



MASTERARBEIT | MASTER'S THESIS

Titel | Title

On the semantic structure of deadjectival verbs

verfasst von | submitted by
Justina Schindler B.A.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 867

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Allgemeine Linguistik:
Grammatiktheorie und kognitive
Sprachwissenschaft

Betreut von | Supervisor:

Univ.-Prof. Dr. Daniel Büring

Abstract

Degree Achievements like *to age*, *to grow* or *to open* are verbs that express a change of state with regards to a gradable property (Dowty 1979). The focus of this thesis are Degree Achievements (DAs) that are derived from gradable adjectives and their semantic properties. I examine the distinction between DAs with positive endstate readings (e.g. *to open* $\approx$ ‘to become open’) and those with comparative readings (e.g. *to weaken* $\approx$ ‘to become weaker’, $\not\approx$ ‘to become weak’). The type of reading for a given DA is dependent on the properties of the underlying adjective: If it is associated with an absolute scale, i.e. one with at least one endpoint, a positive endstate reading arises for the DA. If the adjective scale has no endpoints, the DA will have a comparative reading. Previous analyses have been unable to fully capture this behaviour. I discuss the shortcomings of existing analyses and show that the proposal by Kennedy & Levin (2008) can be adapted to make the correct predictions. My central claim is that the treatment of positive endstate readings in the literature as corresponding to a telic reading is misguided, and that in fact the two must be understood as separate concepts, which interact but are ultimately independent. A positive endstate reading is the result of the relationship between the scale structures of the adjective and its DA; telicity is determined purely by properties of the DA scale. The predictions made by Kennedy & Levin in combination with my perspective on what defines a positive endstate reading are tested against a small set of adjectives and DAs.

Zusammenfassung

Degree Achievements (DAs) wie *to age* ‘altern’, *to grow* ‘wachsen’ oder *to open* ‘öffnen’ sind Verben, die Zustandswechsel eines gradierbaren Prädikats ausdrücken (Dowty 1979). Im Fokus dieser Arbeit stehen die semantischen Eigenschaften von Degree Achievements im Englischen, die von gradierbaren Adjektiven abgeleitet sind. Ich untersuche die Unterscheidung zwischen DAs mit einer positiven Lesart (z. Bsp. *to open* ‘öffnen’ $\approx$ ‘offen werden’) und solchen mit einer komparativen Lesart (z. Bsp. *to weaken* ‘schwächen’ $\approx$ ‘schwächer werden’, $\approx$ ‘schwach werden’). Welche Lesart ein DA erhält, ist abhängig von den semantischen Eigenschaften des zugrundeliegenden Adjektivs. Hat es eine absolute Skala, also eine mit mindestens einem Endpunkt, ergibt sich eine positive Lesart. Bei einer offenen Skala ergibt sich eine komparative Lesart. Bisherige Analysen waren nicht in der Lage, dieses Verhalten vollständig korrekt abzubilden. Ich bespreche die Defizite dieser Analysen und zeige, dass die von Kennedy & Levin (2008) vorgeschlagene Ableitung mit einigen Änderungen die richtigen Vorhersagen trifft. Meine zentrale Behauptung ist, dass die Behandlung der positiven Lesart als grundsätzlich telisch fehlgeleitet ist. Stattdessen müssen die beiden Konzepte klar voneinander getrennt werden. Die positive Lesart ergibt sich aus der Beziehung zwischen den Skalenstrukturen von DA und Adjektiv; Telizität ist ausschließlich abhängig von den Eigenschaften der DA-Skala. Die Vorhersagen, die aus Kennedy & Levin’s Analyse mit den von mir vorgeschlagenen Änderungen resultieren, werden anhand eines kleinen Datensatzes ausführlich verifiziert.

Acknowledgements

I want to thank my supervisor Daniel Buring for sharing his honest enthusiasm, for his trust in my abilities, his encouragement and his helpful metaphors. I am grateful to Sinoël Dohlen for being my partner in life and work, and for our constant exchange of thoughts and ideas. I also thank Kriszta Szendrői for being a great teacher and helping me find my way in the beginning of this project.

I have received valuable feedback from Dalina Kallulli, Berit Gehrke, Marcin Wągiel and the lovely audience at the 75. Studentische Tagung Sprachwissenschaft. The members of Team Bib, who met every week in the library to write and keep each other company, helped me stay motivated and made me feel less alone.

Finally, I want to thank Friederike Schindler, Linde Gläser and Paula Baer for always believing in me and for answering every single phone call, day and night.

Contents

1	Introduction	1
2	Semantics of gradable adjectives	2
2.1	Gradability	2
2.2	Scales and standards	5
2.3	Diagnostics for adjectival scale structure	8
2.3.1	Endpoint diagnostics	9
2.3.2	Standard diagnostics	11
3	Semantics of deadjectival verbs	14
3.1	Degree Achievements	14
3.2	Telicity	15
3.3	DA readings	17
4	Analysing DAs	18
4.1	Properties of an accurate analysis	18
4.2	Dowty (1979)	21
4.3	Abusch (1986)	22
4.4	Hay et al. (1999)	27
4.5	Kennedy & Levin (2008)	29
4.5.1	Kriz (2011) and the availability of positive endstate readings	31
4.6	McNally (2017)	35
4.7	Conclusion	38
5	Empirical perspective	39
5.1	Adjectival properties	39
5.1.1	A note on colour adjectives	46
5.2	DA properties	47
5.3	Discussion	52
6	Conclusion	53
7	Bibliography	54

1 Introduction

Dowty (1979) introduces the term Degree Achievement (DA) for verbs expressing change to a gradable property. It is not surprising that many of these verbs are derived from gradable adjectives.

There is a central semantic distinction in deadjectival Degree Achievements: Some will receive a meaning of the form ‘become *A*’, like in (1-a), while others receive a meaning of the form ‘become more *A* (but not necessarily *A*)’, as in (1-b). The first is called a positive endstate reading, and the second a comparative reading.

- (1) a. The garage door has opened. $\approx$ The garage door has become open.
- b. The effects of the drug have weakened. $\approx$ The effects of the drug have become weaker.

It has been observed that there is a connection between the available DA reading and the semantic properties of the underlying adjective (e.g. Kennedy & McNally 2005:355): Adjectives with minimal or maximal degrees reliably yield positive endstate DAs, while those with fully open scales yield comparative DAs. This relationship is the focus of this thesis. While much progress has been made in the last forty years in the semantics of Degree Achievements, we still lack an exhaustive analysis that fully accounts for the above generalisation.

I propose that a central problem in existing approaches is the conflation of the positive endstate reading with telicity. Most authors treat them as equivalent, predicting that all positive endstate DAs will also be telic, and all comparative DAs will be atelic. As a result, the behaviour of DAs based on adjectives that have a minimal, but no maximal degree (e.g. *to dirty*), which yield atelic DAs with a positive endstate reading, cannot be fully explained. I argue that the defining characteristic of a positive endstate reading is not telicity, but the entailments made by the truth of the DA about the truth of the adjective. With this new notion of the DA readings, I show that the analysis presented by Kennedy & Levin (2008) is actually able to predict Kennedy & McNally’s (2005) generalisation.

The thesis is structured as follows. I begin by discussing some key semantic properties of gradable adjectives and their derived DAs. In section 2, I introduce gradability and its formalisation through scales, endpoints and standards. Sec-

tion 3 introduces the term Degree Achievement, the property of telicity and the distinction between the positive endstate and the comparative reading. Diagnostic tests are provided for all properties. The goal of section 4 is to determine what an accurate semantic analysis of deadjectival DAs should look like. Several influential analyses are discussed, starting with Dowty (1979) and leading up to McNally (2017). I arrive at the conclusion that Kennedy & Levin (2008) is the most accurate framework currently available, with the caveat of revising their conception of what makes a positive endstate reading. In section 5, the predictions made in section 4 are verified for a representative set of adjectives and their derived DAs.

2 Semantics of gradable adjectives

2.1 Gradability

Gradability is a pervasive adjectival property (in English); it is harder to think of a truly ungradable adjective than a gradable one. An adjective qualifies as gradable if it can be true of its argument to varying degrees. For example, it is possible to say that Maya is taller than Nick, but both are tall; *tall* is true of both Maya and Nick, but is true of Maya to a higher degree. Consequently, gradable adjectives can appear in comparative constructions, with degree modifiers such as *too*, *very* or *enough* and with *how* in *wh*-questions.

(2) comparative constructions

- a. This book was more interesting than I thought.
- b. Astronomy was the most interesting class I took in college.
- c. The story is not as interesting as you'd think.
- d. Cross-country skis are longer than downhill skis.
- e. The longest break of my day is during lunch.
- f. My commute to work is now twice as long as it used to be.
- g. The fresher the oyster the better it tastes.
- h. These must be the freshest eggs I have ever eaten.
- i. Our seafood is as fresh as possible.

- (3) degree modifiers
- a. The president's speech was very interesting.
 - b. The class was interesting enough to stay awake.
 - c. These skis are too long for beginners.
 - d. His hair is long enough to braid.
 - e. Her grief was too fresh to joke about.
 - f. My produce is very fresh.

- (4) questions with *how*
- a. How interesting is your job?
 - b. How long is a standard climbing rope?
 - c. How fresh do the eggs need to be for this recipe?

Classic examples of reliably ungradable adjectives include non-subjective (or privative) adjectives like *former* or *alleged*, physical descriptors like *striped* or *hexagonal*, and categorising expressions like *carnivorous* and *nuclear*.

- (5) a. #The 31st president is more former than the 42nd.
b. #He is a very former policeman.
c. #How former is this nuclear reactor?
- (6) a. #This is the most alleged murder that was ever tried.
b. #He is too alleged of a criminal.
c. #How alleged is this information?
- (7) a. #This shirt is more striped than that one.
b. #Your wallpaper isn't striped enough.
c. #How striped do you like your clothes?
- (8) a. #A pentagon is more hexagonal than a square.
b. #Honeycombs are very hexagonal.
c. #How hexagonal are pencils?
- (9) a. #Which is your most carnivorous plant?
b. #Lions are very carnivorous.
c. #How carnivorous is a wolf?
- (10) a. #This bomb is more nuclear than that one.

- b. #Our current bomb is not nuclear enough.
- c. #How nuclear is the energy created here?

Some adjectives may be misidentified as gradable when judged in non-standard uses. A number of the examples in (5)-(10) may be acceptable to certain speakers as a result of coercion. The unacceptability of (7-a) for example could be repaired with the assumption that what is being modified by the comparative is not the shirts' stripedness, but the amount of stripes that are contained in each pattern. (8-b) might be acceptable under the reading that honeycombs are close to being geometrically perfect hexagons; however, this is not the meaning of *hexagonal*, which simply refers to the property of having six sides, not the level of similarity to a hexagon. (10-c) could become acceptable under a reading more akin to 'How much of the energy created here is nuclear?', which again circumvents the problem of *how* actually modifying *nuclear*.

Coercion from ungradable to gradable adjectives is also frequent with nationality adjectives:

- (11) (Sudo 2015:27)
- a. Eric is very Japanese.
 - b. Eric is more Japanese than Yasu.

Here, again, the adjective *Japanese* is not truly compatible with comparison or degree modification; the examples only work if we assume that what is being compared or modified is the level to which Eric behaves like a stereotypical Japanese person or the number of Japanese people in his ancestry, but not his Japaneseness itself. When testing for gradability, it is important not to be misled by cases like these.

Kennedy & McNally (2005:359) treat *dead* as a gradable adjective, based on the uses in (12).

- (12)
- a. half dead
 - b. almost dead

Kearns (2007:74) on the other hand uses it as an example of an ungradable adjective. I too disagree with this classification. While there are of course uses of *dead* with degree-targeting morphology, I maintain that these are metaphoric

uses, and not examples of the literal adjective. *Half dead*, for example, is usually used to mean ‘very exhausted/ill’, not literally ‘halfway dead’. In addition, a person with a terminal illness is in no way more dead than a healthy person, just because they are closer to death. It thus seems misguided to assume a scale of deadness.

Colour adjectives are also often treated as classically gradable, which there is good reason to assume is inaccurate. They are discussed in detail in 5.1.1. In the following sections, I will be sticking to adjectives that are uncontroversially gradable, to avoid complications.

Finally, it should be noted that gradability is not a property exclusive to adjectives. It can also be observed in adverbs, verbs, nouns (s. (13)), and a number of other categories.

- (13)
- a. He ran faster than his sister.
 - b. Peter has grown more than Paul.
 - c. He is even more of an idiot than I expected.

Diagnostics for gradability

- compatibility with comparison
- compatibility with degree modification
- compatibility with *how*-questions

2.2 Scales and standards

It is generally assumed that each gradable adjective is associated with a **dimension** (like height, weight, length, etc.); the **degrees** are ordered along this dimension via an **ordering relation**, forming a **scale**. This explains why ungradable adjectives are unable to be compared, appear with degree modifiers or be used in *how*-questions: They do not come with a scale that could be targeted by these modifications.

We can now take a closer look at some properties in which individual scales can differ. The first one is the presence or absence of **endpoints**, meaning min-

(14) Peter is tall.

If Peter in (14) has a height of 165cm, (14) is usually not going to be judged as true, since the sentence uttered in isolation evokes a comparison among all humans (or maybe among all male humans). But if we now add the context that Peter is actually a ten-year-old boy, the comparison class changes to ten-year-olds; as a result, (14) suddenly becomes true, since 165cm is indeed tall for a fourth grader. So while the presence of endpoints is a stable lexical characteristic that only varies between predicates, the positioning of the standard can be variable within its scale.

Older accounts of gradability often extrapolate from this observation that adjectival standards are generally context-dependent. Kennedy & McNally (2005) however recognise that there are in fact adjectives for which the standard seems to be fixed, or, more precisely, fixed to an endpoint, like those in (15).

(15) (Kennedy & McNally 2005:356)

- a. The spot is visible.
- b. The glass is full.

Something usually counts as visible as soon as it can be seen even a little bit. Thus, a minimal degree of visibility is enough for the positive of the adjective to hold, no matter the context: *visible* has a **minimum standard**. *Full*, on the other hand, has a **maximum standard**: If something is called “full” we expect it to be completely full, not e.g. filled up half way, because if *full* holds of an argument, its maximal degree must be met².

Since the endpoint of a scale does not change depending on the contextual comparison class, these absolute standards are context-independent. Con-

²There are two pitfalls here: First, maximum standard adjectives like *full* often allow for some imprecision (Kennedy & McNally 2005:359); for example, you usually do not fill a cup to the very brim because it would simply be impractical. It suffices to get close to the standard without actually meeting it. The second possible problem is the so called ‘standard shift’ effect: *Full* for a wine glass clearly stops at a lower point than for a coffee mug; while the latter is often filled up as much as possible, the former is only to be filled about halfway. A possible answer to this is that in cases like these, it is not the standard itself that is dependent on the context, but the specifications of the scale and where it has its endpoints. The standard is still only defined as the maximal degree, wherever that may lie for a wine glass scale. Toledo & Sassoon (2011) offer an analysis that can account for this effect by reintroducing context dependency for absolute standards. What is important here is that none of these phenomena prevent us from treating adjectives like *full* as absolute standard adjectives.

sequently, it must be asked how, if not contextually, their standard is fixed. Kennedy & McNally (2005:355) observe that there is a correlation between standard orientation and scale type: Adjectives with at least one endpoint will always have a fixed endpoint standard, and only adjectives with fully open scales have a standard that is context-dependent (Kennedy & McNally 2005:361). We will call the former **absolute adjectives** and the latter **relative adjectives**, a distinction coined by Unger (1975:356), and the corresponding standards will be called absolute and relative standards. Kennedy (2007) interprets this behaviour as the result of **Interpretive Economy** (16).

- (16) **Interpretive Economy:** Maximise the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions. (Kennedy 2007:49)

This principle states a general preference for context-independent meaning contributions. When fixing gradable adjective standards, it ensures that an endpoint standard will be picked whenever possible, since an endpoint is a fixed point available as a lexical property. A midpoint (relative) standard cannot be assigned without referencing the context of the expression, since it is not inherently salient. Therefore, it is generally avoided and only resorted to if there is no context-independent alternative. Thus, while nothing prevents an absolute adjective from having a relative standard, it almost never happens. This gives us a complete typology of five possible adjectival scale types: Completely open with a relative midpoint standard, lower closed with a minimum standard, upper closed with a maximum standard, fully closed with a minimum standard, and fully closed with a maximum standard. The scales are shown in figure 2; standards are marked with a small vertical bar. All scale types are attested in natural language. We will see examples for each type in 2.3.

2.3 Diagnostics for adjectival scale structure

I will now introduce a selection of diagnostic tools that can be used to verify the scale structure and standard orientation of a given adjective. The same will be done on 3.2 and 3.3 for the relevant properties of deadjectival verbs.

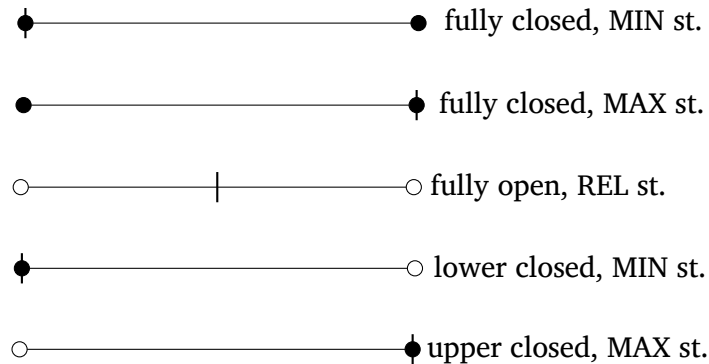


Figure 2: Scale structures with standards

2.3.1 Endpoint diagnostics

The presence of endpoints on an adjectival scale can be made visible with the help of endpoint-oriented degree modifiers such as *completely*, *slightly* and *half* (Kennedy & McNally 2005:354, Demonte 2011:1332, Sudo 2015:28, Lassiter 2015:159). To illustrate this, we can look at adjectives with very intuitive scale structures: *safe*, *dangerous*, *transparent* and *expensive*.

Safe has an upper closed and lower open scale, while *dangerous* has a lower closed and upper open one. It is easy to imagine what it means to be maximally safe and thus minimally dangerous. But there is no such thing as minimally safe, since there is no discernible natural upper limit to the level of danger that can be associated with something. *Transparent* has a fully closed scale; both maximal and minimal transparency are conceivable. Something fully opaque is minimally transparent, while something fully see-through is maximally transparent. *Expensive* has a fully open scale: There clearly is no maximum expensiveness, and while it might be tempting to argue that minimal expensiveness is the state of being free, having no price is not part of the scale of expensiveness (just as having no height is not part of the scale of tallness). Instead, the degrees at the lower end of the scale become smaller and smaller without ever reaching zero (e.g. Lehrer 1985:420). We have thus established the following scale structures for our for adjectives:

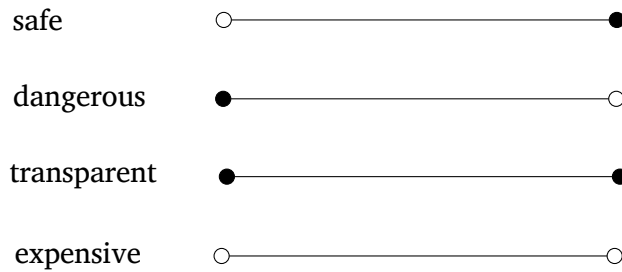


Figure 3: Scale structures

Completely relies on the presence of a maximum; it expresses that the maximum degree on the associated scale has been reached. Thus, it is compatible with upper closed scales (as in *safe* and *transparent*), but incompatible with upper open scales (as in *dangerous* and *expensive*).

- (17)
- a. Nowadays, bungee jumping is completely safe.
 - b. Please don't hitchhike alone in Mexiko; it's #completely dangerous.
 - c. These curtains make no sense! They're completely transparent!
 - d. I simply can't afford these shoes, they're #completely expensive.

Degrees expressed by *slightly* are very close to the lower end of the scale; this means it relies on the presence of a minimum, because without an end it is impossible to determine what is close to it. Thus, *slightly* will be compatible with lower closed, but not lower open scales.

- (18)
- a. You shouldn't go into the forest alone, it's only #slightly safe.
 - b. Your plan is good but slightly dangerous.
 - c. Microwave the rice until it is slightly transparent.
 - d. I do want these shoes but they are #slightly expensive.

The presence of a minimum can also be investigated indirectly via the application of *completely* to the adjective's antonym, based on the idea that antonym scales differ in nothing but polarity (s. 2.2). This means that the presence of a maximum in an adjectival scale entails the presence of a minimum in its antonym. Thus, the grammaticality of *completely safe* implies not only that *safe* has a max-

imum, but that *dangerous* has a minimum³⁴.

Half is dependent on both endpoints, since two points of orientation are necessary to determine a middle. It is thus only compatible with fully closed scales.

- (19)
- a. Mountain climbing is #half safe.
 - b. Be sure to take your phone with you, this part of the city is #half dangerous.
 - c. The bathroom door is half transparent so that you can see whether the light is on without intruding on anyone's privacy.
 - d. I might get these shoes but they are #half expensive.

With these three modifiers, we can test for minima and maxima in adjective scales. One should be wary of cases of apparent unexpected unacceptability where the modifier does not actually modify an adjectival scale, but one induced by another meaning component, e.g. the adjective's argument. "The sky is completely blue", for example, does not mean that the sky has reached a maximum level of blueness, but that every part of the sky is blue. We will encounter many such cases in the following sections.

Diagnostics for scale endpoints

- *completely* for maxima
- *slightly* for minima
- *half* for both

2.3.2 Standard diagnostics

A standard is the degree which must be exceeded in order for the positive of the adjective to apply. This means that if a positive adjective is truthfully predicated of an argument, it can be concluded that the argument meets the standard of the adjective's scale. Depending on the type of standard, there are different kinds of entailments that should hold. From their felicity or infelicity, we can conclude

³Provided that they are true antonyms, which I will not take the time to show here.

⁴The same instinct is reflected in Hay (1998)'s distinction of bounded vs. unbounded range adjectives.

whether we are dealing with a certain type of standard or not.

There are three types of entailment tests for standards we can employ. The first is about comparatives (Kennedy & McNally 2005:360): A MAX standard adjective in a comparative construction ‘*x* is more *A* than *y*’ entails that *y* (the entity with the lesser degree of *A*-ness) is not *A* (s. (20-a)). This is because any degree on a MAX standard scale for which there exists a higher degree must be below the standard.

- (20) a. Going during the day is safer than going at night.
 $\models$ Going at night is not safe.
 b. Cave diving is more dangerous than wreck diving.
 $\not\models$ Wreck diving is not dangerous.

A MIN standard adjective as in (20-b) would not lead to such an entailment. Instead, it would entail that *x* (the entity with the higher degree of *A*-ness) is *A*, which in turn does not hold for a MAX standard adjective. The reason is that any degree on a MIN standard scale for which there exists a lower degree must be above the minimum and thus above the standard.

- (21) a. Going during the day is safer than going at night.
 $\not\models$ Going during the day is safe.
 b. Cave diving is more dangerous than wreck diving.
 $\models$ Cave diving is dangerous.

Adjectives with a relative standard will have neither of these entailments, nor their inverses, since their standards could be anywhere: The relation of the two degrees to each other tells us nothing about their relation to the standard. This is illustrated in (22).

- (22) a. Maya is taller than Nick.
 $\not\models$ Maya is tall. $\not\models$ Maya is not tall.
 $\not\models$ Nick is tall. $\not\models$ Nick is not tall.
 b. My haircut was more expensive than my phone.
 $\not\models$ My haircut was expensive.⁵ $\not\models$ My haircut was not expensive.

⁵This interpretation may be derived based on the pragmatic assumption that phones always exceed the standard of expensiveness. Keep in mind that (22-b) would also be true in a context where I got a used flip phone for 20€ and a haircut for 35€.

⊭ My phone was expensive. ⊭ My phone was not expensive.

The second type of entailment influenced by standard orientation concerns adjectives compatible with partial degree modifiers like *half* or *partially* (Kennedy & McNally 2005:360): If an adjective *A* targeted by such a modifier entails that the unmodified positive *A* does not hold, it must have a maximum standard, since for a maximum standard adjective, a partial degree always lies below the standard, as seen in (23-a) and (23-b). For minimum standards this is never and for relative standards not necessarily the case. For minimum standard adjectives this can be illustrated (e.g. (23-c)), but for relative standard adjectives, it is a moot point, since there are no relative adjectives compatible with partial modifiers in the first place (as these modifiers require endpoints to be meaningful).

- (23) a. This cup is half full. ⊨ This cup is not full.
 b. The door is partially closed. ⊨ The door is not closed.
 c. The door is partially open. ⊭ The door is not open.

The third type of entailment can be observed when contrasting positives with comparatives (Kennedy & McNally 2005:359). For adjectives with a MAX standard, a truthful positive implies that a further increase on the scale is impossible, since the maximal value has already been reached. Thus the unacceptability of (24-a). The same is not true for adjectives with a minimum or relative standard, as illustrated by (24-b) and (24-c).

- (24) a. The professor said her class is full, #although there are still a couple of spots left⁶.
 b. Flying on a plane is dangerous, but still safer than driving a car.
 c. A new laptop is expensive, but getting the old one fixed is even more expensive.

With the help of these entailments, we can identify standard types in the gradable adjectives we will examine in 5. It is important to note that not all of these tests are equally powerful. Firstly, while all three allow us to distinguish MAX from nonmaximal standards, only the first enables us to also reliably distinguish

⁶Again, this might not be as clear under an imprecise reading (essentially reading *full* as closer too *quite full*). We can fix this with a context that forces precision, e.g. 'I emailed my professor to see if I could still register for her class even though I missed the deadline, but she said no.'

between MIN and REL standards. Secondly, while the first and third tests can be applied to all gradable adjectives, the second one is only applicable to those compatible with proportional modifiers (i.e. those with fully closed scales). This means it can help identify MIN standards as non-relative, since any adjective to which it applies must be absolute.

Diagnostics for adjective standards

- ‘ x is more A than y ’ $\models y$ is not A for MAX, $\models x$ is A for MIN, no such entailments for REL
- ‘half/partially A ’ $\models x$ is not A for MAX
- x is A is incompatible with statements about the possibility of a degree increase for MAX (e.g. #but it could be more A /#but it isn’t fully A)

3 Semantics of deadjectival verbs

The following section will give an overview of important semantic concepts associated with deadjectival Degree Achievements. I will first introduce the term Degree Achievement, and then discuss telicity and the distinction between the comparative and the positive endstate reading. Diagnostics will be given at the end of each section.

3.1 Degree Achievements

The term Degree Achievement was first introduced by Dowty (1979:88) as a description for change of state verbs that express change with regards to a gradable property (like age, height or temperature), e.g. (25-a) and (25-b). He observes that in contrast with non-gradable change of state verbs (cf. (25-c)), they are compatible with durational modifiers.

(25) (Dowty 1979:88)

- a. The soup cooled for ten minutes.
- b. John aged forty years during that experience.
- c. Mary became pregnant #for a month.

Degree Achievements (henceforth DAs) include derived as well as lexical verbs; this thesis focuses on DAs derived from adjectives. In the following, *DA* should be read as shorthand for *deadjectival DA*.

There are a number of morphological strategies to derive a DA from an adjective. The most common are the suffixes *-en* (*weak* → *to weaken*) and *-ø* (*empty* → *to empty*). These forms can also appear in combination with prepositions (*to freshen up*, *to calm down*). Less frequent are *-ify* (*diverse* → *to diversify*) and *en-* (*large* → *to enlarge*). In the following, I will steer clear of DAs with prepositions, to avoid interference of the meaning contributed by the prepositions with the applied diagnostic tests (it appears, for example, that the prepositions are able to add a telos). We will encounter mostly *-en* and *-ø* forms.

There are also the forms where derivation surfaces as vowel change, e.g. *full/fill*. Since for such cases it is hard to ensure the verbs are deadjectival and not the other way round, I will not be looking at these forms here.

3.2 Telicity

The distinction between telic and atelic predicates was introduced by Vendler (1967:102), who uses the terms ‘activity’ and ‘accomplishment’ for atelic and telic durative predicates respectively. The defining characteristic of telicity is the presence of an intrinsic endpoint (telos) to the action that is being described: A predicate like *to make a fire*, which is only complete if the fire is made, is telic, while a predicate like *to sleep*, which has no specific endpoint to reach, is atelic. Atelic predicates are homogeneous: on a timeline of the expressed action, there are many points for which the predicate holds, and every subevent of the denoted event will count as such an event by itself. For example, any subpart of a sleeping event will also qualify as a sleeping event, since sleeping for even a second is already sleeping; but not every subevent of building a house will qualify as building a house. If only the foundation has been constructed, a house has not been built.

Diagnostic tests for telicity are well-known and straightforward. Arguably the most popular relates to durational phrases: Atelic predicates like *sleep* are compatible with *for*-phrases but incompatible with *in*-phrases (s. (26)); for telic predicates, it is the other way round (s. (27); Vendler 1967:100,01).

- (26) atelic:
- a. Mo slept for an hour.
 - b. #Mo slept in an hour.
- (27) telic:
- a. #He made a fire for ten seconds⁷.
 - b. He made a fire in ten seconds.

The second test, due to Dowty (1979:58), is based on the interpretations that arise when the verb occurs in combination with *almost*. Telic predicates with *almost* are ambiguous between a reading where the action was almost started and one where it was started and almost finished. Atelic predicates do not display this ambiguity; they can only produce the *almost-started*-reading).

- (28) a. He almost made a fire.
b. He almost slept.

(28-a) can either be interpreted as the subject having almost started to make a fire, or as the subject having been in the process of making a fire but stopped just short of achieving their goal. The subject in (28-b) on the other hand can only be understood as never having started sleeping, since there is no telos in the action of sleeping that could almost be reached.

A third possible test for diagnosing telicity is the compatibility with *to finish*. Only telic predicates are acceptable as complements of *finish*, since atelic predicates have no natural end (e.g. McNally 2017:185).

- (29) a. He finished making a fire.
b. #He finished sleeping.

⁷Here you can get a false positive caused by coercion, e.g. “He wrote his thesis for two weeks.” being read as ‘He worked on his thesis for two weeks.’, where the telic predicate is reinterpreted as atelic. Confusion may also be caused by structural ambiguity, e.g. “He became mayor for ten months.” where *for ten months* modifies *mayor* instead of *become mayor* (as it should for the test to be meaningful).

Diagnostics for telicity

- durational adverbials: *in* with telic predicates, *for* with atelic predicates
- *almost*: ambiguous between ‘almost started’ and ‘almost finished’ for telic predicates, unambiguously ‘almost started’ for atelic predicates
- *to finish*: only telic predicates acceptable as complements

- ### Diagnostics for telicity
- durational adverbials: *in* with telic predicates, *for* with atelic predicates
 - *almost*: ambiguous between ‘almost started’ and ‘almost finished’ for telic predicates, unambiguously ‘almost started’ for atelic predicates
 - *to finish*: only telic predicates acceptable as complements

3.3 DA readings

Every DA derived from an adjective *A* receives one of two readings. One can be paraphrased as ‘become *A*’; it expresses that the positive of the underlying adjective holds, i.e. that the adjectival standard is exceeded. This is the case for *to empty*, as shown in (30-a): If a cup has been emptied, we expect it to be empty. The other reading can be paraphrased as ‘become more *A*’. It entails that the degree to which the underlying adjective applies has increased, but not that the standard has been reached; the positive of the adjective does not necessarily hold. This is the case for *to weaken* (s. (30-b)): If the effects of a drug have weakened, they are expected to be weaker than before, but may still be quite strong.

- (30) a. He emptied the cup. positive endstate reading
 $\models$ The cup is empty.
 ($\models$ The cup is emptier than before.)
 b. The effects of the drug have weakened. comparative reading
 $\not\models$ The effects of the drug are weak.
 $\models$ The effects of the drug are weaker than before.

Following Kriz (2011) I will call the former the **positive endstate reading** and the latter the **comparative reading**. I will use the term **DA reading** to refer to the general property of having one or the other.

A positive endstate reading can be distinguished from a comparative reading by contrasting the perfect and the negated present of the same DA. If they are compatible, the DA has a comparative reading; if not, it has a positive endstate reading.

- (31) a. The gap has widened but it isn't wide.
b. The shirt has dried #but it isn't dry.

Telicity and DA reading are often conflated; *telic* is used to mean 'positive end-state reading' and *atelic* to mean 'comparative reading'. We will see in the following two sections how this is problematic. I propose that the concepts are related, but separate.

Diagnostic for DA readings

- '*x* has *A*-ed but it isn't *A*.' – acceptable for comparative reading, unacceptable for positive endstate reading

With this conceptual basis established, we can now take a look at some analyses of the semantic of deadjectival verbalisation.

4 Analysing DAs

The purpose of this section is to establish a clear picture of what a semantic analysis of deadjectival verbs and their readings should look like. We are interested in how DA reading and telicity are semantically related to each other and to the underlying adjective, and how Kennedy & McNally's (2005) generalisation that relative adjectives tend to produce comparative readings and absolute adjectives positive endstate readings can be accounted for. Empirical support for this idea will be given in 5.2.

4.1 Properties of an accurate analysis

For DAs that are not lexical, but derived from adjectives, it is clear that an analysis of DAs must of course correctly implement the relationship between the adjective semantics and the meaning of the derived verb. So we need to determine what meaning is added to the basic adjective meaning by the process of verbalisation. We have already roughly answered this question in 3.3: verbalisation transforms an adjective meaning *A* into a DA reading 'become (more) *A*', depending on whether the reading is positive or comparative.

First, I should make it explicit that assuming that each DA has only one or at least one dominant reading is a conscious decision (which as we will see is corroborated by a number of existing analyses). There is a claim going back to Abusch (1986) that all DAs are in fact ambiguous between the comparative and the positive endstate reading; I will illustrate why I do not agree with her reasoning in 4.3. I maintain that while there are certainly cases in which a comparative reading can be coerced out of a positive endstate DA, this is by no means equivalent to straightforward ambiguity.

Starting from here, there are two major perspectives that could be taken.

1. The comparative reading and the positive endstate reading are the result of fundamentally different meanings – therefore, **two different semantic structures** must be assumed. This offers a very straightforward explanation of the observed readings; essentially, what you see is what you get.
2. There is **one unified semantics** for DAs. The different readings fall out from how the verbalisation interacts with the lexical semantic properties of the adjective.

While it might require less abstraction, the downside of the approach 1 is that it necessitates an explanation of how and when it is decided which of the two available DA semantics applies in each case. It also ignores the substantial similarities between the two readings, losing points for redundancy. Approach 2 is not only more efficient but has the upside of incorporating the intuition that the readings are two sides of the same coin (as discussed in detail below); this is so essential that it already suffices to make approach 2 the clear favourite. It is not surprising that among the existing analyses discussed here, not a single one assumes two completely independent underlying semantic forms for positive endstate and comparative DAs.

Adopting this approach means we must now explore what a unified semantic structure of verbalisation should look like. There are three ways to go.

- 2a. Assuming a **comparative core**: The semantics added by verbalisation to a gradable adjective encodes the meaning of ‘become more *A*’. This is the default DA reading; positive endstate readings have to be derived from a comparative basis. Examples of this type of analysis is Hay et al. (1999) and Kennedy & Levin (2008).

- 2b. Assuming a **positive core**: The semantics of DAs adds a positive meaning ‘become *A*’. The default DA reading is thus a positive endstate reading, and the comparative readings have to be derived. Dowty (1979) and Abusch (1986) take this approach.
- 2c. Assuming a **vague core**: The semantics of DAs is underspecified in such a way that there is no underlying default reading. Both have to be derived based on the semantic properties of the underlying adjective, the DA’s argument, pragmatic principles etc.

Motivation for a vague core would come from universal ambiguity; since we already ruled this out, we can focus on the other two approaches. The entailments in (30), repeated here as (32), illustrate an important intuition: The positive endstate reading is a stronger version of the comparative reading. The comparative meaning is entailed by both, since becoming *A*, i.e. going from a state of not being *A* to a state of being *A*, logically requires becoming more *A*, i.e. going from a state where *A* applies to a degree d_n to a state where it applies to a degree $d_m > d_n$. The difference between the readings comes down to whether they include the positive meaning or not.

- (32) a. He emptied the cup. positive endstate reading
 $\models$ The cup is empty.
 ($\models$ The cup is emptier than before.)
 b. The effects of the drug have weakened. comparative reading
 $\not\models$ The effects of the drug are weak.
 $\models$ The effects of the drug are weaker than before.

These facts are captured best by a comparative core analysis. Having established this, the next step is to determine what causes the surface differences, i.e. how the comparative reading is strengthened to yield a positive endstate interpretation. There are a number of different possible answers, which will be discussed in the rest of this section. First, however, we should determine what characterises the DA readings.

As we saw above, a positive endstate reading arises when a DA derived from an adjective *A* is interpreted as ‘to become *A*’. This means: if the DA is predicated of an argument, it is entailed that *A* will hold for that argument. Thus, “He

emptied the cup” entails that the cup is empty. In terms of the established semantics for gradable adjectives, that means: if the predicate holds, it is entailed that the adjective’s standard is met.

A comparative reading is present when this is not the case: If the DA is predicated of an argument, we know that the argument’s degree of *A*-ness has increased, but it is not entailed that it has increased enough to exceed the standard; it is not entailed that the positive of the adjective holds. “The effects of the drug weakened” entails that the effects are weaker than before, but not that they are weak. The truth of the predicate allows no entailment about the position of the final degree in relation to the adjective standard.

So, crucially, these readings are the results of what entailment relation holds between the DA and its underlying adjective scale: If something about the DA or the adjective makes it possible to safely predict from the truth of the DA that the adjective standard is met, we arrive at a positive endstate reading. If this is not the case, we arrive at a comparative reading. We will see below that from the analysis proposed by Kennedy & Levin (2008), such predictions can in fact be made.

We have established that an accurate DA analysis should propose a unified comparative semantics and reveal what property allows the correct conclusions to be drawn from the DA to yield a positive endstate reading. With these expectations in mind, we can now proceed to look at different analyses that have been proposed.

4.2 Dowty (1979)

Dowty (1979:90) assumes a positive core (shown in (34)), with the reasoning that a comparative core is not motivated by the morphology of DAs. His analysis is the first of its kind, and not worked out in detail, but is a clear first push towards a unified DA semantics. It is based on the assumption that all gradable adjectives are vague in the sense that they have relative standards; in this case, relative not with regards to an argument-based comparison class, but in comparison to an earlier state. The meaning of a DA is not one extended change of state, but actually consists of a series of punctual state changes linked together, forming the durational event denoted by the DA.

$$(34) \quad \lambda x[BECOME(cool(x))]$$

As discussed above, a positive core analysis like this one is very counterintuitive. The reason this is not considered is the fact that according to Dowty (1979:88), all DA readings are positive. Assuming a comparative core would lead to additional steps in the analysis. I suspect the eventual recognition of comparative readings is the reason why only the earliest analyses assume a positive core.

Abusch (1986:5,6) recognises both the existence of non-punctual and comparative DAs. She designs a semantic structure that can derive more than just momentary state changes. She bases the analysis on Dowty's, keeping his BECOME-core, but makes it context-dependent (s. (35-a)). The context variable c can be filled by the context of the utterance c_u as in (35-b). This then yields the positive end-state reading ' x became cool in the relevant context'. To derive a comparative reading, c would be existentially bound, as in (35-c), resulting in the meaning ' x became cool in some context'. Since we do not know how this relates to the current context, the reading is not positive.

- As mentioned above, Abusch (1986:6) suggests that all DAs (which she calls

22

“inchoatives”) are ambiguous between the positive endstate and the comparative reading. I now want to give a quick overview of her reasoning and explain why I disagree, supporting my instinct that her example *to cool* is in its core a DA with a comparative reading.

Abusch begins with the data in (36), arguing that the fact that *to cool* patterns with the comparative and not the positive of the underlying adjective in the distribution of the given modifiers shows that this DA allows a comparative reading. She then gives (37) as an example of the DA patterning with the positive and not the comparative, which she takes as an indication that the same DA can also receive a comparative reading. Based on this, and the implicit assumption that other DAs would behave similarly, she concludes that DAs in general are ambiguous between the two readings.

	Positive	Comparative	DA
	quite cool	*quite cooler	*has cooled quite
	so cool	*so cooler	*has so cooled
(36)	*a lot cool	a lot cooler	has cooled a lot
	*quite a bit cool	quite a bit cooler	has cooled quite a lot
	*ten degrees cool	ten degrees cooler	has cooled ten degrees
			(Abusch 1986:6) ⁹

	Positive	Comparative	DA
(37)	completely cool	*completely cooler	has cooled completely
	absolutely cool	*absolutely cooler	has cooled absolutely
			(Abusch 1986:6)

This argumentation is flawed in several ways. To start with, I would like to contest the relevance of the unacceptability of “ten degrees cool” in (36). This is not a sign of the general behaviour of adjectives, but a result of the fact that only the positive members of antonym pairs are unmarked with measure phrases (e.g. Seuren 1978, Lehrer 1985): “ten degrees warm” would be perfectly acceptable. Overt measure phrases are not a case where positive adjectives are truly separated from comparatives and DAs.

Furthermore, I think the acceptability judgements in (37) are not accurate. Both “completely cool” and “absolutely cool” strike me as squeamish at best;

⁹Dataset has been slightly shortened.

my intuition is that there is no natural highest point of coolness. The Corpus of Contemporary American English (COCA; Davies 2008) lists no instances of “completely cool” and only a handful for “absolutely cool”, all of which either use a different meaning of *cool* or have an underlying structure where the modifier and the adjective do not actually form a constituent. There are also no entries for “has cooled absolutely”. I must admit that I have no idea what a possible reading of *absolutely cool* would be. As for *completely cool* I suspect that Abusch’s judgement is the result of an argument induced scale: What is being modified by *completely* is not the degree of coolness, but the level to which the argument is affected (e.g. “The house has cooled completely” does not mean that the house is maximally cool, but that every part of the house has been affected by the cooling). With all judgements corrected, the dataset (37) would look like (38): the DA, comparative and positive all pattern together.

	Positive	Comparative	DA
(38)	*completely cool	*completely cooler	*has cooled completely
	*absolutely cool	*absolutely cooler	*has cooled absolutely

It must be acknowledged here that there does exist a use of *to cool* where it can be modified by *completely*, which is specific to the context of food production, like (39). In these cases, the verb has the very particular meaning of letting something that has been heated realign either with room temperature or a temperature fit for consumption. Such clearly defined end goals mean that in these cases, the scale of *to cool* gets a maximum, which is the reason why *completely* is suddenly compatible (cf. 2.3.1) and a positive endstate reading arises¹⁰.

- (39) a. The soup has cooled completely.
b. Let the cake cool completely before adding the frosting.

All entries for “cooled completely” in COCA are food related¹¹, which leads me to the conclusion that the upper closed *to cool* is to be viewed as a special case. I will stick to the conviction that in general, *to cool* is a DA with no maximal degree and a comparative reading, like in the examples in (40), which show that “*x* has cooled” can be uttered in contexts where *x* does not qualify as cool.

¹⁰The same idea is proposed by Hay et al. (1999), which I discuss in 4.4.

¹¹Except one, which is from a highly coercive literary context, and shall thus be ignored.

- (40) a. The planet has cooled to a temperature that would allow humanoid life.
 b. The water in the whirlpools has cooled since they started their energy saving campaign.
 c. The air has cooled because of the rain; it's only 25°C now.

In spite of their being not representative for the use of *to cool* and the behaviour of DAs in general, examples about cooling soup are omnipresent in DA semantics research. They originated in Lakoff (1965), and resurface in Dowty (1979), Abusch (1986), Kennedy & Levin (2008), Kriz (2011), and many other works. I suspect that this is one of the reasons why the idea of all DAs being ambiguous was able to survive until the present. It seems rather unlucky that this vein of research was started on such a misleading example.

As we saw in 2.3.1, whether *completely* can modify a scale is dependent on the presence of a maximum. The fact that it turns out not be compatible with the positive or the DA in (38) supports the idea that neither has a maximum, i.e. that the adjective *cool* is relative and the DA *to cool* has a comparative reading¹². We can further reinforce this idea by replacing *cool* in (38) with something clearly relative like *to widen* (41) or something clearly upper closed like *to empty* (42).

	Positive	Comparative	DA
(41)	*completely wide	*completely wider	*has widened completely
	*absolutely wide	*absolutely wider	*has widened absolutely
	Positive	Comparative	DA
(42)	completely empty	*completely emptier	has emptied completely
	absolutely empty	*absolutely emptier	*has emptied absolutely

As expected, none of the forms allow *completely* when there is no maximum to be targeted, but the positive and the DA do allow it when there is one. The comparative never allows it, not because of the specific adjective, but because of properties inherent to comparison¹³.

The central point of this argumentation is the following: While the patterns that Abusch observes do exist, they are not a symptom of variable DA readings. In

¹²The relationship between endpoints and DA readings will become more explicit in 4.5.

¹³*Absolutely* seems to be unacceptable for independent reasons, maybe a general incompatibility with verb scales.

(36), the DA patterns with the comparatives because both are comparing something: comparatives compare the degree at issue to some other degree, and DAs compare the final to the initial degree of their event. This is the case for every DA, no matter its reading, further supporting the idea of the comparative core. Positives on the other hand have no comparative meaning component, and thus do not allow modifiers that measure differences; instead they allow modifiers that refer to the positioning of a single degree on a scale. In (37), the DAs pattern with the positives because the presence or absence of a maximum is always the same for both forms¹⁴, no matter the reading of the DA, as was illustrated in (41) and (42).

Abusch's final argument in support of ambiguity is the contrast in (43). According to her, the compatibility with *considerably* in (43-b) implies gradability and a comparative meaning. I have no issue with this.

- (43) (Abusch 1986:6)
- a. The weather has finally cooled.
 - b. The weather has cooled considerably.

The use with *finally* in (43-a), she says, "implies an end state" in the reading of the DA. This in my opinion is a gross misinterpretation. *Finally* clearly does not modify the DA, but the entire clause: (43-a) means that after a period of waiting, the event of the weather cooling has now occurred. What has ended is the process anticipating the weather change; not the weather change itself. Whether the event that was anticipated has a positive endstate or not is irrelevant; in fact, here I would argue it is comparative. What (43-a) essentially means is that after a long period of waiting, the event has now occurred which enables us to say that the temperature is lower than it used to be; not by any means maximally low.

From all of the above, I conclude that assuming ambiguous DA readings is not motivated by Abusch's data. An additional problem with Abusch's analysis is the fact that there is no way to derive the correct readings for DAs with explicit measure phrases, such as (44), since the semantics does not allow measuring the difference between the start and end degrees.

¹⁴Again, a detailed explanation of the reasons will be given in 4.5.

(44) The gap widened 2cm.

As we will see in the following sections, Hay et al. (1999) and Kennedy & Levin (2008) improve on this.

4.4 Hay et al. (1999)

Hay et al. (1999) make the leap towards a comparative core. They acknowledge that there are insufficient arguments in favour of a systematic ambiguity between different DA readings (Hay et al. 1999:133) and instead opt for an analysis that assumes a universal comparative semantics for all DAs, which on its own yields an atelic/comparative reading¹⁵ but is augmented to a telic/positive endstate reading under the influence of material from the semantic and pragmatic context.

The semantic contribution made by verbalisation of a gradable adjective is the INCREASE function, shown in (45).

$$(45) \quad \llbracket \text{INCREASE}(A)(x)(d)(e) \rrbracket = 1 \text{ iff } A(x)(\text{SPO}(e)) + d = A(x)(\text{EPO}(e))$$

(Hay et al. 1999:132)

(46) telic reading:

- a. The gap widened by 2cm.
- b. $\exists e[\text{INCREASE}(\text{long}(\text{rope}))(2\text{cm})(e)]$

(47) atelic reading:

- a. The gap widened.
- b. $\exists e, d[\text{INCREASE}(\text{long}(\text{rope}))(d)(e)]$

The function takes a gradable adjective meaning A , an individual x , a degree d and an event e and returns True if the degree to which A applies to x at the starting point (SPO) of e added to d is equal to the degree to which A applies to x at the end (EPO) of e . So INCREASE picks out the degree by which the argument's A -ness has increased. This degree is called the **difference value**.

Hay et al. (1999:137) assume a general preference for telic readings, because they are more informative than atelic readings. The deciding factor for telicity

¹⁵Hay et al. only talk about telic and atelic readings, but I maintain that the terms *positive end-state reading* and *comparative reading* are more suited to what they are talking about, as discussed in 4.5.1.

lies in the boundedness of the difference value: If a specified difference value is available, then the DA is telic (s. (46)); if not, the degree argument is existentially bound (s. (47)) and the DA is atelic. There are a number of different parts of an expression that could contribute a *d*. The most straightforward is an explicit measure phrase, as in (48-a), giving the exact difference value, or a degree modifier like *completely* (s. (48-b)), which picks out a specific point on the scale as the final degree (Hay et al. 1999:133,34).

- (48) a. The gap widened by 2cm.
b. The clothes dried completely.

Additionally, the underlying adjective scale can provide a difference value by having a maximum¹⁶, like in (49), which provides an upper bound for the difference value, leading to a telic reading (Hay et al. 1999:136).

- (49) The rope straightened.

If the semantics cannot provide a difference value, then according to Hay et al. (1999:136) it can also be inferred from the context, similarly to what I proposed in 4.3 for *to cool* in cooking contexts.

- (50) Kim lowered the blinds (completely).

Since for sentences like (50), there is a practical endpoint, the DA in this context can be interpreted as having a maximum, even if it doesn't usually. This maximum can then act as an upper bound for the difference value, resulting in a telic reading. Since in these cases, telicity is the result of an implicature, it can easily be cancelled:

- (51) a. Kim lowered the blinds, but not completely.
b. The soup has cooled, but only slightly.

This helps explain the apparent readiness with which some DAs seem to change readings, which was interpreted by Dowty and Abusch as a sign of underlying ambiguity: It is a pragmatic effect on top of atelic semantics.

¹⁶Hay et al. work with the bounded- vs. unbounded-range distinction (Hay 1998), which classifies gradable adjectives not according to the presence or absence of endpoints in general, but the presence or absence of maxima alone. The relevance of minima is not yet acknowledged.

All these factors taken together mean that the only truly atelic cases of DAs are those that do not have an inherent maximum, occur without explicit degree morphology and in a context that does not provide a pragmatically relevant endpoint.

By assuming a comparative core, including the difference value in the semantics and dispensing with Abusch’s blanket ambiguity, this analysis is a big step in the right direction.

4.5 Kennedy & Levin (2008)

Kennedy & Levin (2008) follow Hay et al. (1999) in assuming a comparative core and a semantics that references a difference value. Theirs is the first analysis to fully explain why DAs based on adjectives with maxima are usually telic, while DAs based on relative adjectives are atelic.

According to Kennedy & Levin (2008:12), gradable adjectives denote **measure functions** of type $\langle e, d \rangle$, which take an individual and give back the degree it is associated with on the adjectival scale. If the adjective occurs in combination with overt degree morphology like in (52-a), a phonologically empty function μ , shown in (53), compares this degree to the overtly given degree, returning True if it exceeds it. If there is no overt material, like in (52-b), the adjective is instead combined with the phonologically empty *pos*, shown in (54), which compares the degree to the adjective’s standard, returning True if it is exceeded.

- (52) a. Maya is 2m tall.
b. Maya is tall.

$$(53) \quad \llbracket \mu \rrbracket = \lambda g \in D_{\langle e, d \rangle} \lambda d \lambda x \lambda t. g(x)(t) \geq d \quad (\text{Kennedy \& Levin 2008:23})$$

$$(54) \quad \llbracket pos \rrbracket = \lambda g \in D_{\langle e, d \rangle} \lambda t \lambda x. g(x)(t) \geq \text{standard}(x) \\ (\text{Kennedy \& Levin 2008:13})$$

Degree Achievements are built in parallel to the adjectival comparative, so we must first understand how the comparative works. Kennedy & Levin (2008:17) analyse the comparative as a so-called **difference function** $m_d^\uparrow$, s. (56), which is similar to the original adjective scale except for the fact that it is essentially cut off at the degree to which we are comparing. So for example in “Maya is

taller than Nick”, the difference function expressed by *taller than Nick* would be a version of the scale associated with *tall*, but with a minimum at Nick’s height. The comparative morpheme *-er* (57) takes Nick’s height and the scale of *tall* and returns the corresponding difference function. This, in turn, also occurs in combination with either μ or *pos*, depending on whether an overt difference degree is expressed or not, as in (55-a) vs. (55-b).

- (55) a. Maya is 10cm taller than Nick.
b. Maya is taller than Nick.

(56) **Difference function** (Kennedy & Levin 2008:17)

For any measure function m from objects and times to degrees on a scale S , and for any $d \in S$, $m_d^\uparrow$ is a function just like m except that:

- i. its range is $\{d' \in S | d \leq d'\}$, and
ii. for any x, t in the domain of m , if $m(x)(t) \leq d$ then $m_d^\uparrow = d$.

(57) $\llbracket \text{-er} \rrbracket = \lambda x \lambda g \in D_{\langle e, d \rangle} \cdot g_{g(x)}^\uparrow$ (Kriz 2011:76)¹⁷

For deriving DAs, Kennedy & Levin propose a very similar process: The adjective scale is copied onto the DA, but cut off at the degree that the argument is associated with at the beginning of the process expressed by the DA. This is called a **measure of change function** m_Δ , which is shown in (59). So in “The gap widened”, the scale of *widen* would be the scale of *wide* but cut off at the degree to which the gap was already wide at the beginning of the widening process, reflecting the instinct that DAs in their core express comparison (cf. section 4.3). The verbalising affix (here *-en*) takes on the task of creating the measure of change function from the adjective and the initial degree (s. (60)). The measure of change function then appears with verbal versions of μ in cases like (58-a) and *pos* in cases like (58-b), which are shown in (61) and (62) respectively.

- (58) a. The gap widened 2cm.
b. The gap widened.

¹⁷Kennedy & Levin never give an explicit denotation for this morpheme, but Kriz (2011) infers accurately what it should look like. The same goes for (60).

- (59) **Measure of change** (Kennedy & Levin 2008:18)
 For any measure function m , $m_{\Delta} = \lambda x \lambda e. m_{m(x)(init(e))}^{\uparrow}(x)(fin(e))^{18}$.
- (60) $\llbracket -en \rrbracket = \lambda g \in D_{<e,d>}. g_{\Delta}$ (Kriz 2011:77)
- (61) $\llbracket \mu_v \rrbracket = \lambda g \in D_{m_{\Delta}} \lambda d \lambda x \lambda e. g(x)(e) \geq d$ (Kennedy & Levin 2008:24)
- (62) $\llbracket pos_v \rrbracket = \lambda g \in D_{m_{\Delta}} \lambda x \lambda e. g(x)(e) \geq standard(g)$
 (Kennedy & Levin 2008:19)

So in this account, DAs have their own scales, which are derived from the underlying adjective, and consequently must also have standards. These are assigned according to Interpretive Economy, repeated here as (63), and Hay et al.'s principle of informativity (Hay et al. 1999:137), which states that telic readings are always preferred over atelic readings. A telic reading arises when the DA has a maximum standard, since in that case there is only one possible endstate (Kennedy & Levin 2008:8); so effectively, the informativity principle expresses a preference for maximum standards over non-maximal standards in DAs.

- (63) **Interpretive Economy:** Maximise the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions. (Kennedy 2007:49)

The way in which DA scales are derived means that they are always absolute, since every DA scale has at least a minimum. The presence of a maximum is dependent on the presence of a maximum in the underlying adjective scale. If a maximum is available, it is always the standard, because of informativity; if not, the minimum is the standard. So in sum, Kennedy & Levin (2008) predict only two types of DA scales, shown in 4: A fully closed one with a maximum standard, resulting in a telic reading; and a lower closed one with a minimum standard, resulting in an atelic reading.

4.5.1 Kriz (2011) and the availability of positive endstate readings

Kriz (2011) judges Kennedy & Levin (2008) as generally accurate, but criticises their apparent inability to explain the behaviour of lower closed scale adjectives.

¹⁸ $init(e)$ and $fin(e)$ return the initial and final time intervals of the DA event, parallel to Hay et al.'s SPO and EPO.

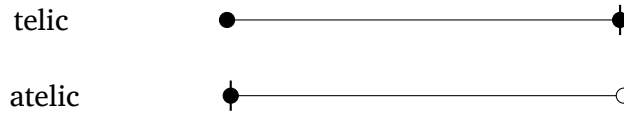


Figure 4: DA scale types in Kennedy & Levin (2008)

As has been established in section 2.2, these should yield DAs with positive endstate readings. Kennedy & Levin do not discuss such predicates, but if their analysis in the given state were applied to them, Kriz argues, it would derive a comparative reading. This is because Kriz continues the tradition of conflating the notions of telicity and DA reading. I argue that if we treat the positive endstate reading as an independent property, detached from being telic, we can in fact arrive at the correct predictions based on Kennedy & Levin's measure of change.

Under the assumption that all positive endstate DAs are telic and vice versa, Kennedy & Levin's analysis does indeed seem problematic. To have a positive endstate reading, a DA must be telic; to be telic, it must have a MAX standard; to have MAX standard it must have a maximum; and for a DA to have a maximum, the underlying adjective must come with a maximum. Since this is not the case for lower closed adjectives, there seems to be no path to a positive endstate reading. To solve this problem, we must renounce the idea that telicity is a prerequisite for a positive endstate reading, and instead focus on the properties of the reading by itself. As outlined in 4.1, the core of a positive endstate reading is the entailment that if the DA holds, the adjective must hold. The absence of such an entailment creates a comparative reading. It is clear what it means for the adjective to hold: Its standard must be exceeded. For Kennedy & Levin, since DAs have their own scales, a DA also holds when its standard is met. We are thus interested in whether the fact that the DA standard is met can tell us anything about the status of adjective standard. So the key to a positive endstate reading in this analysis is the relationship between the DA standard and the underlying adjective standard. We will now explore what generalisations can be made with regards to this relationship.

For absolute adjectives, we have four scale types to look at: lower closed with a MIN standard, upper closed with a MAX standard, fully closed with a MIN standard and fully closed with a MAX standard. The analysis states that to derive

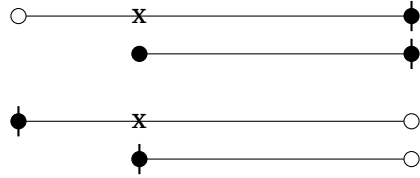


Figure 5: Upper closed and lower closed adjectives scales and their derived DA scales acc. to Kennedy & Levin (2008)

the DA scale, the adjective scale is cut off at the event's initial degree, which becomes the new minimum. Everything below the initial degree is removed, the rest becomes the DA scale. A new standard is determined following Interpretive Economy with the addition that MAX standards are preferred over MIN; this means that DAs have a MAX standard whenever possible, i.e. whenever the underlying adjective scale can provide a maximum.

Upper closed adjective scales will result in a fully closed DA scale. Both have maximum standards, since for the adjective scale the maximum is the only available endpoint and for the DA scale it is preferred over the minimum as standard. Thus, DA and adjective have the same standard. Lower closed scales are not mentioned by Kennedy & Levin, but it can be predicted that in this analysis, they will result in lower closed DA scales; both will have MIN standards, since neither has any other endpoint available, but the standard of the DA scale can lie above the adjective standard if the argument starts out with an initial degree that is higher than the minimum. (Both of these processes are illustrated in figure 5; x marks the initial degree on the adjective scale, which translates into the lower end of the DA scale.) For fully closed scales (fig. 6), the DA will also have a fully closed scale, where the standard is always the maximum. So the standard is either the same as the adjective's, if the underlying standard was a MAX too, or raised from the MIN to the MAX.

So, crucially, for all absolute adjectives, the derived DA standard will be either equal to or above, but never below the original standard. This is why such DAs have positive endstate readings: Any degree that exceeds the DA standard must necessarily also exceed the adjective standard, since the adjective standard can never be higher than the DA standard. Thus, in Kennedy & Levin's analysis, the truth of the DA entails the truth of the positive adjective for all DAs based on absolute adjectives, just as expected.

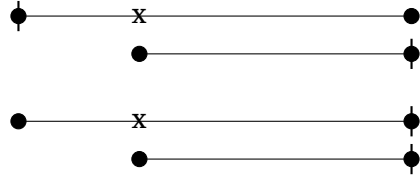


Figure 6: Fully closed scales and their derived DAs acc. to Kennedy & Levin (2008)

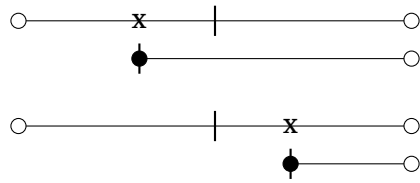


Figure 7: Relative scales and their DAs according to Kennedy & Levin (2008)

For relative adjectives, as shown in figure 7, this is not the case. Since relative adjectives do not come with a maximum, the DA standard will always be the new minimum, i.e. the initial degree. This can be equal to or above the adjective standard, but, crucially, could also be below it. This means the DA can be true without the adjective being true, and we get a comparative reading.

With this understanding of the formal basis of the DA readings, it becomes clear that Kriz's critique of the Kennedy & Levin analysis is inaccurate. While he is correct in remarking that the analysis does not explicitly discuss the behaviour of lower closed scales, and indeed does not provide a way of deriving a telic reading for the corresponding DAs, the presence of a positive endstate reading is in fact neatly predicted for them, if we are willing to review our notion of DA readings and conceptually separate them from telicity.

If telicity and DA readings are separate concepts, we must wonder what the relationship is between them. Being telic means to have only one possible end state. This means the DA standard must be the maximum; if it's any lower, there will be points above the standard, giving us multiple possible endstates. So a telic DA is any DA with a maximum standard. An atelic DA is one with a minimum standard. Having a positive endstate reading means the DA implies the truth of the adjective, i.e. having a standard that is above or equal to that of the adjective. Clearly this is true for all MAX standard DAs: No matter the

underlying adjective, if the DA standard is the maximum of the scale, it can not possibly be below the adjective standard. So every telic DA must also have a positive endstate reading. It is however not the case that every positive endstate DA must also be telic, since there are ways for the DA to entail the truth of the adjective without having a maximum standard; namely, when both the adjective and the DA have a minimum standard, as is the case for lower closed adjectives. Thus, we predict that there are DAs that are telic and have a positive endstate reading, DAs that are atelic and have a comparative reading, and DAs that are atelic and have a positive endstate reading; but no DAs that are telic and have a comparative reading. The empirical validity of this will be verified in section 5.2.

4.6 McNally (2017)

McNally (2017) takes a closer look at how the standards of degree achievements are fixed and in what way they are influenced by the underlying adjective. She disagrees with Kennedy's claim that absolute standards must always be endpoints and midpoint standards are context-dependent. Instead, she says, what makes a standard absolute is the readiness with which it can be determined, no matter at which point of the scale. She recognises the need for separating the concept of telicity from the DA reading, but, in contrast with my own ideas, comes to the conclusion that Kennedy & Levin's analysis is untenable in its given form.

The starting point of McNally's analysis is the distinction of **rule-based** and **similarity-based classification**, which goes back to Hahn & Chater (1998) and which she originally adapted for adjective scales in McNally (2011). She assumes that the decision of whether a DA holds is made via rule-based classification, which entails comparison with an abstract prototypical entity, rather than similarity-based classification, where a comparison class is partitioned based on between item comparison (McNally 2017:188,89). A predicate is associated with an abstract set of satisfaction conditions, which need to be fulfilled in order for the predicate to hold. To establish these conditions, it needs to be determined what degree on the adjective scale must be reached for the DA to apply. If this information is easily available, for example if the degree in question is an endpoint, or if it is salient in the context, a telic reading is possible. If it is hard to determine, like for relative adjectives in uninformative contexts, an atelic read-

ing arises.

McNally is motivated in her analysis by a set of examples that she takes as evidence that DA telicity is not a direct consequence of the adjective's scale structure. First, she cites data from Kearns (2007) and Fleischauer (2013) that are supposed to show the existence of telic readings with a nonmaximal DA standard. We have seen above in 4.5.1 that there should be DAs with minimum standards that have positive endstate readings (namely those based on lower closed adjectives), but they are not predicted to be telic, so this is unexpected.

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

(Kearns 2007:22)

(65) Der Zustand hat sich stabilisiert, er ist aber nicht vollkommen
the condition has itself stabilised he is but not completely
stabil.
stable

'The condition has stabilised, but it is not completely stable.'

(Fleischauer 2013)

I agree with Kennedy & Levin (2008:9) that example (64) is a clear case of an argument induced scale: What is being modified by *completely* here is not the darkness of the sky, but the amount of sky that the darkening has affected; so the first and second clause in (64) are statements about two different scales, which prevents the denied completeness from clashing with the DA's maximum standard. It is hard to find an example where this reparation can be ruled out in order to show the unacceptability that should arise; I can only think of (66), though I am not completely sure about the properties of *really* with gradable adjectives.

(66) The sky darkened in an hour, #but it wasn't really dark.

In any case, I don't agree that (64) can be interpreted as proof that *darken* can occur with a nonmaximal standard.

Fleischauer's example in (65) cannot be discounted as easily; there is clearly no argument induced scale at work here. However, it can be explained as a case of imprecision, in which case it would be sufficient to reach a degree near the standard for the positive of the adjective to hold. According to Kennedy &

McNally (2005:359), this is a common property of maximum-standard adjectives. So the positive of the adjective may still be implied by the DA; but it is satisfied by near-total stability and is thus compatible with the second clause. Example (65) might be a valid argument for nonmaximal standards in telic DAs, but only if one rejects the idea of imprecision, which McNally does not argue against.

McNally also recalls Dowty's soup cooling example, repeated in (67-a), and more examples like it where a maximum for the DA seems to be inferred from the context ((67-b) and (67-c)).

- (67) a. The soup cooled in 10 minutes.
 b. The gelatin will finish hardening around the fruit and candy, creating a solid cell model. (McNally 2017:184)
 c. She finished sharpening her pencil before she answered him. (McNally 2017:184)

She claims that examples like these are too ubiquitous to be analysed as conventionalised irregular standards (McNally 2017:185). This claim is so broad it is hard to argue for or against. Even if the standards in these cases really couldn't be conventionalised, I do not see why it should not be possible for them to be inferred from the context for every utterance individually.

Another important motivation for her questioning Kennedy & colleagues is the behaviour of DAs derived from adjectives with minimum standards (McNally 2017:185). She uses *awaken* as an example, observing that this verb is reliably telic, as shown in (68), and receives a positive endstate reading, as shown in (69).

- (68) (McNally 2017:186)
 a. The baby is awakening. $\models$ The baby has awoken.
 b. The baby finished awakening.
 (69) The baby awoke.
 a. $\approx$ The baby became awake.
 b. $\not\approx$ The baby became more awake.

According to McNally, Kennedy & Levin would predict DAs like this one to al-

low a comparative atelic reading, since it has a minimum standard. I am confused by the choice of this example, since the adjective *awake* has a maximum (“completely awake”; cf. Sudo 2015:31, Kennedy & McNally 2005:361), which McNally (2017:195) herself is aware of. Cases like these are in fact covered by Kennedy & Levin via the informativity principle that prefers telic over atelic readings: As long as a maximum is available, the corresponding DA will always be telic and have a positive endstate reading. So this is not a gap in their analysis; the actual question should be about adjectives with a minimum standard that do not have a maximum, which we already discussed in 4.5.1. McNally ignores the existence of the informativity principle until the very last page of her paper, where she mentions it briefly. From what I understand, she dislikes it because as the result of a pragmatic principle, such positive endstate readings should be cancellable, which she claims they are not. While this may be a valid point to make, it deserves a more elaborate illustration than is given by McNally, especially since without it, there is no problem with (69) and McNally’s criticisms are no longer motivated. The invalidity of the informativity principle as a basis for the positive endstate reading will have to be reliably established before Kennedy & Levin’s proposal is rejected entirely.

To conclude, while McNally’s analysis in itself is not invalid, I do not agree that it is actually necessary to divert from Kennedy & Levin (2008) in such a substantial way.

4.7 Conclusion

I accept Kennedy & Levin (2008) as a largely accurate analysis, with the caveat of separating positive endstate and comparative readings from telicity and atelicity, as explained in 4.5.1. With this slight modification and our new understanding of what defines a positive endstate reading, I was able to show that the correct predictions about the readings and about lower closed adjective DAs actually fall out of the Kennedy & Levin analysis. For this reason and the arguments detailed in 4.6, I do not believe a drastic reanalysis of Kennedy & Levin’s fundamental ideas is necessary. We will now have a look at how this theoretical picture holds up in practise.

5 Empirical perspective

To consolidate the results of the above discussion, I now want to look at a number of natural language examples. I will examine multiple examples representing each adjectival scale type and their DAs with regards to all relevant properties, using the established diagnostic tools and discussing every inconsistency we may encounter. I think a detailed and thorough approach like this is essential to gain a firm grasp on the topic. For a truly reliable picture of the empirical facts, an extensive study with a varied set of adjectives and a large number of judgements would be needed; but this section can nonetheless give a first indication of which aspects of our present system work and which still need more attention.

5.1 Adjectival properties

We will start with a set of adjectives covering the typology of scale types we established in figure 2. We will use the diagnostics introduced in 2.3.1, 2.3.2, 3.2 and 3.3. The full inventory is given in 8, and each adjective's characteristics are discussed in detail in the following.

I will begin with the relative adjectives, since their behaviour is, for the most part, unproblematic. The relative adjectives I will use are *wide*, *weak*, *short*, *cheap* and *narrow*. I will not make the effort to show their gradability, since I do not expect the reader to doubt it; if necessary, it could easily be proven with any of the tests we saw in section 2.1. We expect all of these adjectives to have open scales; and they do, as can be seen from their ungrammaticality with the modifiers *completely*, *slightly* and *half* (cf. section 2.3.1).

- (70) a. I think the truck will fit! The road is #completely wide.
b. The gap is #slightly wide.
c. The bicycle lane should be #half wide.
- (71) a. My influence on his actions is #completely weak.
b. The effects of that drug are #slightly weak.
c. My headache is not too bad; it's #half weak.
- (72) a. The match I pulled was #completely short.
b. Her dress is #slightly short.

- c. Cut the string to be #half short.
- (73) a. Beer in Germany is #completely cheap.
b. The pizza there is great and #slightly cheap.
c. The boots I want are #half cheap.
- (74) a. The pipe must be #completely narrow.
b. The staircase is #slightly narrow.
c. German pavements are often #half narrow.

If Kennedy & McNally (2005:355) are correct, these adjectives should have relative standards, since there are no endpoints available. This is also the case. We can show that the standard is relative by applying the first diagnostic we saw in 2.3.2: A comparative expression ‘*x* is more *A* than *y*’ should imply that *y* is not *A* for maximum standards, that *x* is *A* for minimum standards, and neither for relative standards. (75) – (79) show that for all the adjectives, no entailments of this kind exist, meaning they do not have endpoint standards.

- (75) a. The new road is wider than the old one.
b. $\nVdash$ The old road is not wide.
c. $\nVdash$ The new road is wide.
- (76) a. Tim is weaker than Paul.
b. $\nVdash$ Paul is not weak.
c. $\nVdash$ Tim is weak.
- (77) a. The red rope is shorter than the green rope.
b. $\nVdash$ The green rope is not short.
c. $\nVdash$ The red rope is short.
- (78) a. Blueberries are cheaper than strawberries.
b. $\nVdash$ Strawberries are not cheap.
c. $\nVdash$ Blueberries are cheap.
- (79) a. The cycle path is narrower than the sidewalk.
b. $\nVdash$ The sidewalk is not narrow.
c. $\nVdash$ The cycle path is narrow.

We can thus conclude that relative adjectives behave as expected.

I will now move on to absolute adjectives, which are less straightforward in their behaviour. First, let us look at *clean*, *dry*, *flat* and *clear*, which have upper closed scales. If these scales have maxima but no minima, they should be compatible with *completely*, but not with *slightly* or *half*. This is illustrated in (80).

- (80) a. I made sure the kitchen is completely clean.
b. Use that table over there, at least it's #slightly clean.
c. The bathroom is ?half clean.
- (81) a. The twigs must be completely dry to burn well.
b. My bathing suit is #slightly dry.
c. His shirt is ?half dry.
- (82) a. The surface is completely flat.
b. The curve is #slightly flat.
c. The ground is #half flat.
- (83) a. The spring water is completely clear.
b. This plastic sheet is #slightly clear.
c. The matted glass is #half clear.

In the majority of cases, the adjectives behave exactly as predicted: They are acceptable only with *completely*. It seems that “half clean” and “half dry” are not as bad as expected. Sentences like (84) are easily acceptable.

- (84) a. You're not done with the kitchen yet! It's only half clean.
b. Don't take down the towels yet; they're only half dry.

I think that in these cases, the adjective is actually interpreted with a process reading; *clean* no longer measures the level of cleanliness, but the point to which a cleaning process has progressed. It should thus be assumed to involve a scale like the ones proposed for DAs in Kennedy & Levin (2008), spanning from the initial degree of cleanliness to the maximum of the cleanliness scale. Since this scale is fully closed, it is not surprising that *half* becomes compatible. This idea is supported by the fact that the reading goes away when the argument is not

something that usually changes its state, e.g. (85).

(85) The cake is #half dry.

The fact that a closed DA scale is only possible for an adjective with a maximum would imply that such process readings are unavailable for adjectives with lower closed or relative scales. I can think of no counterexamples to that generalisation at this point, so I accept that the DA-like process reading is in fact what we are dealing with here. The examples in (84) thus do not mean that *clean* and *dry* in their default interpretations have fully closed scales.

Since the maximum is the only endpoint of these adjectives, following Kennedy & McNally (2005) we would expect them to have maximum standards. This means that for comparative constructions '*x* is more *A* than *y*', they should entail that *y* is not *A*, since any degree for which there exists a higher degree on a maximum standard scale must be below the standard. We can see in (86) that this bears out for all of our examples.

- (86)
- a. The floor is cleaner than the table. $\models$ The table is not clean.
 - b. My socks are dryer than my shoes. $\models$ My shoes are not dry.
 - c. The landscape in the north is flatter than the landscape in the south.
 $\models$ The landscape in the south is not flat.
 - d. The water in the pool is clearer than that in the birdbath. $\models$ The water in the birdbath is not clear.

Next, I will look at some adjectives with fully closed scales. I will discuss *open* and *awake*, which are generally accepted as having minimum standards, and *empty* and *straight*, which are assumed to have maximum standards. First, we can verify the presence of both endpoints in all four adjectives by checking the acceptability with all three of our degree modifiers.

- (87)
- a. I could see everything; the door was completely open.
 - b. Leave the door slightly open so I can hear!
 - c. The chest was half open.
- (88)
- a. To drive a car you should be completely awake.
 - b. The baby was #slightly awake, babbling.
 - c. I was half awake when you called me.

- (89) a. The classroom is completely empty.
- b. Let me top you off! Your glass is #slightly empty.
- c. Your cup is half empty.
- (90) a. The border between Colorado and Wyoming is completely straight.
- b. The road is #slightly straight.
- c. My back is half #half straight.

This is where things start to get problematic. *Open* is the only one of the four adjectives which behaves as expected. *Awake*, *empty* and *straight* are all unacceptable with *slightly*, contra Kennedy & McNally's predictions, even though they all supposedly have minima (cf. Sudo 2015:31, Kennedy & McNally 2005:361 for *awake* and *straight*, Sudo 2015:29, Demonte 2011:1331, Lassiter 2015:159 for *empty*; additionally, the grammaticality with *half* implies the existence of a minimum for *empty* and *awake*)¹⁹. *Straight* does not even allow modification by *half*²⁰. Nouwen (2011) observes that this is a frequent phenomenon with fully closed scales. I can think of no interfering pragmatic effect or different reading that might be responsible for this behaviour. If there are none, we must either take this as evidence that the adjectives in question are in fact open at the lower end or revise our understanding of the semantics of *slightly*, a path which has been followed e.g. by Sassoon & Zevakhina (2012).

A strong argument in favour of the minima in these adjectives can be made by showing that their antonyms have maxima (cf. section 2.2): *completely* is acceptable with *asleep*, *full* and even *bent* if the right context is given.

- (91) a. He actually got up and walked around the house while being completely asleep!
- b. The theatre was completely full.
- c. Danny stops, lifts the spoon. It's completely bent.

(COCA: Ocean's Twelve, 2004)

In addition, accepting the fact that they are lower open would leave us with

¹⁹These judgements are corroborated by COCA, which shows no entries for "slightly awake" and "slightly straight", and only two metaphorical uses of "slightly empty".

²⁰Replacing *half* with *partially*, which is sometimes used as a similar diagnostic, significantly improves (90-c), but only because then, a reading with an argument induced scale becomes available.

no explanation for their grammaticality with *half*. In light of these facts, I will maintain the perspective that *awake*, *empty* and *straight* have fully closed scales. Their unacceptability with *slightly* must be attributed to some independent factor.

With regards to standard orientation, on the other hand, everything works as expected. In the comparison test, all of our adjectives give the desired entailments, supporting the idea that *open* and *awake* have minimum standards, while *empty* and *straight* have maximum standards.

- (92) a. The left window is more open than the right window.
 b. $\nVdash$ The right window is not open.
 c. $\models$ The left window is open.
- (93) a. He is more awake now than he was five minutes ago.
 b. $\nVdash$ He was not awake five minutes ago.
 c. $\models$ He is awake now.
- (94) a. The salsa bowl is emptier than the guacamole.
 b. $\models$ The guacamole bowl is not empty.
 c. $\nVdash$ The salsa bowl is empty.
- (95) a. The dirt path is straighter than the the paved road.
 b. $\models$ The paved road is not straight.
 c. $\nVdash$ The dirt path is straight.

We can thus move on to the final adjective type, which is the lower closed scale. We will use *dirty*, *diverse*, *rough* and *sick*. These adjectives should allow modification by *slightly*, since they have minima, but not *completely* or *half*, since they lack maxima. This largely bears out.

- (96) a. Don't walk barefoot on the street, it's #completely dirty.
 b. Can I get a new spoon? This one is slightly dirty.
 c. I should wash my hands, they are #half dirty.
- (97) a. The cast of this show is #completely diverse.
 b. At least his ideas are #slightly diverse.
 c. Your sources of income are only #half diverse.
- (98) a. The raw wood is #completely rough.

- b. The sole of a pointe shoe is slightly rough to ensure a good grip on the stage floor.
 - c. The fabric in this sweater is #half rough.
- (99)
- a. It was the worst cold of my life; I was #completely sick.
 - b. After dinner, I was feeling slightly sick.
 - c. I am getting better but I'm still #half sick today.

“Completely dirty” can be acceptable in cases where an argument induced reading is possible, e.g. (100), so one must take care to avoid incremental arguments when trying to observe the unacceptability of *completely* with the adjective scale. I hope that the example in (96-a) does a good enough job of this.

- (100) His pants were completely dirty.

Since these adjectives have lower closed scales, they would all be expected to have minimum standards, which is in fact the case.

- (101)
- a. His hands are dirtier than his face.
 - b. $\nVdash$ His face is not dirty.
 - c. $\models$ His hands are dirty.
- (102)
- a. The cast of *The Awakening* is more diverse than the cast of *Brothers*.
 - b. $\nVdash$ The cast of *Brothers* is not diverse.
 - c. $\models$ The cast of *The Awakening* is diverse.
- (103)
- a. Your hands are rougher than mine.
 - b. $\nVdash$ My hands are not rough.
 - c. $\models$ Your hands are rough.
- (104)
- a. Patty is sicker than her brother.
 - b. $\nVdash$ Patty's brother is not sick.
 - c. $\models$ Patty is sick.

Thus, we end up with the inventory of gradable adjectives shown in figure 8. We will examine the corresponding DAs in section 5.2.

	upper closed	upper open
lower closed	open, awake, empty, straight	dirty, diverse, rough, sick
lower open	clean, dry, flat, clear	wide, weak, short, cheap, narrow

Figure 8: Adjective Inventory

5.1.1 A note on colour adjectives

Some adjectives are not uncontroversially classifiable with regards to gradability. A notable point of discussion are colour adjectives. Various authors treat them as gradable (e.g. Hansen & Chemla 2017, McNally 2011), and indeed they do appear with many of the above diagnostics for gradability (s. Demonte 2011:1333).

- (105)
- a. The grass is always greener on the other side.
 - b. These are the whitest teeth I have ever seen; they're blinding.
 - c. I don't like this lipstick, it's too pink.
 - d. His nose was just red enough to compromise the appearance of seriousness. (COCA: Cold vengeance by Douglas Child, 2011)
 - e. The sky is very blue today.
 - f. How red are my eyes? Can you tell I've been crying?

There is also a very obvious real world scale associated with colours in the spectrum of visible light. However, this scale differs from classic adjectival scales in an important respect: It is not obvious which way on the colour scale is up and which is down. What the above examples seem to express are various modifications of the intensity of a colour; its closeness to the speaker's representation of the prototypical shade referred to by the adjective, the so-called **focal colour** (Berlin & Kay 1969), e.g. a prototypical red, yellow, etc. Let's say the scale associated with the adjective *green* is a section of the visible spectrum, with focal green as its standard. There is no one-to-one correspondence between a degree-modified expression like "not green enough" and a point on this scale – the corresponding shade could lie either on the yellow or on the blue side of

the standard. The usual meaning of *not...enough* as ‘too low on the scale’ can suddenly send us both ways. This might be fixed by proposing that the scale associated with colour adjectives is in fact not a section of the visible spectrum, but a derived scale measuring the negative distance from a focal scale on the visible spectrum, in any direction, with the focal colour as the maximum. Interpreted in this way, colour adjectives start to resemble the examples of coercion with nationality adjectives we saw in (11): What is being graded is the level of similarity to a prototypical representation of an entity with the given property, not the degree to which the property itself applies. If this is the right direction of analysis, colour adjectives could in fact be reclassified as ungradable. Demonte (2011:1333-34) takes a similar perspective.

Whether the focal colour is a fixed shade or derived from a context-dependent comparison class I cannot conclusively determine here. While there are data suggesting the universality of focal colours across cultures and languages (e.g. Regier et al. 2005, Abbott et al. 2016), they all seem to require varying levels of flexibility. This could be accounted for with full blown context-dependency or the assumption of an absolute scale in combination with limited granularity and/or imprecision. Because of their atypical behaviour, I have refrained from using colour adjectives as examples here.

5.2 DA properties

We will now take a closer look at the behaviour of the DAs derived from the adjectives above, starting with the relative adjectives. Following Kennedy & Levin (2008), we would expect their DAs to have no maxima, and thus a minimum standard, meaning they should be atelic. The *in/for*-test does not reflect this as cleanly as one would hope: As can be seen in (106), only *to widen* and *to narrow* show clear atelic behaviour, while *to weaken* and *to cheapen* seem to allow both *in*- and *for*-duratives.

- (106)
- a. The gap widened #*in/for* five minutes.
 - b. The pain weakened *in/for* a day.
 - c. He shortened his speech *in/for* half an hour.
 - d. Meat products cheapened *in/?for* six months (until there was a shortage).

- e. The road narrowed #in/for half a mile.

We are thus forced to consider whether these predicates have variable telicity. I do not think this is the case, for the following reasons. First, upon closer inspection it becomes apparent that the telic readings in (106-b) and (106-d) are the result of additive coercion, meaning a process that is not part of the predicate denotation is added to the reading to force durationality (e.g. Rothstein 2004). “The pain weakened in a day” is read as ‘After a day’s wait, the pain weakened’ instead of ‘The pain weakening process lasted a day until it was finally complete’. A similar reading arises for “Meat products cheapened in six months”. The fact that this is necessary indicates that the bare uncoerced predicate is not telic itself. Secondly, I would argue that the telic reading in (106-c) is produced from a contextual standard, like in the soup cooling examples discussed in 4.3: Shortening a speech almost always involves reaching a set maximal length. Again, this does not mean the predicate *to shorten* is inherently telic.

Finally, other telicity tests show consistent atelic behaviour for all of the above DAs. Dowty’s *almost*-test works very well: None of the adjectives give rise to an ambiguity between the *almost-started*- and the *almost-finished*-reading; all clearly only allow the *almost-started*-reading, as is expected for atelic verbs.

- (107) a. The gap almost widened.
- b. The pain almost weakened.
- c. He almost shortened his speech.
- d. Meat products almost cheapened.
- e. His eyes almost narrowed.

In addition, none of the predicates is a possible complement of *to finish*. (108-c) may be acceptable, but again only with the support of a contextual standard.

- (108) a. #The gap finished widening.
- b. #The pain finished weakening.
- c. ?He finished shortening his speech.
- d. #Meat products finished cheapening.
- e. #His eyes finished narrowing.

I conclude that these predicates are in fact all atelic, and that the *in/for*-test is failing here for independent reasons. We will return to this problem below, in

the discussion on DAs based on lower closed adjectives, which show the same behaviour. For now, we can move on to the DA reading.

According to my understanding of positive endstate entailments, DAs based on relative adjectives should receive a comparative reading, meaning the DA can be true without the adjective being true. (109) shows that this is the case for all five DAs, confirming my expectations.

- (109)
- a. The gap has widened but it isn't wide.
 - b. The pain has weakened but it is still strong.
 - c. He has shortened his speech but it's still very long.
 - d. Meat products have cheapened but they aren't cheap.
 - e. The road has narrowed but it's still pretty wide.

Next, we will look at DAs derived from our upper closed adjectives: *to clean*, *to dry*, *to flatten* and *to clear*. We would expect all of these to have maxima and thus maximum standards, making them telic. They should also receive a positive endstate reading, since they are derived from absolute adjectives, meaning their standard necessarily lies on or above the adjective standard.

- (110)
- a. He cleaned the kitchen in/?for an hour.
 - b. My bathing suit dried in/#for twenty minutes.
 - c. The landscape in Central Europe used to be mountainous, but it flattened in/#for a hundred million years.
 - d. The sky cleared in/#for an hour.

(110) shows that these DAs are in fact telic. There seems to be a possible atelic interpretation of *to clean*, which in my understanding results from reinterpretation as 'to clean at' or 'to try to clean'. Other than that, the predicates behave as expected. The *almost*-test also paints a clearly favourable picture; all predicates are ambiguous between an 'almost-started'- and an 'almost-finished'-reading.

- (111)
- a. He almost cleaned the kitchen.
 - b. My bathing suit almost dried.
 - c. Erosion almost flattened the landscape.
 - d. The sky almost cleared.

It should be noted here that for (111-b), (111-c) and (111-d), the *almost-finished*-

reading is much more easily available than the *almost-started*-reading; I expect this is because it is hard to imagine what it means for an action to almost begin if it's not being planned by an agent. In any case, this does not diminish the significance of the test, since it is the availability of the *almost-finished*-reading that distinguishes telic from atelic predicates.

(112) shows that all our examples also have a positive endstate reading, just as expected.

- (112) a. He has cleaned the kitchen #but it's still dirty.
 b. My bathing suit has dried #but it isn't dry.
 c. Erosion has flattened the landscape ?but it isn't flat.
 d. The sky has cleared #but it isn't clear.

It appears that (112-c) is not as clearly unacceptable as the other examples. This may be because of imprecision, *flat* being interpreted as 'completely flat'; in the context of dimensions as large as landscape configurations, it seems especially likely that some imprecision may be admissible.

For DAs based on fully closed scales, we make the same predictions as for upper closed scales, since they too come with maxima, allowing for a telic and positive endstate reading. This time, the *in/for*-test works very well:

- (113) a. The garage door opened in/#for half a minute.
 b. He awoke in/#for a couple of minutes.
 c. She emptied her pint in/#for three seconds.
 d. The builders straightened the road in/#for two weeks.

The *almost*-test corroborates these results.

- (114) a. The garage door almost opened.
 b. He almost awoke.
 c. She almost emptied her pint.
 d. The builders almost straightened the road.

The predicates also clearly show positive endstate behaviour.

- (115) a. The garage door has opened #but it isn't open.
 b. He has awakened #but he isn't awake.

- c. She has emptied her pint #but it isn't empty.
- d. They have straightened the road #but it isn't straight.

Finally, we can look at DAs based on lower closed adjectives. We expect these to have no maximum, and thus be assigned a minimum standard, making them atelic. Interestingly, however, as established in the previous section, they should still have a positive endstate reading, since they are based on absolute adjectives. The results of the *in/for*-test are problematic, similar to those for DAs based on relative adjectives (s. (106)).

- (116) a. He dirtied his pants ?in/?for five minutes.
- b. They diversified the cast ?in/#for one year.
- c. He roughened the soil with a rake #in/?for ten minutes.
- d. He sickened ?in/#for a week.

Again, the *almost*- and *finish*-tests show more consistent results: With *almost*, the predicates show no ambiguity (s. (117)), and none of them are compatible with *to finish* (except for *roughen* with a contextual standard (118-c)). Thus, overall it is warranted to stick to the assumption that these DAs are atelic, especially considering the fact that this is not the only DA type for which the *in/for*-test fails.

- (117) a. He almost dirtied his pants.
 - b. They almost diversified the cast.
 - c. He almost roughened the soil.
 - d. He almost sickened.
- (118) a. #He finished dirtying the carpet.
 - b. #They finished diversifying the cast.
 - c. ?He finished roughening the soil.
 - d. #He finished sickening.

We again encounter the problem of a missing agent with the *almost-started*-reading in (117-d), but as established above, this is not detrimental to the argumentation.

Before we move on to the final discussion, we can quickly verify that in spite of their atelicity, these predicates receive a positive endstate reading.

- (119) a. He has dirtied the carpet #but it isn't dirty.
 b. They have diversified their cast #but it isn't diverse.
 c. He has roughened the soil with a rake #but it isn't rough.
 d. He has sickened #but he isn't sick.

5.3 Discussion

While the testing was not without issues, in general it seems the predictions made by Kennedy & Levin (2008) and my own understanding of DA readings do in fact conform with reality. We have seen that when applying the tests we have been using, one must always keep an eye out for various interference effects, such as argument induced scales, contextual standards, additive coercion, pragmatic inconsistencies and the interactions of tense and aspect with telicity. Wherever possible, a variety of tests should be applied.

We have observed two major inconsistencies. The first is the unacceptability of *slightly* with the fully closed adjectives *awake*, *empty* and *straight*. As we have shown, it is still warranted to assume these scales are lower closed. The incompatibility with *slightly* has been observed before and will need to be explained independently.

The second problem is the inadequacy of the *in/for* test with DAs without maxima. I have no final explanation for why this happens. Conceptually, it should be possible to combine e.g. *to dirty* with a durative adverbial expression: It is perfectly realistic that a dirtying process goes on for some time, continuously increasing the degree of dirtiness of the argument, without an intrinsic endpoint. It seems that these predicates receive a sort of punctual interpretation due to their minimum standard: Even though the action can be durative, it counts as complete as soon as it has been started. As a result, the *in/for*-test is no longer applicable, since it presupposes durationality. This will have to be verified and formalized in future work.

I want to note that I found my intuitions on *to flatten*, *to sicken* and *to roughen* to be way less secure than for the other examples; I have done my best to correctly represent their behaviour, but am very much open to conflicting opinions. Here especially the need for a study of statistically significant extent becomes apparent.

6 Conclusion

The semantic structure of Degree Achievements is a topic under construction. While much progress has been made, there are still several aspects that require improvement.

As we saw in section 4, our understanding of DA semantics has converged on a unified comparative semantics, deriving one default reading for each DA instead of assuming pervasive ambiguity between positive endstate and comparative readings. Kennedy & Levin (2008) is currently the most extensive and accurate account of the behaviour of deadjectival DAs. The authors correctly predict all adjectives with maxima to yield telic DAs with positive endstate readings and all relative adjectives to yield atelic DAs with comparative readings. In 4.5.1, I discussed Kriz's (2011) complaint that Kennedy & Levin (2008) are unable to correctly derive the behaviour of DAs derived from lower closed adjectives, which should yield a positive endstate reading in spite of their atelicity. I showed that this is a result of the fallacy of equating positive endstate with telic readings, and comparative with atelic readings. I proposed that the DA reading should instead be conceived of as a property of the relationship between DA and adjective scale, independent from telicity, which is a property of the DA scale alone.

Section 5 was dedicated to showing that the predictions made by Kennedy & Levin (2008) in combination with my understanding of the DA reading bear out in natural language. We observed two major inconsistencies with the diagnostics established in 2.3, 3.2 and 3.3, the first being the unacceptability of *slightly* with a number of fully closed adjectives, and the second the inconsistent results of the *in/for*-test with upper open DAs. These phenomena deserve more attention in future work. Fortunately, alternative tests show that the adjectives and DAs in question have the expected properties nonetheless.

The next step towards a complete understanding of the semantics of deadjectival DAs must be more extensive empirical work. A significant number of judgements about the DA readings and underlying adjective properties is necessary to fortify the generalisations that have been made. As far as I am aware, no such work exists at present. Cross-linguistic studies should also be conducted.

7 Bibliography

- Abbott, Joshua T., Thomas L. Griffiths & Terry Regier. 2016. Focal colors across languages are representative members of color categories. *Proceedings of the National Academy of Sciences of the United States of America* 113(40). 11178–11183.
- Abusch, Dorit. 1986. Verbs of change, causation and time. Technical Report CSLI-86-50. Stanford, California: Center for the Study of Language and Information, Stanford University.
- Berlin, Brent & Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley, California: University of California Press.
- Davies, Mark. 2008. The Corpus of Contemporary American English (COCA). www.english-corpora.org/coca/. Accessed between May and August 2024.
- Demonte, Violeta. 2011. Adjectives. In Claudia Maienborn, Klaus van Heusinger & Paul Portner (eds.), *Semantics*, chap. 52, 1314–1340. Berlin, New York: De Gruyter.
- Dowty, David R. 1979. *Word meaning and Montague Grammar*. Dordrecht: Reidel.
- Fleischauer, Jens. 2013. Interaction of telicity and degree graduation of change of state verbs. In Boban Arsenijević, Berit Gehrke & Rafael Marín (eds.), *Studies in Composition and Decomposition of Event Predicates*, 125–152. Dordrecht: Springer.
- Hahn, Ulrike & Nick Chater. 1998. Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition* 65(2). 197–230.
- Hansen, Nat & Emmanuel Chemla. 2017. Color adjectives, standards, and thresholds: an experimental investigation. *Linguistics and Philosophy* 40. 239–278.
- Hay, Jennifer. 1998. The non-uniformity of Degree Achievements. Presented at the 72nd Annual Meeting of the Linguistic Society of America.
- Hay, Jennifer, Christopher Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in Degree Achievements. In Tanya Matthews & Devon Strolovitch (eds.), *SALT IX*, 127–144. Ithaca, New York: Cornell University.

- Kearns, Kate. 2007. Telic senses of deadjectival verbs. *Lingua* 117. 26–66.
- Kennedy, Christopher. 2007. Vagueness and Grammar: The Semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30. 1–45.
- Kennedy, Christopher & Beth Levin. 2008. Measure of Change: The adjectival core of Degree Achievements. In Christopher Kennedy & Louise McNally (eds.), *Adjectives and Adverbs: Syntax, Semantics and Discourse*, 156–182. Oxford: Oxford University Press.
- Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381.
- Kriz, Manuel. 2011. *Issues in the Semantics of Degree Achievements*. University of Vienna MA thesis.
- Lakoff, George. 1965. On the nature of syntactic irregularity. Indiana University Doctoral Dissertation.
- Lassiter, Daniel. 2015. Adjectival modification and gradation. In Shalom Lappin & Chris Fox (eds.), *The Handbook of Contemporary Semantic Theory*, chap. 5, 143–167. Newark: Wiley & Sons Inc.
- Lehrer, Adrienne. 1985. Markedness and antonymy. *Journal of Linguistics* 21. 397–429.
- McNally, Louise. 2011. The relative role of property type and scale structure in explaining the behaviour 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, 151–168. Berlin Heidelberg: Springer.
- McNally, Louise. 2017. On the scalar properties and telicity of degree achievements. In Fernández-Soriano, Olga & Elena Castroviejo Isabel Pérez-Jiménez (eds.), *Boundaries, Phases and Interfaces: Case Studies in Honor of Violeta Demonte*, 179–198. Amsterdam: John Benjamins Publishing.
- Nouwen, Rick. 2011. The Semantics of Degree. Massachusetts Institute of Technology Lecture Notes.

- Regier, Terry, Paul Kay & Richard S. Cook. 2005. Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America* 102(23). 8386–8391.
- Rothstein, Susan. 2004. *Structuring Events: A Study in the Semantics of Lexical Aspect*. Oxford: Blackwell Publishing.
- Sassoon, Galit W. & Natalia Zevakhina. 2012. Granularity shifting: Experimental evidence from degree modifiers. In *Proceedings of SALT 22*, 226–246.
- Seuren, Pieter A. M. 1978. The structure and selection of positive and negative gradable adjectives. In *Papers from the Parasession on the Lexicon, 14th Regional Meeting of the Chicago Linguistic Society*, 336–346. Chicago: University of Chicago.
- Sudo, Yasutada. 2015. *Advanced Semantic Theory*. University College London Lecture Notes.
- Toledo, Assaf & Galit W. Sassoon. 2011. Absolute vs. relative adjectives – variance within vs. between individuals. In *Proceedings of SALT 21*, 135–154.
- Unger, Peter. 1975. *Ignorance*. Oxford: Clarendon Press.
- Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, New York: Cornell University Press.