John Newman
ORCID: 0000-0002-9978-0689
University of Alberta
Monash University

Singular and plural preferences among adjectival collocates of Cat and Dog

ABSTRACT

The singular vs. plural distinction in English count nouns is not usually considered as being of any real consequence to the choice of adjectives accompanying these nouns. This study questions this assumption and explores the pre-nominal adjectives occurring with *cat/cats* and *dog/dogs* with a view to identifying the main patterns of co-occurrence with singular forms vs. plural forms. Attributive adjectives occurring before *cat(s)* and *dog(s)* were investigated, relying on a corpus of contemporary American fiction. Applying Distinctive Collexeme Analysis to the corpus results, it was found that coherent groups of adjectives occurred preferentially with the singular or plural of both words. Colour adjectives and evaluative adjectives like *good*, for example, occurred preferentially with the singular forms, while adjectives such as *stray*, *wild*, and *feral* occurred preferentially with plural forms. The usage differences observed in the data can be motivated by reference to a folk model of the world in which animals take their place as house pets, wild animals, or somewhere in between.

1. Introduction

The present study is an investigation into less-studied linguistic patterns associated with singular and plural forms of English count nouns. Some patterns associated with this distinction are obvious and predictable, for example, patterns of verb agreement with syntactic subjects, determiner choice (*this/these*, *that/those*), quantifiers and numerals (*one/many*, *second/two*).
https://doi.org/10.32058/LAMICUS-2021-001

With such patterns, there is no need to resort to a corpus study in order to understand the basic patterns at work. The adjective + noun construction, however, is not an obvious site for studying patterns that might correlate with a singular vs. plural noun distinction. We find, after all, *old dog* and *old dogs*, *big dog* and *big dogs*, etc. The default expectation would be to assume that there is no difference between the classes of adjectives that can occur with singular vs. plural nouns. It is my intention in the present study to explore the corpus-based patterning of adjectives with singular vs. plural nouns, using CAT and DOG as the target nouns.[1] My aim is to show that (probabilistic) differences not only exist, but also that the findings raise interesting questions about how these differences reflect our perceptions of cats and dogs. While my approach is essentially quantitative, the intricacies and variability of adjectival uses and meanings require unusually close attention to the actual examples of usage and the concepts of precision and recall of the results will be very relevant.

The study is structured as follows: Section 2 provides some background to the study and rationale for the choice of CAT and DOG as the nouns of interest; Section 3 describes the data and how it was prepared for analysis; Section 4 is the quantitative analysis; Section 5 is a discussion of the results and the differential semantics associated with singular and plural forms; Section 6 is the conclusion.

2. BACKGROUND

The impetus for the present paper stems from the question of whether the singular and plural forms of count nouns might pattern differently in terms of adjectival collocates. In part, the question arises as a follow-up question in response to Langacker's suggestions about the singular vs. plural noun types in Langacker (1991: 74-81). The main idea in that discussion is that singular and plural categories of nouns are to be strongly differentiated in terms of the *types* they represent: "[...] it is essential to realize that a singular noun and its corresponding plural represent distinct categories and that what counts as an instance is consequently very different in the two cases" (Langacker 1991: 78). To use Langacker's own example, *the seven pebbles* is said to profile one instance of the *pebbles* type, not seven instances of *pebble*. *Pebble* and *pebbles* each has its own distinct type – a "discrete entity" type for the singular vs. a "replicate mass" type for the plural. If one accepts that the singular and plural forms of a (count) noun such as *dog* are quite distinct conceptual categories, then one may rightly inquire as to what properties correlate with this distinction.

---

1   Here and throughout, small caps will be used to refer to a lemma and italics to refer to a word form.

The inflectional differences between singular and plural nouns have not played a prominent role in corpus linguistic research in the last 30 years, even if such an interest is fully in keeping with the corpus-based practices of a pioneer in the field, John Sinclair (cf. Sinclair 1991: 53-64, 154). Textbooks on corpus linguistics will include some discussion of lemma vs. inflected word forms but typically remain agnostic about which level is to be used in any particular analysis. Hunston (2002: 52-56), for example, illustrates methods of interpreting concordance lines by choosing opportunistically either the lemma as the unit of interest (as in the case of SUGGESTION) or the inflected word form (as in the case of *point*) without drawing attention to or problematizing the difference. McEnery, Xiao & Tono (2006: 36) describe the benefit of working at the lemma level as a convenient way of collecting the relevant inflected forms without commenting on whether investigating distributional or collocational differences between the inflected forms might be of any interest.

A number of studies of collocational patterning with the inflected forms of different parts of speech in English have revealed interesting results that point to the value of investigating collocational behaviour at the level of the inflected word form. Newman & Rice (2004), for example, investigates collocational differences associated with the inflectional differences between the *-ing* and past tense forms in the pairs sitting and.../*sat and...,standing and.../ stood and*, *lying and.../lay and...* and finds very different semantic properties associated with the conjoined verb in the *-ing* and past tense forms. A study relevant to the singular vs. plural distinction in English is Katz & Zamparelli (2012), a computational linguistic study of collocates (in the form of 'vector spaces') of 2,114 nouns in a POS-tagged corpus that occur with both singular and plural POS-tags forms. The window of the context for collecting the collocates of each noun was the four previous and four subsequent content words. Although the focus of their study is mass nouns, Katz & Zamparelli (2012: 378) report on other nouns as follows: "We have also uncovered a wide range of other cases in which the singular and plural forms of a noun have different distributions, itself an interesting result (and one frequently ignored by vector space models). Another area where there has been some discussion of distinctive patterning with singular vs. plural nouns in English is research on metaphor (cf. Deignan 2006: 114-121). The research that is most directly relevant to the present study, however, is the investigation into semantic preferences associated with singular *child* vs. plural *children* in American fiction (Newman 2021). Using similar methods to those employed in the present study, Newman explored the experiential motivations for the preferences for different collocates associated with *child* vs. *children*, focusing on three different construction types: adjectives in the attributive position,

nouns grammatically possessed by *child's* and *children's*, and present participial forms of verbs in the position immediately to the right of *child/children*. The study revealed unexpected differences between singularity and plurality: psychological attributes vs. non-psychological attributes; expressive, empathetic stance vs. expressively neutral stance; alienable vs. inalienable possession; motion vs. stationary predicates.

The decision to investigate adjectival patterning specifically with CAT and DOG was influenced by two main considerations. Firstly, there is a human dimension to the world of cats and dogs, suggesting a rich set of attributive adjectives accompanying the nouns CAT and DOG in fictional writing. Obviously, cats and dogs are animals that can interact closely with humans and attitudes towards cats and dogs have been a particular focus of interest among researchers exploring human-animal interaction. Selby & Rhoades (1981), for example, investigated the role of various factors that influence the attitudes towards cats and dogs as companion animals, e.g., emotional needs of pet owners, whether the pet provides protection and security, women's attitudes vs. men's attitudes to their pets, etc. Blouin (2013) introduced the idea of a "humanistic" orientation that humans may adopt towards pets in general, though the study focused specifically on dogs. In the humanistic orientation, pet owners view their pets as surrogate humans and enjoy affective benefits from their close attachments to their pets. In light of the empathy and emotional attachment that humans can feel towards cats and dogs, one might well expect that the words CAT and DOG, as used in reference to these four-legged animals, would occur with a variety of adjectival descriptors reflecting human perspectives on these animals (in contrast to inanimate entities such as DRAWING-PIN, BOOKMARK, HYPOTENUSE, etc. or more abstract entities such as LOGICAL OPERATOR, READINESS, AREA, etc.). This human dimension to the life of (most) cats and dogs suggests that we might encounter a greater range of adjectives than would be the case studying, say, words for fire ants, squid, tree snakes. Secondly, the CAT and DOG words are relatively frequent in usage, leading to a greater possibility that the corpus-based results will be fairly robust.

3. Data

3.1. The corpus

The corpus used for this study was the Fiction sub-corpus of the online Corpus of Contemporary American English (COCA, Davies 2008-). At the time of this study, the Fiction component of COCA consisted of approximately 120 million words and included sampled American fiction writing from

1990-2019. The online site describes the contents of the Fiction component as "short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts".

The decision to restrict attention to just a single genre is a consequence of acknowledging very real differences that exist between the main genres in communication goals, style and content of communication, and characteristics of the writer/speaker and addressee. In broad terms, fiction writing reflects the imaginative skills of authors in consciously creating interesting and engaging fictional scenarios and events. Fiction has been described in more sophisticated terms as offering models of the cognitive simulation of our social worlds (Oatley 1999; Mar & Oatley 2008). Fiction in any case presents a freer, imaginatively enriched use of language compared with other genres, which are less about simulations of our social worlds and more about the real world. It is a genre that is especially relevant when the researcher has an interest in exploring the descriptive potential of adjectives. One may expect a substantive amount of descriptive text in fiction writing to accompany nouns like CAT and DOG as a way of adding interest to a story, person, or object. More specifically, and related directly to the purposes of the present study, it has been observed that the fiction genre makes greater use of attributive adjectives than do news and conversation genres. The *Longman Grammar of Spoken and Written English* (Biber et al. 1999) draws attention to this aspect of the use of attributive adjectives in fiction writing, as quoted in (1).

(1)   Although attributive adjectives functioning as descriptors are used in all four registers, fiction employs a greater number of these than the other registers. The semantic domains of size (e.g. *big*, *little*), colour, and evaluation (e.g. *good*, *nice*) are most important in this register. (Biber et al. 1999: 511)

One can expect, then, that the choice of the Fiction sub-corpus of COCA will increase the likelihood of obtaining solid results from a study of adjectival use with cat and dog in fiction.

An additional consideration when it comes to choosing a suitable corpus for the present study is the fact that a close reading of the contexts of use of *cat(s)* and *dog(s)* is often necessary to identify instances that need to be excluded from further study, as explained in the following section. This means that there has to be some manual editing of search results and the search results from the Fiction sub-corpus allow this editing to be carried out more reasonably than would be the case for all of COCA.

3.2. Data collection and accuracy of search results

The procedure to collect the adjectives was the same for both *dog(s)* and *cat(s)*, relying initially on the adjective part-of-speech tag in the L1 position to identify attributive adjectives. While, in general, some check on accuracy of search results is desirable, it was especially important to do so in the case of this study. Recall two main targets of interest in this study: (i) the focus on *dog(s)* and *cat(s)* as used to refer to four-legged animals, and (ii) adjectival preferences associated with the choice of singular vs. plural nouns. With these two goals in mind, considerable editing of the search results to improve both precision and recall was needed and resulted in the removal of hundreds of hits from the original search returns. I illustrate the kinds of issues that arose through the example of *dog(s)*, though similar examples could be cited for *cat(s)*.

The most glaring example of a false positive in the results is *hot dog(s)*, used always to refer to the sausage-and-bun food item. *Hot dog* is in fact the most frequent type in the singular (409 hits), just as *hot dogs* is in the plural (433 hits). All instances of *hot dog(s)* were removed from the results. While it was easy to make a decision about the exclusion of *hot dog(s)*, identifying the references to four-legged animals, as opposed to humans, was not always so straightforward. This was especially so in the use of *cat(s)* and *dog(s)* in names of characters in movie scripts and in the fantasy sub-genres. Other examples of false positives are shown in (2). These all involve invariant singular or plural forms where there is no freedom to choose either singular or plural. (2a) illustrates the category of singular or "bare" *dog* use in noun-noun compounds where there is no singular-plural alternation allowed. (2b) illustrates the use of *dog(s)* as part of a proper noun such as the name of a person, pub, festival, story, etc. where the form is fixed. (2c) illustrates the use of *dog(s)* in relatively fixed idiomatic expressions where the intended reference is to humans, not four-legged animals. Finally, *like* occurring before a noun is mistagged as an adjective throughout the corpus and all instances of *like dogs* were excluded from the study.

(2)   (a) *dog* as part of a compound with a following noun:
          white dog collar, a good dog name, small dog turds, big dog pile, the mad dog woman, mad dog sickness, other dog teams, etc.

     (b) *dog* as part of a proper name:
          The Old Dog Mission, the Black Dog Inn, the Black Dog of depression, the Tomb of the Black Dog, Black Dog (song), *Mr. Big Dog, Mad Dog* (= fortified wine), *Mad Dog Rinaldi, "Mad Dog" Coll* (nickname), *Dead Dog Creek, the Stray Dog* (= inn), *the Sacred White Dog* (= festival), *Barking Dog Road, Ugly Little Dog* (nickname), *Sleeping Dogs* (= title of fiction), etc.

(c)   *dog/dogs* in idiomatic expressions referring to humans:
       *sly old dog, you can't teach an old dog new tricks, you can't teach
       an old dog new kindness, can't learn an old dog new tricks, let sleeping
       dogs lie,* etc.

Recall was improved by adding to the initial search results a number of cases involving mistakenly tagged adjectives. Two important cases of adjectives that do not show up in the initial results involve *little* and *barking.* The word *little* may be tagged in the relevant tagging system (CLAWS 7) as either an adjective or as a "singular after-determiner". The latter occurs in examples such *as most people have **little** concept of marriage today, so much to do, so **little** time, do as **little** as possible* and contrasts with a "plural after-determiner" such as ***several*** in *we had several things in common, one study of **several**,* etc. All the instances of *little* in *little* dog in the corpus should be adjectives, though some instances are mistagged as the singular after-determiner and hence did not appear in the initial results. In the case of *barking,* some instances of its use with *dog(s)* are mistagged as verb forms when they should be considered adjectival (resulting in 51 instances instead of initial 18). One instance of *stray* was mistagged as verbal in *I played to stray dogs and old ladies and babies in bassinet.*

As a result of this editing of the search results, the initial counts were adjusted from a tally of 4,579 instances of adjective + *dog* to 4,090, and from 2,286 instances of adjective + *dogs* to 1,383. The initial tally of 2,286 instances of adjective + *cat* was reduced to 2,159 and the tally of 682 for adjective + *cats* was reduced to 652.

4. Analysis

4.1. Critical adjectives

The ten most frequently occurring adjectives with each of *dog(s) and cat(s)* were identified and these became the focus of analysis. Choosing to single out the ten most frequent adjectives with each form is not simply a matter of prioritizing the most frequent members of a category, although it does do that. An important consideration about the top ten adjectives with each noun is that their frequencies represent the critical, highest-ranked values of a Zipfian-like distribution, i.e., a distribution characterized by a very small number of items with high frequencies and a very large number of items with low frequencies, in particular, frequencies of one and two. In order to better appreciate the contribution of the top ten adjectives, consider Figure 1, showing the cumulative percentages of the 30 most frequent adjectives with each of the four nouns. "Cumulative percentage" of an item refers here

to acumulative percentage of all the adjective tokens with either *dog, dogs, cat,* or *cats*, based on the total number of tokens of the adjective + noun. The shaded box in Figure 1encloses the data pertaining to the 10 highest ranked items and corresponds roughly to the steep slope characteristic of a Zipfian curve. In the case of adjective + *dog,* these top ten adjectives account for 31% of all the adjective tokens; in the case of adjective + *dogs,* the top ten adjectives account for 40% of all the adjectives. The top ten adjectives, then, are critical in their disproportionate contribution to the usage of the construction.
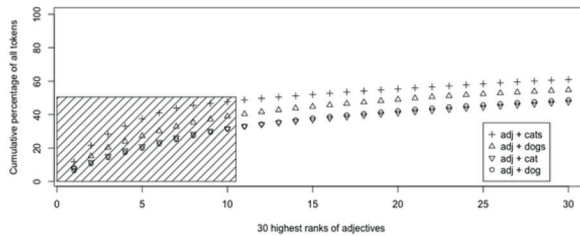

Figure 1. Cumulative plots for the 30 most frequent adjective types occurring with dog(s) and cat(s)

The plots show the cumulative proportion of adjective tokens out of the total number of adjectives occurring with the noun. For example, the ten highest ranked adjectives with *cats* account for approximately 50% of all adjective tokens with *cats*. The shaded area encloses the data captured by the top ten adjective types for each noun

The ten most frequent adjectives with *dog(s)* and *cat(s)* (in the sense of four-legged animals) are shown in Table 1 and Table 2, after necessary editing of the search results per Section 3.2. It is these adjectives that will be the specific focus of interest in the remainder of this study.

Table 1. Top ten attributive adjectives by frequency for dog and dogs after editing

Rank	dog		dogs	
	Adjective	Freq	Adjective	Freq
1	little	263	other	118
2	old	172	stray	99
3	black	150	wild	73
4	good	124	barking	51
5	(an)other	107	big	46
6	big	104	little	38
7	small	102	feral	35
8	stray	93	small	27
9	dead	85	dead	26
10	white	82	rabid	20

Rank	cat		cats	
	Adjective	Freq	Adjective	Freq
1	black	169	feral	77
2	orange	82	stray	65
3	white	79	other	42
4	big	70	black	32
5	gray	56	big	29
6	dead	55	dead	23
7	stray	48	wild	19
8	old	37	Siamese	10
9	little	32	old	8
10	Siamese	30	great, large	7

4.2. Collostructional analysis

Alongside the identification of critical adjectives in terms of frequency of occurrence, there is a need to systematically identify whether it is *dog* or *dogs* that each adjective is more attracted to, bearing in mind the different total frequencies of adjective + *dog* and adjective + *dogs*. After editing of search results as explained in Section 3.2, there remained 4,090 instances of adjective + *dog* in the corpus and 1,383 instances of adjective + *dogs*, i.e., almost three times more adjectives with the singular forms than with the plural forms. This is the "expected" ratio of an adjective occurring with *dog* compared with *dogs*. *Rabid*, for example, occurs more frequently with *dog* (27) than *dogs* (20), but it still occurs less often than what one expects (only 1.35 times more frequently with *dog* than *dogs*), so *rabid* is by this reasoning more attracted to *dogs* than *dog*. The type of statistical analysis that targets this kind of data, where the focus is on the competition for collocates between two similar constructions (here, an adjective + *dog* construction and an adjective + *dogs* construction) is Distinctive Collexeme Analysis (DCA, cf. Gries and Stefanowitsch 2004, Gries 2012). "Collexemes" refers to the items occurring in a specific slot within a construction. DCA calculates the expected frequencies of the collexemes in the two constructions, the direction of the attraction (here, to *dog* or *dogs*), and returns a measure of this degree of attraction. In the present case, the collostructional strength is based on the Log Likelihood measure. The DCA was carried out using the Collostructions package in R (R Core Team 2019, Flach 2017).

The results of the collostructional analysis of adjectives occurring with *dog(s)* and *cat(s)* are shown in Tables 3 and 4. In these tables, "Assoc.noun" refers to the noun that an adjective is more attracted to, "Coll.strength" refers to the collocational strength, and "SIGNIF" refers to statistical significance. The significance level is indicated by asterisks thus: ***** = significant at $p <$ .00001, **** = significant at $p <$ .0001, *** at $p <$ .001, ** at $p <$ .01, * at $p <$ .05, ns = not significant. The adjectives that show statistically significant attraction have been highlighted.

Table 3. Collostructional analysis results of the adjective + dog(s) construction

	Collexeme	Observed Freq with *dog*	Expected Freq with *dog*	Observed Freq with *dogs*	Expected Freq with *dogs*	Assoc. noun	Coll. Strength	SIGNIF
1	*old*	172	134.5	8	45.5	*dog*	58.15	*****
2	*black*	150	122.6	14	41.4	*dog*	30.98	*****
3	*little*	263	224.9	38	76.1	*dog*	30.96	*****
4	*good*	124	103.1	14	34.9	*dog*	20.59	*****
5	*white*	82	70.2	12	23.8	*dog*	9.12	**
6	*small*	102	96.4	27	32.6	*dog*	1.37	ns
7	*dead*	85	83	26	28	*dog*	0.21	ns
8	*big*	104	112.1	46	37.9	*dogs*	2.28	ns
9	*rabid*	27	35.1	20	11.9	*dogs*	6.71	**
10	*barking*	36	65	51	22	*dogs*	44.10	*****
11	*feral*	7	31.4	35	10.6	*dogs*	63.10	*****
12	*stray*	93	143.5	99	48.5	*dogs*	63.14	*****
13	*(an) other*	107	168.1	118	56.9	*dogs*	79.40	*****
14	*wild*	23	71.7	73	24.3	*dogs*	110.88	*****

Table 4. Collostructional analysis results of the adjective + cat(s) construction

	Collexeme	Observed Freq with *cat*	Expected Freq with *cat*	Observed Freq with *cats*	Expected Freq with *cats*	Assoc. noun	Coll. Strength	SIGNIF
1	*orange*	82	64.5	2	19.5	*cat*	30.85	*****
2	*gray*	56	44.5	2	13.5	*cat*	18.27	****
3	*white*	79	65.3	6	19.7	*cat*	16.24	****
4	*little*	32	26.1	2	7.9	*cat*	7.59	**
5	*black*	169	154.4	32	46.6	*cat*	6.96	**
6	*old*	37	34.6	8	10.4	*cat*	0.80	ns
7	*Siamese*	30	30.7	10	9.3	*cats*	0.07	ns

8	*great*	19	20	7	6	*cats*	0.20	ns
9	*large*	14	16.1	7	4.9	*cats*	1.12	ns
10	*dead*	55	59.9	23	18.1	*cats*	1.69	ns
11	*big*	70	76	29	23	*cats*	2.03	ns
12	*(an) other*	54	73.7	42	22.3	*cats*	20.48	*****
13	*wild*	6	19.2	19	5.8	*cats*	31.49	*****
14	*stray*	48	86.8	65	26.2	*cats*	64.39	*****
15	*feral*	19	73.7	77	22.3	*cats*	145.87	*****

4.3. NASTY LITTLE DOG TYPE

While the focus of the present study is on the adjective that occurs immediately before CAT and DOG, it is worth mentioning also the occurrence of the adjective + *little* + noun construction. The "double adjectives" construction is associated above all with the adjective little as the second adjective. There are 51 cases of another adjective used before *little dog*, such as *nasty little dog* (5), *ugly little dog* (2), *scruffy little dog* (2), *nice little dog* (2), *repulsive little dog* (1), *pretty little dog* (1), etc. There are just 7 cases with the plural *little dogs*, e.g., *cute little dogs* (2), *awful-looking little dogs* (1), *nasty little dogs* (1), etc. Given that the overall proportion of singular *little dog* to plural *little dogs* is 263/38, i.e. 6.9 times more frequent, the proportion with the extra adjective (51/7 = 7.3) is close to what one would expect statistically. There are just 5 cases of adjective + *little cat(s)*: *thin little cat, stupid little cat, long-legged little cat, mischievous little cat, blue little cat,* each occurring just once. The OED entry on *little* includes a mention of this construction with little as "implying endearment or appreciation, or tender feeling on the part of the speaker", adding also "coupled with an adjective expressing such feelings, as *pretty little, sweet little,* etc." As the examples show, however, the corpus results are by no means restricted to such positive emotions and include examples such as *nasty little dog, repulsive little dog,* and *awful-looking little dogs,* none of which could be described as a term of endearment.

5. DISCUSSION

The group of adjectives attracted to the plural {*feral, stray, wild, rabid, barking*} refer to animal types that clearly have some significance for humans – humans normally avoid these kinds of animals and are afraid of them in varying degrees. One would not normally make a special effort to interact with such animals beyond removing them or hunting them or even destroying them.

They represent a kind of non-human, disconnected "other" that we keep at a distance, unlike house pets, and this may play a part in our preference to refer to them only as a collective, without the need or desire to individualize them. Apart from the general collectivizing tendency with these categories of animals, feral and wild animals have the further characteristic of living as a pack. This could be viewed as an additional motivation for the association of these adjectives with the plural forms, supported by multiple occurrences of collective phrases in the data such as a pack of feral dogs, *a pack of wild dogs*, *a tribe of feral dogs*, *clans of wild cats*, *herds of wild cats*, etc.

The colour adjectives account for most of the adjectives significantly attracted to the singular forms. The clear preference for the colour adjectives to occur with the singular forms *dog* and *cat* shows that authors are using the colour adjectives to elaborate more on the appearance of an individual dog or cat than a collective of dogs or cats. Authors appear to have little interest in introducing groups of like-coloured dogs or cats into their fiction. The colour of a dog or cat does not generally play any particular part in the story-telling, apart from lending more character and interest to the animal. The exception is the blackness of a cat which does, of course, have special connotations relating to bad luck, something slightly evil, witchcraft, etc. The reference to the colour black, in this case, has direct relevance to the story line, rather than simply being an incidental or trivial descriptive elaboration of the noun. In (3a), *black cat* refers to a Halloween costume, where the blackness of the cat is an essential part of the costume. In (3b), the blackness of the cat is required to help create the image of a witch. (3b) explicitly connects black cats to witches. (3e) It is only *black* that has a special connotation when used with *cat*.

(3) a. Once again, Staci was dressed as a <u>black cat</u> and Randolph had come dressed as Batman.
  b. "All she needs is a black cat and a pointed hat," Charles muttered. "And a broom on the end of that stick."
  c. "C'm on, you don't have to be a scholar to know that witches' familiars are usually <u>black cats</u>, though not always so rotund."

The adjective *little* calls for its own special discussion. In the data under discussion it is attracted to singular *dog* and *cat* in highly significant ways. *Little* is semantically complex, as pointed out above. Both *little* and *small* refer to size, but *little* can be associated with a more emotional (either positive or negative) kind of engagement of the writer or speaker than is found with *small*. The examples in (4) illustrate this nuance. In (4a), the comforting sensation of being cuddled is compared to how a cat can be cuddled by humans. In (4b), expressions such as *wagging its tail* and *trotting along beside her* suggest happiness in a dog and comfortableness with its human companion. (4c)

contrasts the happiness in a little dog's face in stark contrast to the image of animals being eaten by humans. Finally, in (4d), the character is saying warm goodbyes to various humans, animals, and inanimate entities. In all these cases (and there are many others), there is a degree of emotional engagement with the dog which makes the choice of *little*, rather than *small*, seem more appropriate. The lack of any similar emotional nuance in the use of *small* can also be seen in the occurrence of the expressions *the size of a small dog* and *the size of small dogs*, where it is purely the physical size of the dog that is being described. There are no matching examples of any expressions like *the size of a little dog* or *the size of little dogs* in the corpus.

(4) a. She let herself be cuddled, like a <u>little cat</u>,...
     b. All true, even when this attractive, well-built woman about his age, fifty or so, came into breakfast one morning with a <u>little dog</u>, wagging its tail like a sharp stick, trotting along beside her.
     c. But increasingly the thought of eating animals fills me with moral gloom. I look at my <u>little dog</u>'s happy face and feel sick.
     d. Philomela is going to see her sister. Good-bye Nurse. Good-bye Cook. Good-bye <u>little dog</u>. Good-bye birds. Good-bye Papa, good-bye.

The adjective *good* turns out to be significantly attracted to the singular form, but only *dog* (124 instances), not *cat* (16 instances). In part, the preference for *good dog* rather than *good cat* may reflect a difference in how we regard the different behaviours of the animals, with dogs seen as more responsive to, and more worthy of, reward and praise than cats. This attitude is evident in the many cases where a dog is being addressed and praised, as exemplified in (5a-d). A cat can be addressed in a comparable way in the corpus, as shown in (5e), but there are far fewer examples.

(5) a. "That's a <u>good dog</u>." Gingerly Phillip shook the filthy paw and scratched the dog's silky ears.
     b. "Nice dog! <u>Good dog</u>! Nobody's going to hurt you!" the girl called out weakly.
     c. That's a good girl, <u>good dog</u>. Push, sugar lump. Push those puppies out.
     d. She had never been so glad to see them, so grateful for their instant obedience. "<u>Good dogs</u>." She stopped to fondle them.
     e. "There's a <u>good cat</u>."

While *old* is more attracted to both *dog* and *cat*, the attraction only reaches statistical significance with *dog*. Unlike the colour adjectives with *dog*, the oldness of a dog is a topic that is further developed in the text, rather than being some passing, incidental description. The examples in (6) illustrate how the old age of a dog is a topic spread over a larger stretch of discourse.

(6)   a.   And General – he's an <u>old dog</u> – on his last legs, you know. Why should he suffer?
      b.   That's a hell of an <u>old dog</u>. Candy: I had him ever since he was a pup.
      c.   "Plenty of life left in you yet, old dog."

Finally, the adjective *other* has a significant attraction to the plural forms *dogs* and *cats*. Recall that both *another* and *other* were included as a total for *(an) other* occurring with a singular noun, in contrast to just *other* occurring with a plural noun. Even with the larger total for *(an)other* that resulted from this decision, it is still *other* that has a significant attraction to the plural forms. *Other* is used most typically, then, to differentiate a plurality of *dogs* and *cats*, rather than an individual *dog* or *cat*. In other words, *other* points preferentially to *entities* that are not *x*, rather than to *an entity* that is not an *x*.

   The main findings with respect to significant attraction of adjectives to nouns may be summarized as follows:

   (i)   *Dogs* and *cats* attract adjectives relating to wild animals and animals in states or conditions that are dangerous or threatening.
   (ii)   *Dog* and *cat* attract colour adjectives; usually the colour of the animal is incidental and not further referenced in the text (*black cat* is the exception).
   (iii) *Dog* and *cat* attract the adjective *little*, an adjective that can have some emotional component in addition to referring to size.
   (iv)   *Dog* attracts the adjective *good*, with *good dog* playing a particular role in dialogue where a human is addressing a dog.
   (v)   *Dog* attracts the adjective *old*, with the age of the dog often being a topic in the surrounding text.
   (vi)   *Dogs* and *cats* attract the adjective *other*.
   (vii) The adjective + *little* + noun construction is the most common two-adjective sequence with both *dog* and *cat*, used as a term of endearment and non-endearment.

It can be seen that a number of these specific findings involve the classes of attributive adjectives identified by Biber et al. (1999: 511) as being especially

frequent in fiction, i.e., adjectives referring to colour, size, and evaluation. To that extent, the findings from the collostructional analysis are in accord with Biber et al. (1999), but the collostructional analysis goes further by showing the different roles that singular vs. plural nouns play in these patterns.

Most of these findings can be understood as reflecting a difference in how authors conceptualize domesticated vs. wild cats and dogs, as represented in Figure 2. The diagram represents an anthropomorphic view of the world in which the human domain is central and the world of wild animals lies outside of this domain. The human domain in this figure refers to the world most closely associated with human experience and includes family and other household members, family pets, friends, activities associated with a person's residence or place or work, place of recreation etc. Within this domain, there is a preference for the singular forms *dog* and *cat* when used with an attributive adjective. The adjectives used in this domain can highlight various properties of a dog or cat that suggest an emotional attachment to the animal. This pattern of association of singular number and adjectives describing emotional engagement is parallel to the finding in Newman (2021: 17-18) that singular *child*, but not plural *children*, is associated with adjectives of endearment (*dear, precious, poor*). *Good dog* is a familiar form of praise from a human (typically the owner) to a dog, as exemplified in (5) above, found in stretches of dialogue where a human is typically addressing the dog as if the dog is comprehending the words. The preference for human ways of talking to a dog, as opposed to a cat, may reflect the different attitudes of pet owners to dogs and cats, with dogs being viewed as in a more "mutualistic" relationship with their owners while cats are viewed as being in a "parasitic" relationship (Selby & Rhoades 1981: 136). *Little dog/cat* and *old dog* suggest some emotional, affectionate engagement by the narrator or fictional character, as illustrated in (4) and (6). Overall, these uses of adjectives with singular *cat* and *dog* all reflect an engagement with, and interest in, the life and well-being of the dog or cat, even if there are some finer distinctions to be made between cats and dogs. The colour adjectives, for the most part, are incidental to the story telling, but function to add descriptive interest to the individual dog or cat. Outside of the home one finds animals living in the wild, animals that have never been a part of the human domain. This domain one might label as "the wild". Here the preference is for plural dogs and cats, and the adjectives that occur with them indicate their wildness (*feral dogs, stray cats*, etc.), without signalling any emotional engagement with the animals. In between the human domain and the wild domain is a kind of borderland inhabited by dogs and cats that are not entirely wild but present threats, annoyances, or danger to humans, as exemplified by the preference for plural forms such as *barking dogs, rabid dogs, stray cats*, etc. This tripartite division into human domain, a domain between human and wild, and the wild domain is com-

parable to a similar division of the animal kingdom proposed by Donaldson & Kymlicka (2011) into Domesticated, Liminal (i.e., wild, but living in the midst of human settlement) and Wild.
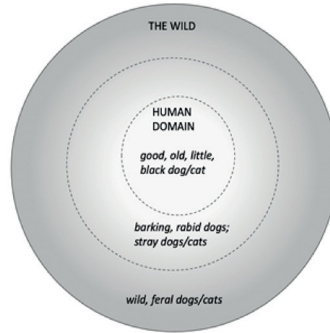

*Figure 2.
Adjective +
inflected form
preferences in the
human domain,
in the wild, and
in between*

6. Conclusion

The study of the singular vs. plural distinction of *cat(s)* and *dog(s)* has revealed collocational patterns that could not be simply predicted on the basis of a one vs. many difference. In the adjective + noun combinations that were investigated, the singular vs. plural distinction is aligned with a human vs. wild dimension where the singular is the preferred form in the human domain and plural is preferred in the wild domain. Along with this preference in number choice, the preferred adjectives accompanying these nouns also contrast in the extent to which they reflect attributes of animals interacting harmoniously with humans, prototypically house-pets, as opposed to animals that live in a world apart from humans. The adjectival distinctions accompanying singular vs. plural nouns reflect conceptualizations of the roles of cats and dogs in society, just as attributive adjectives with MAN and WOMAN are revealing in what the adjectival patterning tells us about social perceptions of gender, as studied in Pearce (2008). The different roles of women and men in society, their biological differences, the different ways each is portrayed in private and public discourse, are all well known and it is not altogether surprising that adjectives collocating with WOMAN and MAN reflect some of these differences. This is not to diminish the value of Pearce's study, which reveals fascinating details about adjectival preferences with respect to appearance and personality traits. But the difference between one dog and many dogs, or one cat and many cats, is not something that occupies our attention in the same way as gender differences do and so the distinctive behaviours of attributive adjectives with DOG and CAT are more surprising. To return to Langacker's distinction of singular vs. plural types, the present study shows that there are indeed some collocational correlates

of the two types with *cat(s)* and *dog(s)*. Singular number represents a "discrete entity" type with preference for adjectival collocates relating to an intimate, human domain, in which house-pets are fondly treated; plural number, or the "replicate mass" type, preferentially occurs with adjectival collocates relating to the wild, set apart from ordinary human experience, and a domain in which individual animals do not enter into close relationships with humans.

It is possible to further refine and nuance the patterns summarized in Figure 2. Within the human domain, one can identify dogs as somewhat more privileged than cats in their degree of interaction with humans, with *good dog*, but not *good cat*, being a preferred combination with the singular form. The preference for *old dog*, rather than *old cat*, could also be seen as reflecting a greater interest in, and empathy for, dogs.

Lexical items such as CAT and DOG are by no means representative of the whole lexicon and I make no generalization about singular vs. plural preferences of English adjectives beyond what I have reported above. But one might well expect that lexical items in some domains would present comparable differences in terms of their preferences to occur with singular vs. plural forms (cf. Newman 2021). Lexical domains of interest would include the world of humans and objects that one interacts with closely. A quick check of some adjective + noun combinations in the Fiction corpus of COCA reveals further potentially interesting items of interest: *sexual partner* (20th most frequent combination with the singular) vs. *sexual partners* (2nd highest combination with the plural); *little flower* (most frequent combination with the singular) vs *little flowers* (10th most frequent combination with the plural); *next-door neighbor* (most frequent combination with the singular) vs. *new neighbors* (most frequent combination with plural). The ranking in terms of raw frequencies is only part of the story and collostructional analysis is needed to properly compare the attraction to the singular vs. plural forms.

Data Availability

All data underlying this research are available at https://osf.io/nc2ba/. The data package includes all (annotated) concordance lines, frequency lists of all adjectives, input files to the DCA, and the R code for the plots in Figure 1 and the DCA.

References

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, Edward Finegan 1999: *Longman Grammar of Spoken and Written English*. Harlow, UK: Pearson Education Limited.

Blouin, David D. 2013: Are dogs children, companions, or just animals? Understanding variations in people's orientations toward animals. *Anthrozoös* 26.2, 279-294.

Davies, Mark 2008-: *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present. URL:* https://www.english-corpora.org/coca/. ED: 15 November. 2021.

Deignan, Alice 2006: The grammar of linguistic metaphors. In: A. Stefanowitsch and S. Th. Gries (eds.) 2006: *Corpus-based Approaches to Metaphor and Metonymy. Berlin/New York: Mouton de Gruyter, 106-122.*

Donaldson, Sue, Will Kymlicka 2011: *Zoopolis: A Political Theory of Animal Rights.* Oxford, UK: Oxford University Press.

Flach, Susanne 2017: *Collostructions: An R Implementation for the Family of Collostructional Methods. Package version v.0.1.0. URL:* https://sfla.ch/collostructions/. ED: 15 November. 2021.

Gries, Stefan Th. 2012: Corpus linguistics: Quantitative methods. In Carol A. Chapelle (ed.) 2012: *The Encyclopedia of Applied Linguistics. Oxford: Wiley-Blackwell, 1380-1385.*

Gries, Stefan Th., Anatol Stefanowitsch 2004: Extending collostructional analysis: A corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics* 9.1, 97-129.

Hunston, Susan 2002: *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Katz, Graham, Roberto Zamparelli 2012: Quantifying count/mass elasticity. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz, Alex Trueman (eds.) 2012: *Proceedings of the 29th West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Proceedings Project, 371-379.

Langacker, Ronald W. 1991: *Foundations of Cognitive Grammar. Vol. II: Descriptive Application.* Stanford, CA.: Stanford University Press.

Mar, Raymond A., Keith Oatley 2008: The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science* 3.3, 173-192.

McEnery, Tony, Richard Xiao, Yukio Tono 2006: *Corpus-based Language Studies: An Advanced Resource Book.* London/New York: Routledge.

Newman, John 2021: *Child and children in a corpus of American fiction: Contrasting semantic preferences and their experiential motivations. Cognitive Semantics 7.1, 1-30.*

Newman, John, Sally Rice 2004: Patterns of usage for English sit, stand, and lie: A cognitively-inspired exploration in corpus linguistics. *Cognitive Linguistics* 15.3, 351-396.

Oatley, Keith 1999: Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology* 3.2, 101-117.

OED. *Oxford English Dictionary Online*, accessed through the University of Alberta Libraries, August 2020.

Pearce, Michael 2008: Investigating the collocational behaviour of *man* and *woman* in the BNC using Sketch Engine. *Corpora* 3.1, 1-29.

R Core Team 2019: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/. ED: 15 November. 2021.

Selby, Lloyd A., John D. Rhoades 1981: Attitudes of the public towards dogs and cats as companion animals. *Journal of Small Animal Pratice* 22.3, 129-137.

Streszczenie

Preferencja liczby pojedynczej lub mnogiej przymiotnikowych kolokatów rzeczowników *cat* [kot] i *dog* [pies]

Liczba gramatyczna angielskich rzeczowników policzalnych nie jest zwykle traktowana jako czynnik istotnie wpływający na wybór przymiotnika, z którym taki rzeczownik występuje. Są przecież przymiotniki takie, jak *big* [wielki], które mogą łączyć się z rzeczownikami w liczbie pojedynczej (*a big dog*) oraz mnogiej (*big dogs*). Niniejszy artykuł bada, czy istnieją interesujące różnice dotyczące preferencji niektórych typów przymiotników co do liczby gramatycznej. W tym celu artykuł bada przymiotniki występujące w pozycji prenominalnej z rzeczownikami *cat/cats* oraz *dog/dogs*. Artykuł jest poniekąd rozwinięciem uwag Langackera (1991: 74-81) na temat podstawowych rodzajów rzeczownika: policzalnych i niepoliczalnych. Według Lagnackera rzeczowniki policzalne I niepoliczalne należą do różnych rodzajów. W przykładzie Langackera, *the seven pebbles* [siedem kamyków] profiluje jedno wystąpienie rodzaju *pebbles* [kamyki], a nie siedem rodzajów *pebble* [kamyk]. Rzeczowniki *pebble* i *pebbles* przedstawiają sobą osobne rodzaje – „odrębna całość" w przypadku liczby pojedynczej i „nieciągła masa" w przypadku liczby mnogiej.

Wybór rzeczowników *cat(s)* i *dog(s)* do analizy motywowany był dwoma względami: po pierwsze, świat psów i kotów ma również ludzki wymiar, dzięki czemu można spodziewać się bogatego zestawu przymiotników

przydawkowych łączących się z tymi rzeczownikami w fikcji literackiej; po drugie zaś, słowa *cat(s)* i *dog(s)* mają dość dużą częstotliwość występowania, co uprawdopodabnia trafność wyników opartych na danych korpusowych.

W analizie przebadano przymiotniki przydawkowe występujące przed rzeczownikami *cat(s)* i *dog(s)* w części korpusu Corpus of Contemporary American English zawierającej współczesną amerykańską beletrystykę. Wybór ten motywowała świadomość, że ilość przymiotników występujących w beletrystyce jest większy, niż w innych gatunkach tekstów. Wstępne wyniki wyszukiwania ograniczono do tych, gdzie desygnatami rzeczowników *cat(s)* i *dog(s)* są 'czworonogie zwierzęta', pomijając wyrażenia typu *hot dog* (przekąska), nazwy własne typu *Black Dog Inn, Dead Dog Creek*, etc., czy przymiotniki poprzedzające nazwy złożone (*a white dog collar, a good dog name*). Trzecia grupa wyników nieuwzględnionych w analizie obejmuje przypadki idiomatycznego użycia badanych rzeczowników, w których to wyrażeniach zastosowanie liczby pojedynczej bądź mnogiej jest często z góry określone, i które nie odnoszą się do żadnych konkretnych psów czy kotów, np. *sly old dog, let sleeping dogs lie*. Aby usprawnić czasochłonną edycję wyników wyszukiwania, do analizy wybrano dziesięć przymiotników najczęściej występujących łącznie z rzeczownikami *cat, cats, dog, dogs*, jako że przymiotniki te odpowiadały za od 30 do 50% wszystkich łączących się z nimi przymiotników.

Tak przygotowane wyniki wyszukiwania poddano analizie koleksemów dystynktywnych, w wyniku której stwierdzono, że spójne grupy przymiotników wykazywały preferencję występowania łącznie z liczbą pojedynczą lub mnogą badanych rzeczowników. Istotne prawidłowości stwierdzone w analizie to: (i) rzeczowniki *dogs* i *cats* w liczbie mnogiej przyciągają przymiotniki związane z dzikimi zwierzętami bądź zwierzętami w warunkach lub stanie uznawanym za niebezpieczny (*feral, stray, wild, rabid, barking*); (ii) rzeczowniki te w liczbie pojedynczej łączą się z przymiotnikami określającymi kolor (*orange, grey, white, black*); (iii) w liczbie pojedynczej rzeczowniki *dog* i *cat* łączą się z przymiotnikami *little, good* i *old*, które sugerują empatyczną orientację ze strony człowieka. Należy zauważyć, że prawidłowość dotycząca łącznego występowania liczby pojedynczej i przymiotników oznaczających zaangażowanie emocjonalne jest spójna ze stwierdzenie Newmana (2021: 17-18), że rzeczownik *child* w liczbie pojedynczej, a nie *children* w liczbie mnogiej, łączy się z przymiotnikami wyrażającymi uczucia (*dear, precious, poor*).

Widoczne w wynikach badania różnice w łącznym użyciu konkretnych przymiotników z liczbą pojedynczą bądź mnogą analizowanych rzeczowników można przypisać ludowemu obrazowi świata, w którym zwierzęta mogą być usytuowane w *domenie człowieka*, w *dziczy*, czy też *gdzieś pomiędzy*. Domena człowieka oznacza świat najbliżej związany z doświadczeniem ludzkim. Należy do niej rodzina, inni członkowie gospodarstwa domowego,

zwierzęta domowe, przyjaciele, czynności związane z miejscem zamieszka-
nia, pracy, rekreacji, itd. W tej domenie występuje preferencja użycia liczby
pojedynczej *dog* i *cat* z przymiotnikami przydawkowymi. Przymiotniki uży-
wane w odniesieniu do tej domeny podkreślają takie cechy psa lub kota, które
sugerują więź emocjonalną ze zwierzęciem. Użycie takich przymiotników
z liczbą pojedynczą rzeczowników *cat* i *dog* odzwierciedla zainteresowanie
i zaangażowanie w życie i dobrostan psa czy kota, mimo że można wyróżnić
pewne różnice między kotami a psami. Przymiotniki określające kolor prze-
ważnie nie są zbyt istotne dla fabuły, ale wzbogacają opis poszczególnego psa
czy kota. Domena dziczy wykazuje preferencję dla liczby mnogiej badanych
rzeczowników, zaś przymiotniki z nimi występujące określają ich dzikość
(*feral dogs*, *stray cats*, itp.) i nie sygnalizują zaangażowania emocjonalnego
w związku ze zwierzęciem. Pomiędzy tymi domenami istnieje swoisty pas
graniczny zamieszkały przez psy i koty, które nie są całkowicie dzikie, ale
stanowią zagrożenie czy też irytują ludzi, co ilustruje widoczna preferencja
dla form liczby mnogiej, jak *barking dogs*, *rabid dogs*, *stray cats*, itp. Ten trójpo-
dział na domenę człowieka, domenę graniczną, i domenę dziczy wykazuje
podobieństwa do podziału królestwa zwierząt na Udomowione, Liminalne
(czyli dzikie, ale zamieszkujące obszary zasiedlone przez człowkeka) i Dzikie
proponowanego przez Donaldson & Kymlickę (2011).

W nawiązaniu do Langackerowskiego zróżnicowania rodzajów na poje-
dyncze i mnogie, niniejsza analiza wskazuje, że istnieją pewne kolokacyjne
korelaty tych dwóch typów w przypadku rzeczowników *cat(s)* i *dog(s)*. Liczba
pojedyncza odpowiada rodzajowi „odrębna całość" i preferuje kolokaty przy-
miotnikowe związane z bliską nam domeną człowieka, gdzie zwierzęta do-
mowe są hołubione. Z kolei liczba mnoga, czy też „nieciągła masa", preferuje
kolokaty przymiotnikowe związane z dziczą, nie związane z codziennym
doświadczeniem człowieka, ale wskazujące na domenę, gdzie poszczególne
zwierzęta nie tworzą bliskich związków z człowiekiem.

Bibliografia

Donaldson, Sue, Will Kymlicka 2011: *Zoopolis: A Political Theory of Animal Rights.*
        Oxford, UK: Oxford University Press.
Langacker, Ronald W. 1991: *Foundations of Cognitive Grammar. Vol. II: Descrip-
        tive Application.* Stanford, CA.: Stanford University Press.
Newman, John 2021: *Child* and *children in* a corpus of American fiction: Con-
        trasting semantic preferences and their experiential motivations. *Cog-
        nitive Semantics* 7.1, 1-30.

Tłum. Małgorzata Paprota