A-adjectives, statistical preemption, and the evidence: Reply to Yang (2015)

Adele E. Goldberg
Princeton University

Jeremy K. Boyd
University of California, San Diego

Adele E. Goldberg (corresponding author)
adele@princeton.edu
Department of Psychology
Green Hall
Princeton University
Princeton, NJ 08540

Jeremy K. Boyd
jboyd@ucsd.edu
Human Development Program
University of California, San Diego
9500 Gilman Dr.
La Jolla, CA 92093-0115

Abstract

A certain class of English adjectives known as *a*-adjectives resists appearing attributively as prenominal modifiers (e.g., *??the afraid boy, ??the asleep man*). Boyd & Goldberg (2011) had offered experimental evidence suggesting that the dispreference is learnable on the basis of categorization and *statistical preemption*: repeatedly witnessing predicative formulations in contexts in which the attributive form would otherwise be appropriate. The present paper addresses Yang (2015)'s counterproposal for how *a*-adjectives are learned, and his instructive critique of statistical preemption. The counterproposal is that children receive evidence that *a*-adjectives behave like locative particles in occurring with certain adverbs such as *far* and *right*. However, in an analysis of the 450 million word COCA corpus, the suggested adverbial evidence is virtually non-existent *(e.g., *far alive; *straight afraid*)*. In fact, these adverbs occur much more frequently with typical adjectives (e.g., *far greater, straight alphabetical*). Furthermore, relating *a*-adjectives to locative particles does not provide evidence of the restriction, because locative particles themselves can appear as prenominal modifiers (*the down payment, the outside world*). The critique of statistical preemption is based on a 4.3 million word corpus analysis of child directed speech that suggests that children cannot amass the requisite evidence before they are three years old. While we clarify which sorts of data are relevant to statistical preemption, we concur that the required data is relatively sparsely represented in the input. In fact, recent evidence suggests that children are not actually cognizant of the restriction until they are roughly ten years old, an indication that input of an order of magnitude more than 4.3 million words may be required. We conclude that a combination of categorization and statistical preemption is consistent with the available evidence of how the restriction on *a*-adjectives is learned.

Keywords: a-adjectives, statistical preemption, categorization

1.   *A*-ADJECTIVES

There are certain restrictions on language that do not follow from general semantic, phonological, or syntactic facts in any obvious way. One such case involves a class of adjectives which begin with a syllabic schwa ("a") and resist appearing prenominally in attributive position (Bolinger 1971, Huddleston & Pullum 2001, Larson & Marušič 2004). We refer to these as *A*-ADJECTIVES; examples are provided in (1):

1. A-ADJECTIVES
    a. ?? the/an asleep child   e. ?? The/an alive monster
    b. ?? the/an afraid man   f. ?? The/an ablaze building
    c. ?? the/an alone boy   g. ?? the/an afloat ship
    d. ??the/an aware woman   h. ?? the abloom flowers

Near synonyms (2) and non-*a*-adjectives with similar phonology (3) readily appear attributively:

2. Semantic near-synonyms
    a. the/a sleeping child   e. the/a living monster
    b. the/a scared man   f. the/a burning building
    c. the/an isolated boy   g. the/a floating ship
    d. the/a mindful woman   h. the/a blooming flowers

3. Phonologically related non-*a*-adjectives
    a. the/an adult male   c. the/an acute sickness
    b. the/an astute comment   d. the/an aloof woman

The key distinction between the *a*-adjectives in (1) and the non-*a*-adjectives in (3) appears to be that each of the words in (1) is morphologically segmentable into *a*- plus a semantically related stem (Boyd & Goldberg 2011, Coppock 2008). Other adjectives with similar phonology, but which are not segmentable, fall outside the *a*-adjective category. For example, /dult/ in *adult* is not an English morpheme, and neither are the syllables /stut/, /cyut/ and /luf/ in *astute, acute* and *aloof*. We can conclude that while *a*-adjectives are partly defined by the way they sound, the category does not reduce to phonological or semantic characteristics.

The restriction on *a*-adjectives is diachronically motivated by the history of many of the adjectives as prepositional phrases. For example, *asleep* comes from the Old English prepositional phrase, *on sleep* (Boyd & Goldberg 2011; Long 1969). As prepositional phrases, it made sense that they would not occur attributively. Today's speakers, however, are generally unaware of the historical origin of these adjectives, and yet they implicitly recognize and respect their unusual distributional pattern. Moreover, certain adjectives have been assimilated to the restricted subcategory of *a*-adjectives. For example, *afraid* was never a prepositional phrase, but it nonetheless resists attributive use (Boyd & Goldberg 2011).

The unusual distribution of *a*-adjectives poses a clear learnability challenge. How do speakers learn to avoid using these adjectives in prenominal attributive position? Notice that positing an invisible feature or some type of underlying syntactic structure is unhelpful, unless the learner has some way of inferring the invisible feature or underlying structure from the surface input. Without evidence that *a*-adjectives are unlike other adjectives, some type of indirect negative evidence of the restriction is needed (Goldberg 2011a, Yang 2015).

2.   LEARNING THE RESTRICTION VIA CATEGORIZATION AND STATISTICAL PREEMPTION (BOYD & GOLDBERG 2011)

Statistical preemption or "blocking" of a target form is the process of learning to avoid a potential target form because a competing form has been consistently witnessed instead in

contexts in which the target form would otherwise have been appropriate. This is the widely accepted view of how children learn to avoid morphological overgeneralizations such as *goed*: *went* is consistently witnessed in contexts in which *goed* would otherwise have been appropriate (Aronoff 1976). The mechanism involved is simple error-driven learning of the sort that has been a mainstay in psychology for decades (Rescorla & Wagner 1972). A number of researchers have proposed that statistical preemption is capable of scaling up to account for certain non-occurring syntactic formulations that have readily available competing alternatives (Ambridge et al. 2012, Brooks & Tomasello 1999, Brooks & Zizak 2002; Clark 1987; Goldberg 1995, 2006, 2011b, Poser 1992, Robenalt & Goldberg 2015).

Boyd and Goldberg (2011) investigated the role that statistical preemption might play in the acquisition of the restriction on attributive use of *a*-adjectives. Three experiments with the same general structure were reported. Before performing a production task, and under the guise of providing instructions, undergraduate participants witnessed the experimenter use two adjectives attributively three times each,; and two other adjectives predicatively, three times each, in a relative clause. The production task required speakers to identify one of two contrasting animals, which elicited the use of certain target adjectives. The target adjectives included real *a*-adjectives (e.g., *asleep, afloat*), novel *a*-adjectives (e.g., *ablim, adax*), real near-synonyms (e.g., *sleepy, floating*) and novel non-*a*-adjectives (e.g., *chammy, flitzy)*. The contrastive context in the production task advantaged an attributive description (e.g., *The sleepy cow*) over a predicative use in a relative clause (e.g., *The cow that's sleepy*), but either was possible and both were used.

Filler items were included and counterbalanced so that each critical item was preceded half of the time by fillers that reliably elicited attributive responses (prototypical adjectives; e.g., *fast*, *old*, and *strong*), and the other half by fillers that reliably elicited relative clause responses (third-person present tense verbs such as *bites*; e.g., *The cat that bites*). These were included in order to guard against spurious priming effects.

Experiment 1 provided a baseline measure of how speakers tended to use each type of adjective. During the exposure phase, the two adjectives witnessed attributively and the two adjectives witnessed predicatively were all non-*a*-adjectives, and none were used as target adjectives in the production task. Results demonstrated that speakers do in fact avoid using real *a*-adjectives attributively, as compared with non-*a*-adjectives. Speakers also avoid using novel *a*-adjectives attributively to some extent, although not as strongly as they do for real *a*-adjectives. The latter fact implies that the undergraduate participants implicitly recognized an abstract *a*-adjective *category* that is capable of weakly attracting new members like *ablim* and *adax*, whose morphology was opaque.

Experiment 2 provided the key test of statistical preemption. A new group of undergraduates was presented with three preemptive (predicative) uses of each of two novel *a*-adjectives (e.g., *The hamster that's ablim*). They witnessed the same number of attributive uses of different novel adjectives that did not serve as target adjectives during the production task (*The tooky hamster*). Participants then described scenes as in the first experiment. The preemptive exposure had a clear effect: there was a dramatic reduction in attributive responses for all four novel *a*-adjectives, and no reduction for novel non-*a*-adjectives.

A third group of speakers provided evidence that speakers are savvy about what counts as a preemptive context. In this final experiment, two novel *a*-adjectives were again witnessed in relative clauses, but this time there was an independent motivation for the relative clause use, in that the novel *a*-adjectives were conjoined with a complex adjective phrase, e.g. (*The hamster that's ablim and proud of himself)*. Since complex adjective phrases are independently

unacceptable in attributive position (e.g., *The proud of himself hamster*), learners should not assume that *ablim* is responsible for the relative clause use.[1] Participants behaved shrewdly in the face of this "pseudo"-preemptive exposure: it was essentially ignored, and descriptions were provided much as they were in the first experiment.

A post-test determined that participants were unaware of the relevant manipulations. Thus a combination of statistical preemption and categorization seems to be effective as a way of implicitly learning restrictions on use.

3.   YANG (2015)'S CHALLENGE

Yang (2015), hereafter Y15, offers a counterproposal for how the restriction on *a*-adjectives can be learned. He argues that speakers witness *a*-adjectives occurring with certain adverbs such as *right* or *straight,* and that this induces learners to treat *a*-adjectives like locative particles (e.g., *out, in, over*; see also Bruening 2011a,b).[2] It is suggested that a parallel with locative particles explains *a*-adjectives' resistance to attributive use.

Y15 argues further that statistical preemption is not a viable approach to learning the restriction. In a search of a 1.9 million-word corpus of three-year-old children's speech, he finds no attributive uses of *a*-adjectives, and concludes that children know the restriction on *a*-adjectives by the time they are three years old. In a search of a slightly larger subset of 4.3 million-words of child-directed speech from CHILDES (MacWhinney 2000), he finds that the majority of *a*-adjectives within the corpus do not appear in relative clauses. On this basis, Y15 argues that the required evidence for statistical preemption is unavailable to young children.

Y15 also finds a handful of adjectives (e.g. *careful*) that only occurred predicatively in the corpus, but are not actually restricted from attributive use, and certain other adjectives that only appeared attributively in the corpus but can actually occur predicatively as well (e.g., *green*). On the basis of these facts, Y15 argues that statistical preemption would lead to certain erroneous restrictions.

In order to account for the experimental findings of Boyd & Goldberg (2011)'s second experiment, which purported to show evidence in favor of statistical preemption, Y15 suggests

---

[1] Certain *a*-adjectives—e.g., *abud* in (a) below—receive an ongoing process interpretation and obligatorily occur with prepositional phrases (c). The required prepositional phrase provides an independent explanation for the constraint against their appearing attributively (b).

a. The tree is abud with green shoots.
b. ??An abud tree is a beautiful thing to see. (Yang 2015)
c. ??The tree is abud.

On the other hand, the *a*-adjectives under discussion here—*asleep, afraid, awake, aware, alive, afloat*—do not require prepositional phrases, and neither do they necessarily share the same ongoing process interpretation (pace Y15; Coppock 2008: 181). Thus, appeal to the productive *a*- aspectual morpheme does not account for the restriction, except perhaps as support for a more general *a*-adjective category.

[2] We follow Yang's terminology here in referring to these as particles, but in fact, we are persuaded by Huddleston & Pullum (2001)'s arguments that there is no principled reason to distinguish most of them from prepositions.

that because speakers witnessed relative clause structures being used in the exposure phase, they were simply primed to increase their own relative clause use during the production task, thereby depressing the rate of attributive responses. We respond to Y15's proposal and critique of statistical preemption in turn below.

4.   *A*-ADJECTIVES *SEEM* LIKE ADJECTIVES, NOT PARTICLES OR PREPOSITIONAL PHRASES

Instead of relying on statistical preemption, Y15 suggests that there exists evidence that *a*-adjectives pattern with locative particles (e.g., *up, down, on, in, out*), and unlike typical adjectives; the restriction on prenominal attributive use is argued to be learned on this basis. Relatedly, Bruening (2011) argues that *a*-adjectives are prepositional phrases underlyingly, not adjectives, and that this is why they resist occurring attributively. Table 1 classifies the relevant cases into typical adjectives (Class 1), which readily occur attributively; *a*-adjectives (Class 2), which begin with a morphologically segmentable unstressed schwa and resist occurring attributively; prepositional phrases (Class 3); and locative particles (Class 4). For Y15 or Bruening's line of argumentation to be valid, children must witness evidence that *a*-adjectives (Class 2) systematically pattern like prepositional phrases (Class 3) or locative particles (Class 4), and unlike typical adjectives (Class 1).

1. TYPICAL ADJECTIVES	2. *A*-ADJECTIVES	3. PREPOSITIONAL PHRASES:	4. LOCATIVE PARTICLES
red	*asleep*	*on the table*	*up*
sleepy	*afloat*	*into the room*	*down*
floating	*afraid*	*to the house*	*on*
full	*alone*	*at two o'clock*	*in*
huddled	*ablaze*	*in the mind*	*inside*
pinkish	*abloom*	*inside the box*	*around*
absurd	*alive*	*out of the city*	*away*
acute	*awake*	*around the ring*	*across*
aloof	*aware*	*away from them*	*out*

Table 1: Examples of typical adjectives that may appear attributively; *a*-adjectives which resist attributive use; prepositional phrases; and locative particles.

It should be clear that classifying *a*-adjectives with locative forms does not address the learnability issue unless one provides a way for learners to *recognize* that *a*-adjectives pattern with locative forms (Goldberg 2011a). Y15 and Bruening (2011) propose one such critical piece of evidence. In particular, *a*-adjectives, prepositional phrases, and locative particles, but not regular adjectives are said to appear with "*right*-type" adverbs: *right, well, far* and *straight* (Y15, 2015: 15). Y15 suggests that "probably not all *a*-adjectives may be used with *right*-type" modification… [but]... such adverbial modification can not appear at all with typical adjectives" (2015: 8).

In actuality, *right*-type adverbs only very rarely occur with only a subset of *a*-adjectives, as is demonstrated by a search of the vast 450 million word Corpus of Contemporary American English (COCA; Davies 2008). Table 2 includes an exhaustive list of *a*-adjectives that occur

even once with the suggested adverbs (left panel) in COCA. With the single exception of *well aware*, no *a*-adjective appears regularly with any of the proposed adverbs, and many *a*-adjectives (e.g., *asleep, ablaze, afloat, abloom*) do not occur at all. At the same time, many typical adjectives *do* regularly occur with *right*-type adverbs as shown in the right panel of Table 2.

Attested uses in the 450 million word COCA corpus

EXHAUSTIVE combinations of "right type" adverbs + *a*-adjectives (class 2)	#	NON-exhaustive combinations "*right*-type" adverbs + **typical adjectives (class 1)**	
		Straight winning	14
		Straight party-line	7
Straight + a-adjective[3]	0	Straight alphabetical	2
		Far greater	1300
Far astray	6	Far worse	720
Far adrift	4	Far better	719
Far alone	3	Far different	526
		Far higher	386
		Right good (dialect)	28
Right afraid	9	Right proud (dialect)	11
Right alone (?)	2	Right fine (dialect)	8
Right alike (?)	2	Right honorable	7
		Well pleased	70
Well aware	1591	Well worthwhile	12
Well alone	12	Well familiar	16
Well alive	1	Well used	17
Well afraid	1	Well early	7

Table 2. Attested instances of *right*-type adverbs with all *a*-adjectives (left), and with a subset of typical adjectives that may be used attributively (right). Searches were performed on COCA using strings such as, e.g., *straight* [j*], where [j*] ranges over all adjectives.

Perhaps the reason that *right*-type adverbs were thought to be diagnostic was that certain locative prepositions and adverbs (*across, around, away, apart ahead*) were classified as *a*-adjectives in Y15's analysis, and these forms can occur with *right*-type adverbs. But *across, around, away, apart*, and *ahead* are standardly analyzed as locative particles, not adjectives. Conflating these instances with the class of *a*-adjectives begs the question as to whether locative particles and *a*-adjectives have identical distributions. We count only forms that are tagged as *adjectives* in COCA and Google dictionary as *a*-adjectives here. Thus we can conclude that the "the signature evidence that relates *a*-adjectives to locative particles and prepositional phrases" (Yang: 2015: 15) is not compelling.

No other positive evidence by which to learn the restriction on *a*-adjectives is suggested. Relevantly, even if *a*-adjectives *were* linked to the distribution of locative particles, the link could not be used to learn the critical restriction on *a*-adjectives, because locative particles themselves can sometimes occur in attributive position, as the examples in Table 3 demonstrate.

---

[3] That is, no *a*-adjectives are returned by an exhaustive COCA search of "straight [j*]" (i.e., *straight* followed by any adjective).

the near future	the out lesbian
the past year	a down mood
the outside world	the above examples
the inside track	

Table 3: Examples of locative particles that appear attributively[4]

To summarize, *right*-type adverbs do not provide evidence that *a*-adjectives pattern with locative particles and unlike other adjectives. Moreover, even if *a*-adjectives did pattern with locative particles, it would not explain their avoidance of attributive use because locative particles themselves do not uniformly show the relevant constraint against attributive use.

Bruening's (2011a) proposal that *a*-adjectives are actually prepositional phrases underlyingly could in principle offer an explanation of *a*-adjectives' unusual distribution, since prepositional phrases rarely if ever occur attributively (cf. also Coppock 2008 for a proposal to treat *a*-adjectives as phrasal). In fact, we had suggested that the diachronic status of *a*-adjectives as prepositional phrases motivated their synchronic distribution (Boyd & Goldberg 2011: 62). But for *a*-adjectives to actually *be* prepositional phrases synchronically, it would require that *a*-adjectives patterned with prepositional phrases generally, and unlike other adjectives. And, if the only evidence is based on unacceptable formulations—that is, on *non*-occurrences in certain constructions—then the learnability question remains. That is, as both Y15 and Goldberg (2011a) emphasize, unless there exists positive evidence that *a*-adjectives are underlyingly prepositional phrases, learners would still have to learn a *restriction* (or a set of restrictions), and that requires recourse to some type of indirect negative evidence, such as statistical preemption. The only positive evidence Bruening (2011a,b) had offered was co-occurrence with *right*-type adverbs, but as we've seen, these adverbs do not distinguish the class of *a*-adjectives from other adjectives (Table 2). That is, *right*-type adverbs do not provide a way to learn that *a*-adjectives form a special class together with prepositional phrases and distinct from typical adjectives.

In fact, there exists evidence in favor of treating *a*-adjectives as adjectives—i.e., with Class 1—rather than with prepositional phrases (Class 3) or locative particles (Class 4). Importantly, like the vast majority of other adjectives*,* and unlike prepositional phrases and particles, *a*-adjectives semantically designate a property that is not primarily locative.

Distributional tests of grammatical categories are notoriously imperfect (Croft 1991), but being a direct complement of the verb *seem* provides a classic test for adjective status in English (Jackendoff 1972; Lakoff 1970), and the judgments in (4) indicate that *a*-adjectives are adjectives according to this test (Goldberg 2011a):

4.a. The child seemed alive/afraid/afloat/alone/aghast.
b. ?The child seemed on the table/at two o'clock/out of the house.
c. ?The child seemed in/on/out/inside/around/ahead/about.

---

[4] There are many more particle combinations that play the role of modifiers within compounds with the stress pattern that is characteristic of compounds including *the up escalator, the down escalator, the on button, the away game, the on ramp, the in crowd, the up side*, *the down position*.

Bruening (2011b) correctly observes that this restriction is not categorical, as *seem* does at times appear with particles or prepositional phrases (usually cases that semantically designate a non-spatial property like *under the weather*). But a COCA search reveals that directly following *seem,* there are almost four times as many adjectives (> 11,000) as prepositions/particles (< 3000), once *to* infinitives and *like* phrases are discarded.

Moreover, if we assume that conjunction is more likely to involve constituents of the same grammatical category than distinct grammatical categories (Chomsky 1957; Pearl & Sprouse 2013; Yang 2015), then it is also relevant that *a*-adjectives are much more frequently conjoined with other adjectives than with prepositions/particles. Table 4 provides the search results on the relevant conjunctions from COCA.

a-adj.	order <a-adj.> and <adj.>	order <adj> and <a-adj.>	**Total conjunctions with** ADJECTIVES	order <a-adj> and < P>	order <P> and <a-adj.>	**Total conjunctions with** PREPOSITIONS
alive	1213	240	**1453**	100	0	**100**
alone	273	261	**534**	244	0	**244**
afraid	40	178	**218**	7	2	**9**
aware	72	90	**162**	1	3	**4**
asleep	24	36	**60**	14	0	**14**
afloat	4	8	**12**	1	0	**1**

Table 4: Search results from COCA for the conjunction of particular *a*-adjectives with adjectives (both orders: <*a*-adjective > and [j*]; [j*] and <*a*-adjective >); and with prepositions (both orders: <*a*-adjective > and [i*]; [i*] and < *a*-adjective>). Note that the first order includes conjunctions with prepositional phrases.

As Bruening (2011b) notes, there are certain prepositional phrases that are inseparable and have adjective-like meanings, such as *at ease* and *on fire*. These cases are quite adjective-like in their distribution and therefore they behave like *a*-adjectives in many ways: they can readily appear directly after *seem* and can be conjoined with uncontroversial adjectives. They also resist attributive prenominal position *(*the at ease man)*.[5] But, unlike *a-adjectives*, these inseparable collocations are recognizable as prepositional phrases because each involves a

---

[5] Bruening (2011b) emphasizes a claim that the restriction on both *a*-adjectives and inseparable prepositional phrases like *at ease* holds regardless of whether they are used prenominally or postnominally. But it is not clear that there is a postnominal restriction, nor that the postnominal uses are best characterized as attributive and not predicative. The following examples are from COCA:

    a. Finally, an hour later- " Click. A photograph of a house on fire.
    b. I saw a woman alone at one of the tables,
    c. He gives the impression of a man at ease,
    d. Vemon pulled tenderly, as if dragging a man asleep and trying not to wake him.

recognizable preposition and a recognizable noun. That is, we know from much work in morphology, that the subparts of strings are analyzed to the extent that the frequency ratio of the subpart to the whole string is high (Hay & Baayen 2005). Because *at* is used as a highly frequent preposition independently of the string *at ease*, and *ease* is also regularly used as a noun in other contexts, the subparts: *at* + *ease* will be recognized as preposition + noun—that is, as a prepositional phrase. Their resistance to being used in prenominal attributive position follows from this fact.

This is not true of *a*-adjectives. While we have noted that *a*-adjectives involve a recognizable root, the *a*- prefix is not independently used as a preposition *(*a the roof)*, and very few of the roots can be used independently as nouns (**the fraid; *the ware; *the live; *the lone*). Thus learners have no reason to consider *a*-adjectives to be prepositional phrases, aside from their resistance to appearing prenominally. We therefore need an account of how learners come to implicitly recognize this resistance.

Thus, the preponderance of the evidence (semantic and distributional) suggests that *a*-adjectives as discussed here and in Boyd & Goldberg (2011) are adjectives. Since the vast majority of other adjectives readily appear in attributive, prenominal position—including adjectives with very similar meanings and similar phonology—the fact that this circumscribed class of *a*-adjectives does not, requires learners to recognize a restriction on the basis of the language they witness. Boyd & Goldberg (2011) had suggested that a combination of statistical preemption and categorization offers a viable route to learning the restriction on *a*-adjectives.[6]

5.   ADDRESSING PROPOSED CHALLENGES TO STATISTICAL PREEMPTION + CATEGORIZATION
Recall that Y15 had found that the 4.3 million word corpus did not contain relative clause uses of all *a*-adjectives. But it is important to bear in mind that preemptive contexts for the restriction on *a*-adjectives are not provided by only (or all) relative clause uses of these adjectives. The discourse context of *any* predicative use of an *a*-adjective could be relevant as long as an attributive use would normally be favored in that context. In addition, not all relative clause uses provide preemptive contexts, because a relative clause can be motivated by factors that are independent of the adjective (recall Boyd & Goldberg, experiment 3). What is required is that the attributive use could be expected, but some other formulation is repeatedly witnessed instead. The context-dependent nature of statistical preemption makes it is difficult to quantify using automated techniques (but see Goldberg 2011b).

---

[6] The restriction on *a*-adjectives has alternatively been suggested to follow from a general constraint against right-headed forms in attributive prenominal position (Larson & Marušič 2004: 270, note 2; David Pesetsky, personal communication 7/15). While such a restriction might be used to rule out the attributive use of prepositional phrases (**the on top cherry*), and complex adjectives (**the proud of itself fox*) (Williams 1981), the restriction on *a*-adjectives does not readily yield to this analysis. Even if we accept the idea that *a*- is a derivational morpheme that can be used to turn roots *fraid, live, ware*, etc. into adjectives, and even if we overlook the issues that are raised by treating derivational morphemes as heads (Arcodia 2012), we run up against counterexamples to the intended constraint. For example, the initial morphemes in *uphill, downstream, inside, online* would seem to turn roots into adjectives in a very parallel way, and yet these forms do not avoid prenominal attributive uses (e.g., *an uphill battle, the downstream fish, the inside lining, the online chat)*.

Moreover, the "statistical" part of statistical preemption refers to the idea that learners are able to make stronger inferences about restrictions on a word's use when they have more data. For instance, a child who has heard an adjective used once attributively and once predicatively should not be expected to entertain any strong intuitions that it might be unacceptable in one position or the other. In contrast, hearing the adjective used once attributively and 20 times predicatively (in contexts in which an attributive use would normally be appropriate) would support a constraint against attributive use, and hearing it used once attributively and 2,000 times predicatively would support an even stronger constraint. Thus it is important that an appropriately sized corpus is investigated for any meaningful inferences to be drawn.

The corpus considered by Y15 was quite small, since children of high SES parents witness roughly ten times the amount of input considered—45 million words—by the time they are four years old (Hart & Risley 1995). Even children who receive only the most impoverished input witness 13 million words during their first four years. Moreover, Y15's corpus analysis considered only relative clause uses of adjectives, likely because this was the type of preemptive exposure that was provided to learners in Boyd & Goldberg (2011). As noted above however, any predicative use is potentially relevant.

5.1. THE RESTRICTION ON *A*-ADJECTIVES IS LEARNED LATE
Nonetheless, we can concur with Y15 that three year olds are not likely to have witnessed or absorbed sufficient evidence to reliably constrain their use of *a*-adjectives. The children may have only witnessed *a*-adjectives used predicatively, but they are unlikely to have witnessed a sufficient number of instances of these *a*-adjectives in contexts where attributive use would have been at least as appropriate. Moreover, statistical preemption requires that learners are able to implicitly recognize when one form might have been appropriate but another was consistently witnessed.

We expect statistical preemption to be a slow and gradual process that can take years if the relevant data is sparse, as it is in this case. In fact, there is evidence that the restriction on *a*-adjectives is actually learned quite late. Hao (2015) ran a version of Boyd and Goldberg's (2011) experiment with children who ranged from 6-17 years old. She found that children younger than ten years old did *not* use *a*-adjectives (*asleep, alive, afloat* and *afraid*) attributively significantly less than other adjectives (*floating, frightened, living,* and *sleepy*), whereas older children displayed a more adult-like pattern. These results are displayed in Figure 1.
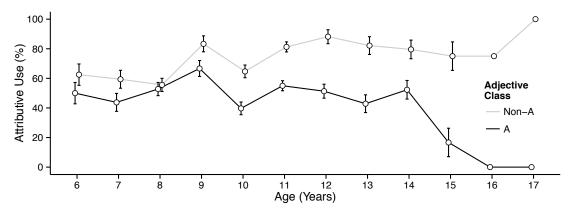

Figure 1. Results based on Hao (2015) for the production task used in Boyd & Goldberg (2011) with children aged 6-17.

Thus children appear to require a least an order of magnitude more evidence than the corpus considered in Y15. Note, too, that Y15's observation that three-year-old children did not use *a*-adjectives attributively does not entail that they actively avoided attributive use. These young children may have simply used *a*-adjectives in contexts in which they had previously witnessed them being used (e.g., Boyd & Goldberg 2012). Hao's (2015) evidence that the restriction on *a*-adjectives is not evident until children are at least 10 years old is consistent with the idea that the restriction is learned slowly and gradually, as children repeatedly witness the relatively rare contexts in which *a*-adjectives are used predicatively when an attributive use would otherwise have been appropriate.

5.2. CATEGORIZATION AT DIFFERENT LEVELS OF GENERALIZATION
Recent experimental and modeling work indicates that both children and adults track distributional information at multiple levels of generalization (Perfors et al. 2010, Wonnacott et al. 2008, Wonnacott 2011). For adjectives, this means that learners gradually learn how individual adjectives are used, while at the same time they learn how adjectives are used more generally. For example, a high level generalization—supported by the majority of adjectives—is that English adjectives can appear either attributively or predicatively. This level of generalization is relevant because it explains the existence of productive, creative uses of adjectives.

Intermediate levels of generalizations also exist, and the size of an identifiable cluster of cases that behave similarly correlates with the strength with which new instances are associated with the generalization (e.g., Granger 1990; Dąbrowska 2008). Thus, once a category of *a*-adjectives is implicitly recognized, it facilitates the extension of the category to tokens that are judged to be similar, in a probabilistic way (see Boyd & Goldberg 2011; Experiment 1).

Thus, once we consider the appropriate type and amount of input and appreciate the fact that witnessed items cluster in similarity space, the handful of adjectives that only occurred in a relative clauses (e.g., *careful*), or never did (e.g., *green*) in a 4.3 million word corpus should not be alarming. Since *green* can appear predicatively in English (*the frog is green*), it is highly unlikely that the attributive uses in the corpus of child-directed speech occurred in contexts in which a predicative use would have been otherwise preferable. Therefore, it is unlikely that the attributive uses served as preemptive evidence. Moreover, g*reen* is a member of a large class of color terms, unrestricted phonologically or morphologically, which all freely occur both attributively and predicatively. The existence of these analogous adjectives supports the use of *green* attributively via categorization. Finally, there is little doubt that even young children will witness *green* repeatedly used predicatively soon enough (e.g., *Kermit is green*).

Careful is used attributively in certain expressions that are primarily spoken by adults to adults (*careful attention/consideration/review*), and it remains to be seen at what age children readily use it attributively themselves. But parallel to what is true for *green*, since *careful* can be used attributively, it is unlikely that the predictive uses in the corpus served as clearly preemptive. Instead, they likely were motivated independently. For example, children no doubt hear the phrase *be careful* quite often, but this should not count as preemptive evidence, because an attributive paraphrase is unnecessarily cumbersome (*Be a careful child*!) and therefore would not be as appropriate.

Thus an adjective (or predicate) may, under the right circumstances, be used creatively in a target construction, if the adjective is novel, or even if it has only been witnessed appearing in a

construction that potentially competes with the target construction (e.g., Bybee 2010; Perek & Goldberg, to appear; Suttle & Goldberg 2011; Wonnacott et al. 2008). The factors that determine whether an adjective is used productively in this way include (a) whether and how frequently the target construction has been witnessed occurring with a range of similar words (categorization), and (b) how frequently the word has systematically been witnessed in the competing construction when the target construction would otherwise have been appropriate (statistical preemption). Experimental manipulations allow us to control these factors, and much more work is needed in order to adequately quantify them.

5.3 Structural priming does not explain the experimental results

Y15 suggests that speakers avoided using novel *a*-adjectives attributively in Boyd & Goldberg (2011)'s second experiment due to structural priming. However, the same number of attributive and predicative exemplars was witnessed in each of the three experiments during exposure, but novel *a*-adjectives only *dramatically* resisted attributive use in Experiment 2. Moreover, recall that during the production task of each experiment, fillers that reliably elicited attributive and non-attributive uses were interspersed (and counterbalanced) before each target adjective, in order to guard against priming effects.

Perhaps Y15 intended that specifically witnessing two novel *a*-adjectives used in relative clauses during the exposure in Experiment 2 is what primed their use in relative clauses during the production task. There is in fact, a recognized "lexical boost" in structural priming when specific words are used in both the prime and target. However, this boost typically does not last beyond a single filler trial (Chang et al. 2006, Hartsuiker et al. 2008, Kaschak & Borreggine 2008), and thus cannot account for the increased use of relative clauses during the production task, since several trials and a short break intervened. Moreover, lexically specific priming would not explain why two other novel *a*-adjectives that were *not* witnessed in a preemptive context during exposure nonetheless resisted attributive use just as much as the two novel *a*-adjectives that had been witnessed. To summarize, is not clear how structural priming can provide an adequate fit to the three experiments reported in Boyd & Goldberg (2011).

On the other hand, the proposal that statistical preemption works in conjunction with categorization argues that (1) learners use statistical preemption to infer that novel *a*-adjectives are dispreferred in attributive position, (2) this restriction generalizes to unwitnessed novel *a*-adjectives via an implicit *a*-adjective category, and (3) that acquisition of the restriction on novel *a*-adjectives in Experiment 2 was because the adult participants only needed to implicitly recognize the novel *a*-adjectives as members of the *a*-adjective category—a category that has been learned on the basis of statistics that have been built up over many years.

6 CONCLUSION

Language can be used creatively by combining existing constructions in new ways. But there exist certain restrictions on such combinations that do not follow from functional factors (e.g., clashes of semantics or information structure), nor from phonological constraints, nor from a lack of available constructions to license the combination. These "embarrassing exceptions" as Baker had called them (1979: 547) have challenged linguists for quite a long time (Ambridge et al. 2008, Bowerman 1988, Braine 1971, Brown & Hanlon 1970, Gennari & MacDonald 2008, Goldberg 1995, Lakoff 1970, Pinker 1989). The restriction against using certain adjectives attributively—*alive, afraid, alone, aware, afloat, asleep*—is such a case of an embarrassing

restriction, as neither the semantics nor the phonology of these words explains the restriction, and of course the vast majority of adjectives do readily appear attributively.

We have seen that the proposed evidence that *a*-adjectives pattern like locative particles or prepositional phrases and unlike typical adjectives in terms of occurring with *right*-type adverbs is lacking, even in the very large, 450 million word COCA corpus. Moreover, since locative particles *can* appear attributively (Table 3), relating *a*-adjectives to locative particles would not explain the restriction.

The available evidence, including experimental findings (Boyd & Goldberg 2011), and the fact that the restriction is learned quite late (Hao 2015), is consistent with the idea that the restriction emerges from error-driven learning in the form of statistical preemption. That is, there are certain contexts when attributive adjective uses are expected. If predicative uses, whether relative clauses or simple predications, are consistently witnessed instead in such contexts, speakers will eventually learn to avoid the attributive use. We assume this is a slow process since it takes time for the relevant evidence to accrue. At the same time, once a restriction on an identifiable subcategory emerges, learners can and do assimilate new members to the restricted category. Thus, combination of categories at different levels of generalization, and a recognition that statistical preemption is context dependent serves to ensure that our knowledge of language allows us to be creative, but not unconstrained.

Acknowledgements

References

AMBRIDGE, BEN, J.M. PINE, C.F. ROWLAND & F. CHANG. 2012. The roles of verb semantics, entrenchment and morphophonology in the retreat from dative argument structure overgeneralization errors. *Language* 88.45-81.

AMBRIDGE, BEN, JULIAN M. PINE, CAROLINE F. ROWLAND & CHRIS R. YOUNG. 2008. The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition* 106.87-129.

ARONOFF, MARK. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

ARCODIA, G.F. (2012). Constructions and headedness in derivation and compounding. *Morphology*, *22*(3), 365–397. doi:10.1007/s11525-011-9189-2

BAKER, CARL L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10.533-81.

Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic* (Vol. 8). John Benjamins Publishing.

BOLINGER, DWIGHT. 1971. *The Phrasal Verb in English*. Cambridge, MA: Harvard University Press.

BOWERMAN, MELISSA. 1988. The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? *Explaining language universals*, ed. by J. Hawkins, 73-101. Oxford: Blackwell.

BOYD, JEREMY K. 2014. Statistical pre-emption. *Encyclopedia of Language Development*, ed. by P.J. Brooks & V. Kempe, 608-10. Thousand Oaks, CA: SAGE Publications.

BOYD, JEREMY K. & ADELE E. GOLDBERG. 2011. Learning what *not* to say: The role of statistical preemption and categorization in *a*-adjective production. *Language* 87.55-83.

BOYD, JEREMY K. & ADELE E. GOLDBERG. 2012. Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners. *Journal of child language*, *39*(3), 457.

BRAINE, MARTIN D.S. 1971. On two types of models of the internalization of grammars. *The ontogenesis of grammar*, ed. by D.I. Slobin, 153-86. New York: Academic Press.

BROOKS, P.J. & MICHAEL TOMASELLO. 1999. How children constrain their argument structure constructions. *Language* 75.720-38.

BROOKS, P.J. & O. ZIZAK. 2002. Does preemption help children learn verb transitivity? *Journal of Child Language* 29.759-81.

BROWN, ROGER & CAMILLE HANLON. 1970. Derivational complexity and order of acquisition in child speech. *Cognition and the development of language*, ed. by J.R. Hayes, 11-53. New York: John Wiley and Sons.

BYBEE, JOAN. 2010. *Language, usage and cognition*. Cambridge University Press, 2010.

BRUENING, BENJAMIN. 2011a. A-Adjectives are PPs, not Adjectives. *http://lingcomm.blogspot.com/2011/03/adjectives-are-pps-not-adjectives.html*

BRUENING, BENJAMIN. 2011b.

CHANG, FRANKLIN, GARY S DELL & KATHRYN BOCK. 2006. Becoming syntactic. *Psychological Review* 113.234-72.

COPPOCK, ELIZABETH. 2008. The logical and empirical foundations of Baker's Paradox. Stanford, CA: Stanford University.

CHOMSKY, NOAM. 1957. *Syntactic Structures*. Walter de Gruyter.

CLARK, EVE V. 1987. The principle of contrast: A constraint on language acquisition. Mechanisms of language acquisition, ed. by Brian MacWhinney, 1–33. Hillsdale, NJ: Erlbaum.

CROFT, WILLIAM. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.

DAVIES, MARK. 2008. The Corpus of Contemporary American English: 450 million words, 1990-present. Available at http://corpus.byu.edu/coca.

DĄBROWSKA, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language*, *58*(4), 931-951.

GENNARI, S.P. & M.C. MACDONALD. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58.161-87.

GOLDBERG, ADELE E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

—. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

—. 2011a. Are *a*-adjectives like *afraid* prepositional phrases underlyingly and does it matter from a learnability perspective? https://adele.princeton.edu/pubsbytopic/#Statistical.

—. 2011b. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22.131-53.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of memory and language*, *29*(2), 228-244.

HAO, JESSICA. 2015. Abstraction versus restriction in syntactic learning : An examination of children ' s acquisition of the a -adjective restriction. Senior thesis. Princeton University.

HART, B., & RISLEY, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

HARTSUIKER, R.J., S. BERNOLET, S.. SCHOONBAERT, S. SPEYBROECK & D. VANDERELST. 2008. Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language* 58.214-38.

HARTSUIKER, R.J. & C. WESTENBERG. 2000. Word order priming in written and spoken sentence production. *Cognition* 75.27-39.

HAY, JB AND RH BAAYEN. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences*, *9*(7), 342–348.

HUDDLESTON, RODNEY & GEOFFREY K. PULLUM. 2001. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

JACKENDOFF, RAY. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT press.

KASCHAK, MICHAEL P & KRISTIN L BORREGGINE. 2008. Is long-term structural priming affected by patterns of experience with individual verbs? *Journal of Memory and Language* 58.862-78.

LAKOFF, G. 1970. *Irregularity in Syntax*. New York: Holt, Rinehart & Winston.

LARSON, RICHARD K & FRANC MARUŠIČ. 2004. On indefinite pronoun structures with APs: Reply to Kishimoto. *Linguistic Inquiry* 35.268-87.

LONG, RALPH B. 1969. *The sentence and its parts. University of Chicago Press.*

MACWHINNEY, BRIAN. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.

PEARL, LISA AND JON SPROUSE. 2013. Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. Language Acquisition 20.23–68.

PEREK, F, AE GOLDBERG. TO appear. Generalizing beyond the input: the functions of the constructions matter. *Journal of Memory and Language.*

PERFORS, A, J.B TENENBAUM & E WONNACOTT. 2010. Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language* 37.607-42.

PINKER, S. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

POSER, WILLIAM J. 1992. Blocking of phrasal constructions by lexical items. *Lexical Matters*, ed. by I.A. Sag & A. Szabolcsi, 111-30. Stanford, CA: CSLI.

RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: Current research and theory*.

ROBENALT, CLARICE & ADELE E. GOLDBERG. to appear. Judgment evidence for statistical preemption: It is relatively better to *vanish* than to *disappear* a rabbit, but a lifeguard can equally well *backstroke* or *swim* children to shore. *Cognitive Linguistics*.

SALKOFF, MORRIS. 1983. Bees are swarming in the garden: A systematic synchronic study of productivity. *Language* 59.288-346.

SCHEEPERS, C. 2003. Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition* 89.179-205.

SUTTLE, LAURA & ADELE E. GOLDBERG. (2011). The partial productivity of constructions as induction. *Linguistics*, *49*(6), 1237-1269.

WILLIAMS, EDWIN. 1981. On the notions ''lexically related'' and ''head of a word.''

Linguistic Inquiry 12: 245–274

WONNACOTT, E, E.L NEWPORT & M.K TANENHAUS. 2008. Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology* 56.165-209.

WONNACOTT, E. 2011. Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language* 65.1-14.

YANG, CHARLES. 2015. Negative knowledge from positive evidence. *Language*.