The Distribution of English Adjectival Past Participles

Suzanne van Duym
University of Minnesota

1   Introduction

In this paper I examine the distribution of two particular forms of the English adjectival past participle: the regular (non-negated) form, e.g. *flown*, and the form negated with the prefix *un-*, e.g. *unflown*. These are compared in prenominal attributive position within noun phrases, e.g. *the unflown plane*. It is shown that there are at least two, and perhaps three, distinct patterns of distribution between non-negated and negated forms: some past participles occur primarily without negation, a few appear equally often with and without the negating *un-*, and a final group is found most often with *un-*.[1] To explain why these distribution patterns exist, I examine factors of informativeness, situation aspect, incremental theme, semantic features and historical considerations. The most significant factor appears to be informativeness. I conclude that, while none of these factors is solely able to explain the distribution patterns, all at least potentially contribute to a multifaceted explanation that likely also involves aspects that remain unidentified.

I begin by giving some background information necessary to an understanding of the subsequent data and analysis. Next, methodology and data are presented and briefly discussed. I start the analysis by addressing semantic factors that may be relevant. This is followed by treatments of situation aspect and the related concept of incremental theme. Historical information is then considered. Finally, an examination of informativeness completes the analysis section; the conclusion follows this.

2   Background Information

The English past participle (italicized), in addition to being used as a verbal form in the perfect and passive constructions shown in (1) and (2) respectively, occurs as an adjectival form as shown in (3) below:

---

[1]For other studies that address adjectival past participles requiring some kind of additional modifying element to be acceptable, see Ackerman & Goldberg (1996), Grimshaw & Vikner (1993), Wasow (1977) and Lakoff (1970).

(1)   Perfect past participles

      a. The geese have *flown* south.

      b. The chef has *cooked* the onions.

(2)   Passive past participles

      a. The planes are *flown* by experienced pilots.

      b. The onions were *cooked* by the chef.

(3)   Adjectival past participles

      a. The vase looks *broken.*

      b. The *cooked* onions were added to the soup.

As can be seen in the examples above, past participial forms may or may not be identical to the corresponding preterite forms (bold), shown in (4) (Huddleston & Pullum 2005:34).

(4)   Preterite forms

      a. The geese **flew** south.

      b. I **broke** the vase.

      c. The chef **cooked** the onions.

In (1a), (2a), and (3a), the past participles differ from their corresponding preterite forms (shown in (4a) and (4b), respectively), while (1b), (2b), and (3b) show the more common situation of a past participle that is identical to its preterite form, shown in (4c). The past participles of central concern in this paper vary as to whether they share the shape of their preterite form, though a majority do, reflecting a general pattern among English past participles as a whole (Huddleston & Pullum 2005:34). However, all of the past participles studied here are of the adjectival type, as seen in (3). (3a) shows an adjectival past participle (APP) functioning as a predicative complement, while (3b) shows an attributive APP as an internal pre-head modifier to a noun. It is APPs within noun phrases of the type in (3b) that will be examined here.

As is the case with some verbal past participles (as in (1) and (2) above), many APPs that are able to function attributively can also take the negative prefix *un-*. This prefix can have either of two distinct meanings, illustrated below:

(5)   *un-* indicating a reversal

      a. She opened the wine then passed around the *uncorked* bottle.

      b. My *untied* shoelaces were caused by a mischievous child with a penchant for detangling knots.

(6)   *un-* indicating absence of a state or property

      a. The *unforeseen* outcome could not have been predicted by anyone.

      b. An *unnamed* baby will not have a birth certificate yet.

The negated APPs in (5) describe nouns that have undergone a process followed by the reversal of that process; this meaning of *un-* is relatively rare in the data used here. In contrast, the APPs in (6) show the much more common meaning of the prefix; these APPs describe a noun as lacking the property indicated by the corresponding non-negated APPs. Because instances of the prefix as in (5) are so rare in the data and do not seem to be relevant, the distinction between these two senses of *un-* will not be further discussed.

3   Methodology

Having established that APPs in general occur both with the *un-* prefix (negated) and without it (non-negated) in noun phrases, I compared the distribution of the prefixed and unprefixed forms for a wide variety of APPs. The Corpus of Contemporary American English (COCA) was used as the data source and was searched for each form within a phrase containing *a*, *an*, or *the* followed by the APP and then any noun.[2] Example search queries are shown below:

(7)   Sample search queries

      a. Query: a|an|the broken [nn*]

      b. Meaning: *a* or *an* or *the* followed by *broken* followed by any noun

      c. Sample retrieval: a broken leg

      d. Sample retrieval: the broken dishes

      e. Query: a|an|the unbroken [nn*]

      f. Meaning: *a* or *an* or *the* followed by *unbroken* followed by any noun

      g. Sample retrieval: an unbroken promise

      h. Sample retrieval: the unbroken vase

(COCA)

---

[2]The requirement of one of these three determiners before each APP served several purposes: first, it somewhat restricted an extremely large data pool (which might, of course, be viewed as good or bad depending on one's perspective); second, it helped to ensure that the phrases retrieved were, in fact, noun phrases; finally, it reduced (though it did not eliminate) the number of retrieved phrases containing plural nouns, which can be problematic for situation aspect (see section 4.2).

(7a) shows the query for an APP in non-negated form as entered into COCA, and (b) shows this query translated into plain English; (c) and (d) show examples that would be retrieved by this query. (7e) shows the query for an APP in negated form as entered into COCA and (f) shows this query translated into plain English; (g) and (h) show examples that would be retrieved by this query. For each query, the total occurrences for the first 1000 tokens was recorded; in most cases this was equivalent to the total occurrences overall, as very few searches yielded more than 1000 tokens.[3] For each APP, the proportion r of total non-negated to negated forms was calculated by dividing the total number of occurrences of non-negated forms p (for *positive*) by that of negated forms n, i.e., $r = p/n$ (or $r =$ *(occurrences of non-negated forms)/(occurrences of negated forms)*).[4] Thus APPs that occur more often in negated than non-negated form had $r < 1$. Example calculations based on the queries in (7) as well as those for an APP with $r < 1$ are shown in (8):

(8) Sample r calculations

     a. broken: $p^5 = 4080$ $n^6 = 232$ $r = 4080/232 = 17.6$

     b. foreseen: $p = 6$ $n = 104$ $r = 6/104 = .1$

As can be seen above, an APP that occurs more often in non-negated form, such as *broken*, has $r > 1$, while one like *foreseen*, which occurs most often as *unforeseen*, has $r < 1$. An r value of 1 indicates that an APP occurs equally often in negated and non-negated forms.

204 APPs were tested, each in both non-negated and negated form, for a total of 408 search queries. Initially APPs were chosen somewhat randomly as they came to mind or were encountered in various contexts not directly related to this research. However, once a pattern of distribution became apparent, I began to seek out the more elusive APPs that occur predominantly with the *un-* prefix, selecting verbs to test that I thought (generally through my own native speaker intuition) might fulfill this criterion. The full list of verbs tested, along with their p, n, and r values, can be found in the Appendix. The number of APPs with $r < 1$, indicating that they occur more often in negated than non-negated form, was 29, or 14% of those tested. Henceforth these will be referred to as negative APPs. This data can be seen in the first column of the table below, and confirms my initial intuition that such forms are a minority, especially considering that this was not a random sampling but one in which these types of APPs were actively sought. Another 26 APPs, in the second data column below, had r in the 1.0-2.2 range, indicating that their non-negated forms do not strongly predominate over the negated ones. These will be referred to as near-equal APPs.

---

[3]As used here, a token means a specific phrase in which an APP was found; an occurrence is a specific instance of a token being found in the corpus. So, for example, for the APP *unbroken* there are 128 tokens, including *an unbroken string*, *the unbroken chain*, and *the unbroken record* (plus 125 others). For the token *an unbroken string*, there are 6 occurrences, meaning that the phrase is found 6 times in COCA.

[4]All r values were rounded to one decimal place.

[5]The number of occurrences of *broken* yielded by the query shown in (7a).

[6]The number of occurrences of *unbroken* yielded by the query shown in (7b).

It remains unclear whether these are best grouped with the lower or higher r-value APPs, or whether they constitute a distinct distributional group on their own. 2.2 was chosen as the upper bound for this group because it reflects a natural break in the r values of the data. Finally, the vast majority of APPs had r values greater than or equal to 2.6, indicating strong predomination of the non-negated form over the negated form; or r values that could not be calculated due to an n of 0, indicating that the negated form was not found at all in the relevant context in the corpus. Henceforth these APPs with $r \geq 2.6$ will be referred to as positive APPs. The data is summarized in Table 1.[7]

Table 1: APP Distributional Data

	r = 0.0-0.9	r = 1.0-2.2	$r \geq 2.6$ or n = 0 (r undefined)
# of APPs	29	26	149
% of tested APPs	14%	13%	73%

Given this data, the obvious question is, what characteristics of this minority, the negative APPs, cause them to occur predominantly with the negative prefix *un-* as shown in the first column of data?

4 Analysis

4.1 Semantic Fields

Several factors were examined in an attempt to answer the question posed above. The first of these was semantic, as it became apparent very early on in the research that many of the negative APPs seemed to share a small number of semantic fields or conceptual categories. Specifically, words having to do with mental processes appeared to be overrepresented in the negative APP category. Examples of these can be seen below:

(9) Mental process APPs
     unknown, undoubted, unforeseen, unmourned, uncounted, unexamined

Other categories represented by several negative APPs included speech-related activities and acts having to do with physical movement. Examples of these are given below:

(10) Other semantic categories

      a. Speech: unsung, unspoken, unanswered, unasked, unquestioned

      b. Physical: unmanned, unattended, unharmed, unzipped, unmoved

---

[7]Percentages here, and throughout this paper, are rounded to the nearest whole percentage point.

The three categories given in (9) and (10) constituted the most salient semantic fields for negative APPs. However, members of these fields were also found among the positive APPs (though in lesser proportions), as seen below:

(11) Positive APP semantic categories

       a. Mental: excited, predicted, intended, admired

       b. Speech: said, announced, called

       c. Physical: swept, risen, thrown, sent

The fact that these categories are represented in both sets of APPs strongly suggests that these semantic distinctions alone are not sufficient to explain the differences in distribution. In addition, the presence of pairs such as *unforeseen* in (9) and *predicted* in (11a), whose non-negated forms are essentially synonymous, makes it clear that semantic field is not the only factor at work here. However, it is also true that the identified semantic fields are represented to a much greater degree among the negative APPs than the positive ones, as seen in Table 2.[8]

Table 2: Semantic Categories as Represented by APPs

	Mental	Speech	Physical
negative	28%	24%	21%
near-equal	23%	8%	14%
positive	19%	5%	13%

As can be seen above, 28% of negative APPs were classified as belonging to the mental processes semantic category, while only 23% of near-equal and 19% of positive APPs were in this category. Even more contrast can be seen in the speech category: 24% of negative APPs were classified in this category, while only 8% of near-equal and 5% of positive APPs related to speech. Finally, 21% of negative APPs dealt with physical movement, while 14% of near-equal APPs and 13% of positive APPs did so. All three categories are represented most among negative APPs and least among positive APPs, with near-equal APPs falling somewhere in between. This indicates that the identified categories may play a small part in explaining the distributional pattern, if indirectly.

4.2   Situation Aspect

Moving on to a venue in which semantics, syntax, and morphology are known to interact, the next potentially explanatory factor examined was situation aspect, the "semantic domain

---

[8]Note that percentages for any given APP type do not add up to 100% because many APPs fell in none of the identified semantic categories.

of the temporal structure of situations" (Levin 2005, Smith 1997:1). I based my analysis of this primarily on Smith (1997); the basic components of her theory are shown below:

(12) Situation aspect as outlined by Smith (1997:3,20)

       a. State: static, atelic, durative
          know the answer, love Mary

       b. Activity: dynamic, atelic, durative
          laugh, stroll in the park

       c. Semelfactive: dynamic, atelic, instantaneous
          tap, knock

       d. Accomplishment: dynamic, telic, durative
          build a house, walk to school

       e. Achievement: dynamic, telic, instantaneous
          win a race, reach the top

There are five types of situations, shown with example phrases in (a) through (e). Each type is defined by whether it is static or dynamic, telic or atelic, and durative or instantaneous. Though it shows some promise, the viewing of this data through a situation aspect lens is problematic. The main reason for this is that situation aspect, as discussed by Verkuyl (1972), Smith (1997) and others, has to do most prominently with a verb; indeed it has been treated in the past as relating solely to the verb (Smith 1997). As noted earlier, the past participles being examined here are not functioning as verbs but as adjectives. Thus aspect can only be discussed with reference to the verbs from which these APPs are derived, rather than the APPs themselves (Grimshaw & Vikner 1993). In addition, many linguists now accept that aspect cannot be determined based solely on a verb, but rather requires what Smith refers to as a "verb constellation," which includes not only the verb itself but its arguments (Smith 1997:5,17; Verkuyl 1972). As adjectives, APPs do not have the type of arguments found in a verb constellation, nor is there always one clear (generic) set of arguments that the verb from which a APP is derived would have. These factors make talking about situation aspect in relation to APPs messy at best, but they do not render the concept useless.

In order to assess the type of situation aspect that might in theory be associated with each APP, I considered the verb from which it was derived together with the arguments suggested by the noun phrases in which the APP occurred; generally, this meant a subject and a singular object. I avoided using plural objects because they can cause atelicity where it would not otherwise be found as the result of an unspecified number in the object preventing a natural endpoint from arising (cf. *eat an apple* vs. *eat apples*) (Smith 1997:55). Using this method, I attempted to assess whether the verb that each APP was derived from was static, comprising "a single, undifferentiated period" or dynamic, comprising "successive stages"; instantaneous (conceptually, at least), or durative; and telic, having a "change of state which constitutes the outcome, or goal, or the event," or atelic (Smith 1997:19). Very few of the positive or negative APPs were derived from static verbs, with most coming from dynamic verbs instead, as shown below.

(13) Static verb phrases associated with APPs

    a. expect a finding

    b. love a boy

(14) Dynamic verb phrases associated with APPs

    a. examine an assumption

    b. sweep the floor

    c. peel an orange

    d. dig wells

Four of the negative APPs had associated verbs that were positively identified as static; among the positive APPs this number was eight.

A significant proportion (48%) of the negative APPs came from exclusively durative verbs, as opposed to 28% from exclusively instantaneous verbs.[9] Some examples of durative verb phrases associated with these APPs are given in (15).

(15) Durative verb phrases associated with negative APPs

    a. sing a song

    b. eat an apple

    c. ask a question

In contrast, a lower proportion of the positive APPs (40%) were derived from verbs of exclusively durative classification, and this was matched by 40% derived from verbs of exclusively instantaneous classification. Examples of both types are given below.

(16) Verb phrases associated with positive APPs

    a. Durative: use the room

    b. Durative: build a road

    c. Instantaneous: break a vase

    d. Instantaneous: miss the shot

When looking only at verbs whose $n = 0$ (a group included in the positive APPs), the percentage derived from durative verbs is an even lower 26%, compared to a much higher instantaneous verb percentage of 64%.

The telicity of some APPs could not be definitively determined: among the APPs classified, all groups had more telic than atelic associated verbs. Examples of telic and atelic verb phrases are given below.

---

[9]Many APPs, both positive and negative, appeared in multiple senses (whose durativity varied) in the data and as such could not be classified as exclusively durative or instantaneous.

(17) Verb phrases associated with positive APPs

    a. Telic: catch a fish

    b. Atelic: amuse the astronomers

(18) Verb phrases associated with negative APPs

    a. Telic: eat a Twinkie

    b. Atelic: mourn the victims

Among positive APPs, 47% were associated with telic verbs and 27% with atelic verbs; among negative APPs, 62% were associated with telic verbs and 31% with atelic verbs. Due to the fact that significantly more positive than negative APPs were unclassifiable (26% versus 7%), it is difficult to say whether my expectation that negative APPs' associated verbs would have a higher rate of telicity was in fact confirmed. Had every form been classifiable, I may or may not have seen the above-noted 15 percentage point difference between telic verbs associated with positive APPs and those associated with negative APPs. A more robust method of determining situation aspect characteristics of verbs from which APPs are derived is likely needed in order to assess whether these features may play an explanatory role in the current question. This may include treating APPs derived from different senses of the same verb as different APPs, something that has not yet been undertaken in this study.

In summary, positive APPs were most likely to be derived from verbs that were dynamic, durative or instantaneous, and telic; assuming that these characteristics occur together, these are verbs that, by Smith's (1997) classification, would be considered either accomplishments or achievements. Among positive APPs, I classified 47% as either accomplishments or achievements, or possibly varying between the two (depending on the sense of the verb or on uncertain durativity classification). Another 18% had varying or uncertain classification that included accomplishment or achievement and some other classification, e.g. semelfactive or achievement, based on varying or uncertain telicity. When isolating from the group of positive APPs only those with $n = 0$, these percentages remain similar (40% and 24%, respectively).

In contrast, negative APPs were most likely to be derived from verbs that were dynamic, strictly durative, and telic; Smith (1997) would classify a verb having all these characteristics as an accomplishment. Among negative APPs, I classified 21% as accomplishments, and another 17% had varying or uncertain classification that included accomplishment as a possibility. This relatively low total percentage (38%) suggests that the characteristics of dynamic, durative, and telic, while they do predominate in the negative APP category, do not necessarily tend to co-occur. In fact, achievement proved to be the most common classification, with the slightly higher total percentage (strict achievement classifications as well as variable classifications that included achievement) of 41%.

Taken as a whole, this data suggests that situation aspect alone is not very explanatory of the distribution of positive and negative APPs. The single factor most likely to play a role appears to be durativity, especially when one considers the significant difference of this factor

between negative APPs and $n = 0$ APPs. What this data does show is that APPs that can be used in the position studied here, both positive and negative, are often derived from verbs of the accomplishment or achievement type. This is significant largely because the accomplishment type has previously been identified as contributing to modification requirements such as that satisfied by the use of *un-* (Grimshaw & Vikner 1993).

According to Grimshaw & Vikner, all accomplishment verbs have a two-part event structure consisting of a process and a resultant state (1993:144-145). An "obligatory adjunct," a term which they use rather broadly to also include negating prefixes, serves to "identify" one of the sub-events that would otherwise go unidentified (1993:144). They show that this is required whenever the accomplishment verb is one involving creation, or what they call "constructive accomplishments," because in other types of accomplishments, there is not an unidentified subevent (Grimshaw & Vikner 1993:146-147).[10]

In order to test these assertions, I classified all APP-associated verbs that I had identified as accomplishments (including those that were uncertain or may vary) as either constructive or non-constructive. Grimshaw & Vikner would predict that I should find no constructive verbs in the positive APP category, given that the noun phrase context in which I searched the corpus eliminated obligatory adjuncts of other types. However, this was not what I found. Among the positive APPs identified as associated with accomplishment verbs (47 total), 15% were classified as constructive, 60% as non-constructive, and 26% as uncertain or variable. Of the 11 negative APPs identified as associated with accomplishment verbs, 100% were classified as non-constructive. While this is certainly not enough data to disprove the theory of Grimshaw & Vikner (1993), it strongly suggests that something is going on beyond what they identified. More data and examination are needed to say for sure what role situation aspect and event structure factors play in the distribution of APPs studied here.

4.3 Incremental Theme

The next factor examined was inspired by the first negative APP I discovered, which was *uneaten*. I noted that this form, while occurring much more as a negative than a positive ($n = 27$, $p = 4$), actually had another form that was substantially more common than either the positive or negative: *half-eaten*, with 109 occurrences (sum of *halfeaten* (=3) and *half-eaten* (=106)). This distribution, while relatively unique, suggested that the notion of incremental theme, first proposed by Krifka (1987, 1989), Dowty (1987) and Hinrichs (1985), might be worth considering as a factor. The identification of an incremental theme, or a theme whose parts are mapped onto parts of an event by a telic predicate, allows for a further distinction in the situation aspect realm than those made by the criteria discussed above with reference to Smith (1997) (Dowty 1991:567). Unfortunately, the uncertain classification of the telicity of some verbs from which APPs were derived makes the application of this incremental theme

---

[10]Constructive accomplishments are specifically defined as ones in which "the Theme did not exist in its present form before the event occurred," though this is intended to exclude situations in which the new "form" of the theme is the result of a destructive process (Grimshaw & Vikner 1993:146).

distinction perhaps less valuable as a tool than it might otherwise be.

Like telicity and the other situation aspect components, the notion of incremental theme also applies to verbs rather than adjectives, and thus must be related to the verbs from which APPs were derived. To the extent possible, verbs previously labeled as telic were identified as having or not having an incremental theme. The method used was to consider whether, when the event of the verb was partially done, the theme could be considered partially "verbed"; e.g., if the event of eating an apple is partially done, the apple is partially eaten, but one's toes cannot be said to be partially touched when the event of touching one's toes is partially completed (Dowty 1991). This methodology admittedly allows for a good deal of ambiguity and uncertainty especially when only one person is making judgments, and could doubtless be improved with further study.

Some verbs whose themes I was able to identify as incremental or non-incremental, together with an example theme and their most common form of associated APP, are shown below.

(19) Verbs with incremental themes

      a. eat (e.g., an apple); uneaten

      b. shave (e.g., a face); unshaven

      c. read (e.g., a report); read

      d. paint (e.g., a wall); painted

(20) Verbs without incremental themes

      a. attend (e.g., an event); unattended

      b. sell (e.g., a sandwich); unsold

      c. force (e.g., a smile); forced

      d. find (e.g., an object); found

As can be seen above, verbs associated with both positive and negative APPs were found in both the incremental and non-incremental theme categories. However, among the above and other classified verbs, there seemed to be a possible trend that was consistent with my expectations borne out of *eaten*: the verbs from which positive APPs were derived had a slightly lower proportion of incremental themes among them than did the verbs from which negative APPs were derived. However, it also appeared that incremental theme may be strongly correlated with durativity, as can be seen from the data given in Table 3.[11]

---

[11] "% durative" indicates percentage of APPs associated with a durative verb; likewise for "% instantaneous." "% incremental theme" indicates percentage of telic APPs associated with a verb that has an incremental theme; likewise for "% non-incremental theme."

Table 3: Data on Incremental Theme and Durativity

	APPs with $n=0$	positive APPs	negative APPs
% durative	26	40	48
% incremental theme	29	41	50
% instantaneous	64	40	28
% non-incremental theme	71	49	39

More precise and comprehensive identification of incremental themes is still needed to say with certainty whether they play a role in the distribution of positive and negative APPs. If association with an incremental theme verb is determined to affect the distribution of an APP, it will be necessary to investigate why this is the case.

4.4   Word Origins

Another element of a rather different type examined in this study was historical provenance. Because the negative prefix being used, *un-*, is of Germanic origin, it is reasonable to consider whether the negative APPs, which take the prefix much more commonly than the positive ones, also more commonly share a historical background with the prefix. The reasoning behind this is, in part at least, that sharing a common origin with the prefix means a word has existed in the same language as the prefix for a longer time than a word with unrelated origins. This may correlate with the prefix being more productive with that word, and with other negative prefixes being therefore less likely to be affixed to the word. It is not clear that historical factors could explain the lack of positive instances for some forms; origin is offered solely as a possible reason for the abundance or scarcity of *un-* prefixed negative forms relative to the total number of negative forms produced in any way.

Historical information was sought in *The Oxford English Dictionary* (1989) for the verbs from which all studied APPs were derived, and was available for the vast majority of them. Among these it was found that all words were Indo-European and, as expected, that the two most common origins were Germanic and Latin/French (most often, Latin by way of French). My prediction that a higher percentage of the verbs from which negative APPs were derived would be Germanic than Latin/French, and that the reverse would be true for verbs from which positive APPs were derived, was initially strongly supported by a smaller data set. With the current data, however, this is supported only very weakly, as shown below:

Table 4: Origins of Verbs

	Latin/French	Germanic
% of negative APPs	41	52
% of positive APPs	48	46

(*The Oxford English Dictionary* 1989)

The few verbs whose exact origins could not be determined or were neither Latin/French nor Germanic are not included in the calculations above. Though by no means conclusive or sufficient to solely explain the patterns of distribution, this historical information suggests that word origin may still be yet another factor that plays a meaningful role.

4.5 Informativeness

The last factor examined was borne of the observation that many of the forms found infrequently or not at all seemed, to some degree, semantically or pragmatically anomalous. For example, is *an eaten apple* an apple at all? Under usual circumstances, why would one specify a man as *the identified man* when being identified is part of the default or assumed state of men? Why would one say *the unbent wire* when the equivalent phrase *the straight wire* is available? I group the answers to these questions under the category of informativeness.

It is generally recognized by Griceans and neo-Griceans alike that, under most definitions of informativeness, speakers typically prefer the more informative of two otherwise similar utterances of comparable length. Likewise, given two utterances of similar informativeness, the more concise one is usually preferred. This being the case, one should expect that a phrase of the type studied here is rarely or never found if it is no more informative than a simple [Determiner + Noun] phrase. So if an APP is primarily or always found in negated form, it is predicted that the negated form typically contributes to the informativeness of the phrase in which it occurs, while the non-negated one does not. Likewise, if an APP is primarily or always found in non-negated form, it is expected that that form typically contributes to informativeness, while the negated form does not.

4.5.1 Informativeness Defined

Informativeness has been defined in various ways. According to the Gricean view, informativeness refers simply to the amount of information, or semantic content, in an utterance (Grice 1975). This has been interpreted by Grice himself as well as neo-Griceans in terms of a scale of informativeness (Horn 1976), typically involving entailment (Atlas & Levinson 1981, Grice 1978). In the context of theories of categorization, Giora states: "To say that a category member is informative in a given set is to say that it has more features (information) than necessary for category inclusion" (1988:550). For example, within the set of apples, *an apple smaller than a house* is not very informative, because it describes a characteristic

that holds of all apples and thus might be considered necessary for inclusion in the apple category. In contrast, *a green apple* is much more informative because it describes a feature beyond what is necessary to be considered an apple; certainly not all apples are green, so knowing that an item is an apple does not constitute knowing that the item is green. This is not unlike Sperber & Wilson's treatment of relevance as a combination of the amount of semantic content, the extent to which it causes a "positive cognitive effect," and the amount of processing required for this effect (2005:608). In this study I use a definition of informativeness similar to those of Giora (1988) and Sperber & Wilson (2005): put simply, a phrase or utterance is informative to the extent that it provides information not already present via category membership (cf. Giora) or context (cf. Sperber & Wilson).

4.5.2   Assessment of Informativeness

For the assessment of what information can be said to be provided by category membership, I utilize Pustejovsky's notion of qualia structure: "the structured representation which gives the relational force of a lexical item" and constitutes "modes of *explanation* for a word" that "permit a much richer description of meaning" than a purely decompositional or relational view would provide (1998:76, emphasis in original). The qualia that Pustejovsky proposes for a nominal are shown below:

(21) Pustejovskian qualia

     a. Constitutive: the relation between an object and its constituent parts

     b. Formal: that which distinguishes it within a larger domain

     c. Telic: its purpose and function

     d. Agentive: factors involved in its origin or bringing it about

          (Pustejovsky 1998:76)

I argue that the content of these qualia comprises at least the greater part, and perhaps the whole, of information considered in the assessment of informativeness as I have defined it. For example, Pustejovsky gives *reading* as the telic quale for the lexical item *novel*; thus it is expected that *the read novel* is, under most circumstances, not much more informative than *the novel*, and certainly less informative than *the unread novel* (and thus expected to occur rarely if at all). In fact, usage suggests that for this particular case (and generally for nouns with a telic quale of *reading*), this assessment is correct: the query "a|an|the read [*nn]" returns mostly idiomatic phrases and ones in which it is unclear whether *read* is an APP or some other heteronym, while the query "a|an|the unread [*nn]" returns mostly phrases containing nouns with a telic quale of *reading*, e.g., *report*, *newspaper*, *essays*.

Because informativeness here is being assessed based on a relationship between the verb associated with an APP and the noun that occurs in the phrase with the APP, one cannot directly and in isolation classify APPs or their associated verbs regarding informativeness. What can be directly assessed, however, are the types of nouns that typically occur in phrases with each APP. What proportion of them have qualia that would make the negated form

of the APP more informative, and what proportion have qualia that would make the non-negated form more informative? It is expected that APPs classified as positive typically occur with nouns whose qualia make the negated form of the APP relatively uninformative. Likewise, negative APPs should typically occur with nouns whose qualia make the non-negated form of the APP relatively uninformative. In fact, this was the case for 74% of the APPs examined, whose top two noun collocates had qualia of the type described (COCA).

Though it is imprudent to draw any conclusions without examining more of the data, preliminary results suggest that informativeness assessed via noun qualia may be a highly explanatory factor in the distribution of positive and negative APPs. In addition, the significance of informativeness for APPs has been recognized elsewhere; Ackerman & Goldberg make a claim very similar to the one here via their "non-redundancy constraint," which states: "If the referent of the head noun, N, implies a property P as part of its frame-semantic or encyclopedic knowledge, then an APP is not allowed to simply designate P; it must be further qualified" (1996:21). They also note something else that I found to be true of my data, namely that "contrastive contexts are able to rescue APPs from unacceptability" (1996:23). So, for example, while *the read novel* is generally unacceptable, it may appear in a context in which it is necessary to distinguish a novel that has been read (by some particular person) from one that has not, i.e. in direct opposition to *the unread novel*. In this case, it is quite informative. I have not identified which of the APPs in my data occur in contrastive contexts; this remains an avenue for further research.

4.5.3 Other Factors Related to Informativeness

Having defined what is meant here by informativeness and how it is assessed, I present two other factors that I am grouping with informativeness. The first of these is redundancy with resulting lexical blocking (Blutner 2004). For some of the APPs tested, the negative form was essentially (at least in some senses) completely synonymous with some other unrelated form, as in the example previously given with *unbent* and *straight*.[12] In any case such as this, the synonymous forms ought to be, according to the criteria set out above in section 4.5.1, assessed as equally informative. The fact that in these cases the non-APP form is invariably more common, with the APP form often not found at all, can be explained by lexical blocking. Lexical blocking, as discussed by Blutner, is a process in which "the appropriate use of a given expression formed by a relatively productive process [here, APP formation] is restricted by the existence of a more 'lexicalized' alternative to this expression" (2004:501). Aronoff (1976) and Kiparsky (1982) have shown that this process applies not only to "expressions" but also to inflectional and derivational processes such as those involved in the transformation of the simple form of a verb into an APP. As suggested by Blutner, I believe that this blocking is the result of pragmatic factors; specifically, the use of the anomalous or less lexicalized alternative provides information by suggesting that there is a specific reason for the choice

---

[12]Note that *unbent* (in its adjectival function) and *straight* are alike in failing to distinguish between the meanings "never having been bent" and "having been bent, followed by a reversal of the bending." For *unbent*, this is because *un-* in this case can have either of the two meanings in examples (5) and (6).

of that alternative over the commonly used item. Speakers wanting to avoid this suggestion must therefore avoid the alternative; the majority of the time, this is in fact what happens.

Statistical preemption is a related factor that may either co-exist with lexical blocking or provide an alternative explanation for it. Statistical preemption, according to Boyd & Goldberg, "is an implicit inference speakers make from repeatedly hearing a formulation, B, in a context where one might have expected to hear a semantically and pragmatically related alternative formulation, A. The result is that speakers implicitly recognize that B is the appropriate formulation in such a context, and that A is not appropriate" (Boyd & Goldberg 2009:5). This may serve to further explain lexical blocking in that speakers use the statistical rates at which they hear certain expressions to realize that one expression is, in fact, more lexicalized, and therefore use of the other would be pragmatically marked. Used as a different explanation independent of lexical blocking, statistical preemption can function as a model of how the disuse of APP forms and preference for equivalent non-APP forms is perpetuated.

The last factor to be presented here that is related to informativeness is that of conceptual impossibility (or extreme unlikelihood): in most cases, a phrase that represents something impossible in the world, and therefore very difficult or impossible to even imagine, will not be used.[13] It is for this reason that, for example, one never finds the APPs *unchewed* or *unalerted* with the reversal meaning (see (5)); it is impossible to unchew something or to unalert someone. Similarly, the APP *gathered* is used in the examined data only to refer to groups of things or people (e.g., *crowd*, *assembly*); a group not gathered is arguably not a group, making the idea of an ungathered group highly unlikely if not impossible, and in fact the APP *ungathered* is not found in my data. APPs describing conceptual impossibilities were, as a group, absent from the data, as would be expected.

To conclude this section, I note that in my analysis I have examined a myriad of factors encompassing semantic fields, event structure, historical linguistics and pragmatics. They appear to be relevant to varying degrees. However, I suspect that there is a great deal of interconnectedness between many of the elements analyzed here: for example, semantic fields are likely related to event structure in ways whose relevance to this study may not have been discovered yet. Finally, there are likely further important factors that remain to be found.

5   Conclusion

From the varied aspects examined above, it is clear that there is no single factor responsible for the differences in distribution between negative and positive adjectival past participles in attributive function. A semantic factor such as semantic field may play a small part, as suggested by the fact that some fields seem to be overrepresented among the negative APPs (or, from the opposite perspective, underrepresented among the positive APPs). As noted

---

[13]We may perceive as an exception to this a phrase, such as *the unbuilt building*, that represents something that is physically impossible in the world but not conceptually impossible, and thus useful to be able to talk about.

above, this may also relate to one or more of the other factors. Though difficult to apply to an examination of adjectives, certain components having to do with situation aspect may be relevant. Though I had many APPs associated with accomplishment verbs, Grimshaw & Vikner's predictions about these were not borne out in my data. Incremental themes were slightly more common among negative APPs' verbs, but this appears to be strongly correlated with durativity. Historical information suggests a slightly higher percentage of Germanic origins among negative APP verbs than positive, which may relate to the Germanic origin of the negative prefix studied here. Of those examined here, informativeness is the factor with the most explanatory potential: in general, the form (negated or non-negated) of an APP that is predicted to be more informative is what I find occurring most in my data. This portion of my analysis is also strengthened by its general agreement with that of Ackerman & Goldberg (1996).

In order to paint a fuller picture of what is responsible for the distributional pattern noted here, further study is needed. Clearly, 203 forms is only a very small subset of the adjectival past participles in English that can occur with a negative prefix and in attributive function. Examining many more forms would improve the viability of the study. It may also be beneficial to look at negative prefixes besides *un-*; this could lead to support for the historical relationships proposed here. There are undoubtedly other factors at work in this distribution that were not considered here, and these should be explored. One possible angle is to consider the distribution patterns of adjectives not derived from verbs to see if anything similar exists there; if so, this would suggest that verbal characteristics could be less of a factor than previously thought. Another direction for research is to investigate similar forms in other languages, to see whether this distribution is a phenomenon unique to English. Whether it is or not, this information would likely be valuable in determining what other factors to evaluate.

In future research, pursuing an explanation of exactly how semantic factors interact with event structure may be worthwhile. Improved assessment of situation aspect factors would maximize their potential usefulness. Historical information, while mostly complete for the data here, can be expanded as discussed with other prefixes. In addition, it may be informative to examine what verbs of origins besides Latin/French and Germanic do with respect to negative prefixes of various ancestries. Finally, as mentioned in the section on informativeness, given the promising results of the preliminary work, a complete study of qualia-based informativeness is an obvious next step. Other methods of assessing informativeness may also be worth investigating. While this paper has made a significant start by providing analyses of several potential factors in the distribution of positive and negative adjectival past participles, I intend to pursue this study further by examining some of the approaches discussed above.

References

Ackerman, Farrell & Adele E. Goldberg. 1996. Constraints on Adjectival Past Participles. In Adele E. Goldberg (ed.), *Conceptual Structure, Discourse and Language*, 17-30. Stanford, CA: CSLI Publications.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar.* Cambridge, MA: MIT Press.

Atlas, Jay David & Stephen C. Levinson. 1981. It-Clefts, Informativeness, and Logical Form: Radical Pragmatics (Revised Standard Version). In Peter Cole (ed.), *Radical Pragmatics*, 1-62. New York: Academic Press, Inc.

Blutner, Reinhard. 2004. Pragmatics and the Lexicon. In Laurence R. Horn and Gregory Ward (eds.), *The Handbook of Pragmatics*, 488-514. Malden, MA: Blackwell.

Boyd, Jeremy K. & Adele E. Goldberg. Forthcoming. Learning What Not To Say: The Role of Categorization, Statistical Preemption, and Discounting in a-adjectives. (2009 draft).

Davies, Mark. 2008 –. The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. Available online at http://www.americancorpus.org.

Dowty, David. 1987. Aktionsarten, NP semantics, and the structure of events. Paper presented at the joint Association for Symbolic Logic/Linguistic Society of America Conference on Logic and Natural Language, Stanford University.

Dowty, David. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67(3): 547-619.

Giora, Rachel. 1988. On the Informativeness Requirement. *Journal of Pragmatics* 12: 547-565.

Grice, H. P. 1975. Logic and Conversation. In P. Cole & J. L. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, 41-58. New York: Academic Press.

Grice, H. Paul. 1978. Further Notes on Logic and Conversation. In Peter Cole (ed.), *Syntax and Semantics 9: Pragmatics*, 113-128. New York: Academic Press.

Grimshaw, Jane & Sten Vikner. 1993. Obligatory Adjuncts and the Structure of Events. In Eric Reuland & Werner Abraham (eds.), *Knowledge and Language*, vol. II, Lexical and Conceptual Structure, 143-155. Dordrecht, Netherlands: Kluwer Academic Press.

Hinrichs, Erhard. 1985. *A compositional semantics for aktionsarten and NP reference in English.* Columbus: Ohio State University dissertation.

Horn, Laurence Robert. 1976. *On the Semantic Properties of Logical Operators in English.* Bloomington, IN: Indiana University Linguistics Club.

Huddleston, Rodney & Geoffrey K. Pullum. 2005. *A Student's Introduction to English Grammar.* Cambridge: Cambridge University Press.

Kiparsky, Paul. 1982. Word-formation and the Lexicon. In F. Ingeman (ed.), *Proceedings of the 1982 Mid-America Linguistic Conference.* Lawrence, KS: University of Kansas.

Krifka, Manfred. 1987. *Nominal reference and temporal constitution: Towards a semantics of quantity. (FNS-Bericht 17).* Tubingen: Forschungsstelle fur naturliche Systeme, Universitat Tubingen.

Krifka, Manfred. 1989. *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aktionsarten.* Munich: Wilhelm Fink.

Lakoff, George. 1970. *Irregularity in Syntax.* New York: Holt, Rinehart and Winston, Inc.

Levin, Beth. 2005. Semantic Prominence and Argument Realization IV: Is Aspect a Semantic Determinant of Argument Realization? LSA Linguistic Institute class handout.

The Oxford English Dictionary. 2nd ed. 1989. *OED Online.* Oxford University Press. Dec. 2009 < http://dictionary.oed.com.floyd.lib.umn.edu/>.

Pustejovsky, James. 1998. *The Generative Lexicon.* Cambridge, MA: MIT Press.

Smith, Carlota S. 1997. *The Parameter of Aspect.* Dordrecht: Kluwer Academic Publishers.

Sperber, Dan & Deirdre Wilson. *Relevance: Communication & Cognition*, 2nd edn. Oxford: Blackwell.

Verkuyl, H. J. 1972. *On the Compositional Nature of the Aspects.* Dordrecht: D. Reidel Publishing Company.

Wasow, Thomas. 1977. Transformations and the Lexicon. In Peter W. Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal Syntax*, 327-360. New York: Academic Press, Inc.

University of Minnesota – Twin Cities
Department of Linguistics
214 Nolte Center
315 Pillsbury Drive SE
Minneapolis, MN 55455

vanam008@umn.edu

6 Appendix

The numbers after each word indicate positive occurrences p, negative occurrences n, and the ratio of positive/negative occurrences r, respectively. Words are arranged in order of increasing r values.

The symbol * next to a word indicates that I noticed some idiomatic or other use that seemed to skew the numbers or might potentially do so.

A blank r cell indicates that a ratio could not be calculated due to division by zero.

APP	p	n	r
mourned	0	2	0
harmed	0	3	0
noticed	0	12	0
doubted	0	34	0
attended	1	42	0
seen	16	539	0
eaten	4	27	0.1
foreseen	6	104	0.1
answered	11	130	0.1
questioned	13	103	0.1
asked	4	18	0.2
counted	4	23	0.2
discovered	16	71	0.2
sung*	21	114	0.2
tamed	15	51	0.3
interrupted	38	119	0.3
named	108	310	0.3
wanted	167	386	0.4
proved	5	9	0.6
identified	248	402	0.6

APP	p	n	r
spoken	264	428	0.6
explained	59	84	0.7
moved	4	5	0.8
sold	15	20	0.8
examined	56	74	0.8
manned*	135	164	0.8
known	1247	1477	0.8
zipped	6	7	0.9
shaven	25	27	0.9
fastened	1	1	1
served	2	2	1
tasted	2	2	1
impressed	11	11	1
made	80	84	1
expected	1913	2003	1
heard*	21	19	1.1
decided	166	135	1.2
deserved	45	35	1.3
challenged	64	51	1.3
tested	100	74	1.4

The Distribution of English Adjectival Past Participles

APP	p	n	r
sought	15	10	1.5
appreciated	16	11	1.5
tied*	37	24	1.5
opened	178	118	1.5
announced	115	67	1.7
checked*	77	43	1.8
invited	112	61	1.8
lit	271	153	1.8
guarded	109	56	1.9
delivered	21	11	1.9
dreamed	8	4	2
intended	777	382	2
clothed	17	8	2.1
met*	82	39	2.1
paid	492	222	2.2
marked	976	375	2.6
matched	165	60	2.8
hurried	144	47	3.1
listed	211	65	3.2
finished	1559	487	3.2
used	741	205	3.6
brushed	19	5	3.8
read*	46	12	3.8
started	4	1	4
swept	38	8	4.8
licensed	406	78	5.2
kept	34	6	5.7
peeled	52	9	5.8
released	83	14	5.9
loved	110	15	7.3
timed	55	7	7.9
proven	545	67	8.1
satisfied	199	23	8.7
cut	468	50	9.4
approved	276	29	9.5
called	31	3	10.3
educated	664	64	10.4
written	2526	244	10.4
divided	463	44	10.5
developed	926	87	10.6
permitted	48	4	12

APP	p	n	r
corrected	98	8	12.3
loaded	466	37	12.6
heated	818	64	12.8
limited	4780	363	13.2
informed	583	42	13.9
built*	159	11	14.5
said*	179	12	14.9
decayed	45	3	15
cleared	126	8	15.8
interested	205	13	15.8
broken	4080	232	17.6
created	178	10	17.8
guided	427	24	17.8
lighted	366	19	19.3
bitten	20	1	20
returned	81	4	20.3
changed	511	25	20.4
painted	916	44	20.8
contained	42	2	21
caught	22	1	22
chewed	22	1	22
bleached	90	4	22.5
numbered	113	5	22.6
burned	230	10	23
risen*	78	3	26
shaken	81	3	27
reported*	998	37	27
prepared	585	21	27.9
imagined	339	12	28.3
embarrassed	143	5	28.6
diminished	233	8	29.1
needed	590	20	29.5
completed	707	23	30.7
melted	155	5	31
damaged	593	19	31.2
balanced	2559	80	32
found	64	2	32
practiced	221	6	36.8
assigned	300	8	37.5
collected	264	7	37.7
spent	208	5	41.6

APP	p	n	r
concerned	324	7	46.3
described	50	1	50
polished	553	11	50.3
driven	101	2	50.5
charged	157	3	52.3
improved	758	14	54.1
forced	597	10	59.7
wound*	129	2	64.5
hanged	65	1	65
amused	131	2	65.5
done*	267	4	66.8
chosen	596	8	74.5
buried	242	3	80.7
measured	486	6	81
predicted	434	5	86.8
surprised	176	2	88
learned	394	4	98.5
decorated	201	2	100.5
coiled	118	1	118
calculated	483	4	120.8
packed	526	4	131.5
destroyed	136	1	136
hurt	153	1	153
chopped	174	1	174
burnt	183	1	183
observed	811	4	202.8
frightened	408	2	204
hung*	227	1	227
worn	490	2	245
cracked	492	2	246
flooded	261	1	261
excited	276	1	276
fallen	1127	4	281.8
repeated	626	2	313
frozen	1631	5	326.2
noted	463	1	463
faded	498	1	498
split*	1118	2	559
accepted	572	1	572
forgotten	583	1	583
admired	31	0	

APP	p	n	r
allowed	33	0	
annoyed	70	0	
applauded	1	0	
arisen	1	0	
arranged	141	0	
awarded	4	0	
awoken	0	0	
banned	136	0	
bathed	0	0	
bent	229	0	
blown	76	0	
boiled	110	0	
bruised	205	0	
burst	101	0	
dug	3	0	
drawn	86	0	
dreamt	1	0	
encouraged	0	0	
enjoyed	1	0	
forgiven	10	0	
gathered	164	0	
given	4626	0	
judged	9	0	
missed	307	0	
offended	59	0	
offered	56	0	
placed	4	0	
put*	33	0	
rung*	7	0	
scratched	43	0	
sent	3	0	
shot*	314	0	
shut	30	0	
stolen	699	0	
stuck	78	0	
suggested	310	0	
taken	10	0	
torn	499	0	
thrown	44	0	
woken	0	0	
worried	308	0	