Christopher Kennedy (eds), 183–219. Oxford: OUP.

Piñón, Christopher. 1997. Achievements in an event semantics. In *Proceedings of Semantic and Linguistic Theory 7*, Aaron Lawson & Eun Cho (eds), 273–296. Ithaca NY: CLC Publications.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge MA: The MIT Press.

Rothstein, Susan. 2008. Telicity and atomicity. In *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect* [Linguistik Aktuell/Linguistics Today 110], Susan Rothstein (ed.), 43–78. Amsterdam: John Benjamins. doi: 10.1075/la.110.04rot

Sassoon, Galit W. 2013. Between-noun comparisons. In *The Dynamic, Inquisitive, and Visionary Life of ϕ, ?ϕ, and ◊ϕ: A Festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman*, Maria Aloni, Michael Franke & Floris Roelofsen (eds), 286–299. <http://www.illc.uva.nl/Festschrift-JMF> (1 November 2015).

Solt, Stephanie & Gotzner, Nicole. 2012. Who here is tall? Comparison classes, standards and scales. Paper presented at *Linguistic Evidence 2012*. Universität Tübingen, Tübingen: February 9–11, 2012.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 56(2): 143–160. doi: 10.2307/2182371

Author queries

Please provide the complete reference for this citation '(Düsseldorf 2011)' of this chapter.